

Running head: NUMERACY REVIEW

Numeracy: A Review of the Literature

Nathan Dieckmann

Decision Research & University of Oregon

Decision Research Report No. 08-2

May, 2008

Abstract

Numeracy is defined as the ability to understand and use numbers. In addition to basic reading and writing skills, today's consumers need an understanding of numbers and basic mathematical skills to use any numerical information presented in text, tables or charts. This is especially true in many financial and healthcare settings, where a basic understanding of numerical concepts is arguably as important for informed decision making as reading ability. Although systematic research on numeracy has been growing steadily over the last several years, there have been few comprehensive reviews of this literature. In addition to summarizing key findings, a review could serve to identify gaps in our knowledge and suggest paths for future research in the field. The primary goal of this paper is to review current directions in numeracy research, and, in particular, to examine the relationship between numeracy and decision making in health and selected non-health domains.

In recent years, there has been a drastic shift in the way consumers are expected to make important health and financial decisions. Ever increasing amounts of information are made available to the public, with the expectation that consumers will use this information for decision making. For example, employees may be presented with several retirement account plans and presumed capable of making informed decisions about the optimal plan for their goals (Thaler & Sunstein, 2003). Current policy approaches toward Medicare beneficiaries serve as another example of this shift. Here, patients are expected to utilize information about the various Medicare drug options, make trade-offs, and ultimately produce wise choices regarding their care (Hibbard, Slovic, Peters, Finucane, & Tusler, 2001; Woloshin, Schwartz, & Welch, 2004). This change is consistent with a move toward a more consumer-driven approach to healthcare, in which consumers share in both healthcare decision-making and the healthcare associated cost (Hibbard & Peters, 2003).

The information provided to consumers to facilitate these choices come from numerous sources and are presented in a variety of formats. This can be confusing and daunting for even for the most motivated and skilled consumer. Financial and healthcare providers, while generally supplying sufficient information for consumers, often overlook the importance of presenting information in a way that can be easily comprehended and used by the vast majority of consumers (Fagerlin et al., 2004). With the increasingly important decisions being left to the consumer, it is critical that information providers are sensitive to the skills of the consumers they are trying to reach.

Researchers have long recognized the importance of literacy for informed decision making (see Rudd, Colton, & Schacht, 2000). A great deal of our information comes in textual form, particularly in light of today's reliance on the Internet for disseminating information. Thus, individuals with limited literacy skills are at a marked disadvantage in this information age. Findings suggest that the low literate make poorer decisions and have worse financial and healthcare outcomes (Baker, Parker, Williams, Clark, & Nurss, 1997).

Often overlooked has been the pivotal role of numeracy in decision making. Numeracy is defined as the ability to understand and use numbers. In addition to basic reading and writing skills, today's consumers need an understanding of numbers and basic mathematical skills to use any numerical information presented in text, tables or charts. This is especially true in many financial and healthcare settings, where a basic understanding of numerical concepts is arguably as important for informed decision making as reading ability.

An estimate from the National Adult Literacy Survey (NALS; 1992) suggests that approximately half of the US population has only very basic or below basic quantitative skills (Kirsch, Jungeblut, Jenkins, & Kolstad, 2002). More specifically, half of the US population had difficulty identifying and integrating numbers in a lengthy text, or performing a numerical task that required two or more sequential steps to reach a solution. The National Assessment of Adult Literacy (NAAL, 2003; see Kutner, Greenberg, Jin, & Paulsen, 2006) also demonstrated similarly poor numerical skills among the US population, although there was a slight decrease in the percentage of individuals with below basic quantitative skills (from 26% in 1992 to 22% in 2003; National Center for Education Statistics, 2007). These findings are particularly worrisome considering the increasing importance of numerical skills in the contemporary workplace (Parsons & Bynner, 1997).

It is not just the general population that has difficulty with numerical tasks. Several authors have found that even highly educated participants have an inadequate understanding of probabilities, risks and other chance related concepts (Estrada, Barnes, Collins, & Byrd, 1999; Lipkus, Samsa, & Rimer, 2001; Sheridan & Pignone, 2002). For instance, physicians generally believe it is important to provide quantitative risk estimates to their patients. However, when surveyed, they actually felt more comfortable providing verbal estimates of risk, presumably because of a lack of confidence and knowledge concerning the quantitative risk estimates (Gramling, Irvin, Nash, Sciamanna, & Culpepper, 2004).

With the recognition of the importance of numeracy in financial and healthcare settings has come a growth in research efforts dedicated to the measurement of numeracy and the effects of low numeracy on judgment and decision making. This review will take a broad look at current interest and research findings concerning numeracy.

Scope and Goals of the Review

Although systematic research on numeracy has been growing steadily over the last several years, there has not been a published review of this literature. In addition to summarizing key findings, a review could serve to identify gaps in our knowledge and suggest paths for future research in the field.

The primary goal of this paper is to review current directions in numeracy research, and, in particular, to examine the relationship between numeracy and decision making in health and selected non-health domains. In the first section of this review we consider the measurement of numeracy, and also detail the specific conceptualization of numeracy that will be the focus of the remainder of the paper. In the next section we discuss research on the relationship between

numeracy and decision making in health and non-health domains. Then, we review selected research from the cognitive/developmental literature that can further our understanding of numeracy. Finally, we summarize the current state of knowledge concerning numeracy and discuss possible future directions for the field.

Method

Literature Search

Several methods were used to search the literature for potentially relevant research reports. Electronic databases (e.g. PsychInfo, Medline) were used to capture an initial set of potentially relevant research reports. The initial search terms were relatively broad (e.g. Numeracy, Numeric ability, Number ability, etc) resulting in a large number of potential reports. We scanned the abstracts of all of the papers identified in the electronic databases for inclusion in the review. After the initial search, the Web of Science database was used to identify additional reports that had referenced many of the pivotal numeracy papers. Finally, the reference lists of all papers identified by the first two methods were examined for additional papers that were missed by the electronic searches.

Study inclusion

We focused primarily on empirical reports published in peer-reviewed journals or published books. Dissertations and master's theses were excluded from the review. We also excluded papers that reported single-case studies, qualitative studies, and papers with very small sample sizes (e.g. results from interviews with 2 or 3 participants). A few unpublished working papers or other reports were included, but we did not make a specific effort to retrieve unpublished

literature. We think the decision to not specifically seek unpublished reports is justified since the primary purpose of this review was to get a broad sense of our current knowledge concerning numeracy, and to propose directions for further research. This decision does, however, leave open the potential problem of publication bias in the review (also called the file-drawer problem; Rosenthal, 1979).

Narrative review process

We discuss the results of the research reports in a qualitative, narrative fashion, primarily because of the breadth and primary goals of this review. In many situations, a quantitative review using meta-analytic methods can be very informative and produce results not attainable with narrative reviewing methods (see Beaman, 1991; Cook & Leviton, 1980; for additional information on meta-analysis see Cooper & Hedges, 1994). Particularly as the research literature on numeracy grows, future reviewers should consider synthesizing this literature with meta-analytic techniques.

Literature review

Conceptualizing and measuring numeracy

Numeracy defined in the broadest sense is the ability to understand and use numbers. This would include an understanding of the real number line, the ability to compare numbers in magnitude, the understanding of time and money, measurement, estimation, and the ability to perform simple arithmetic. At a somewhat higher level, a broad definition might also include basic logic, performing multi-step operations, a fundamental understanding of chance and basic

statistical principles, and comfort with proportions, fractions, probabilities, and risks.

Researchers have defined and measured numeracy in various ways, often because of differences in their specific research interests and domain of study. For example, Paulos (1988) defines innumeracy as the “inability to deal comfortably with the fundamental notions of number and chance” (pg. 3). He discusses difficulties individuals have in understanding extremely large and small numbers, grasping infinity, correctly using combinations and permutations to calculate quantities, and understanding basic concepts involving chance and probability. Another example comes from the healthcare domain, where researchers are often interested in the ability of the public to understand the risks and benefits of particular medical treatments. These authors often define numeracy as the ability to understand proportions, risks, percentages and probabilities, since these are the forms in which risk and benefit information is most often presented to consumers (Burkell, 2004). Still other researchers have focused on the ability of people to use numbers embedded in text documents and to perform simple arithmetic operations (e.g. Kirsch et al., 2002). Below we will discuss the different ways in which researchers have operationalized and measured numeracy.

Health literacy is often defined as a set of basic skills that individuals must possess to function in today’s healthcare system. The skill set includes both basic reading and numerical skills, although there is some disagreement among researchers about the precise definition of health literacy (Baker, 2006; Speros, 2005). The Test of Functional Health Literacy in Adults (TOFHLA) consists of two components: a 50-item reading comprehension section and a 17-item numerical ability section (Parker, Baker, Williams, & Nurss, 1995). In the numerical section, participants are presented with actual hospital forms and prescription vials and are asked questions that test the ability to “comprehend directions for taking medicines, monitoring blood

glucose, keeping clinic appointments, and obtaining financial assistance.” (Parker, 1995, pg. 538). Another example is the health literacy component of the National Assessment of Adult Literacy (NAAL; National Center for Education Statistics, 2007). The items of the numerical component of health literacy assess the ability to explain and/or perform arithmetic operations on information embedded in a document (i.e. addition, subtraction, multiplication, division).

There have also been attempts to develop shorter measures of health literacy that are more easily administered in real world clinical settings (e.g. Baker, Williams, Parker, Gazmararian, & Nurss, 1999). One example is The Newest Vital Sign (NVS) measure, which consists of 6 questions asking about the information presented in an ice cream nutrition label (Weiss et al. 2005).

Since health literacy involves much more than just numerical skills, we will not present a comprehensive review of this literature. In short, however, research has found that people with poorer health literacy are more likely to use hospital and emergency department services (Baker et al., 2002; Baker et al., 2004), have more physical and mental health problems (Wolf, Gazmararian, and Baker, 2005), do not receive as much preventative healthcare (Scott et al., 2002), and are less knowledgeable about health (Gazmararian et al., 2003; Williams et al., 1998a; Williams et al., 1998b). For more literature concerning health literacy see Rudd et al. (2000).

Much of US population is currently suffering from a poor rate of personal savings, increasing consumer debt, and an increasing number of bankruptcies. Consequently, researchers have begun to focus on the financial literacy of the population (Fox, Bartholomae, & Lee, 2005). Financial literacy refers to the skills and knowledge necessary to make wise financial decisions. For example, the financially literate are able to balance a checkbook, understand and use interest

rates to their advantage (e.g. with a credit card), and purchase appropriate insurance (Beverly & Burkhalter, 2005). Financial literacy includes both basic numerical skills and knowledge about the workings of the financial world, and in this sense it is not a strict measure of numeracy.

Measures of financial literacy include questions about personal savings rates, use of retirement accounts, facts about interest rates/sales tax and budgeting, and paying bills on time (Hilgert, 2003; Lucey, 2005; Perry & Morris, 2005). In short, researchers have found that people with higher levels of financial literacy are better able to manage credit cards, better able to invest and save (Hilgert et al, 2003), and have greater levels of wealth (Lusardi & Mitchell, 2006).

Both health and financial literacy are relatively broad constructs that involve a range of skills including reading ability, domain specific knowledge, and numerical skills. Many researchers, however, have developed more specific measures that focus on numerical skills. In the health domain, several authors have argued that the numeracy aspect of health literacy has been overshadowed by the literacy component (Golbeck, Ahlers-Schmidt, Paschal, & Dismuke, 2005; Montori & Rothman, 2005). In an attempt to clearly distinguish health numeracy from the larger concept of health literacy, Goldbeck et al. (2005) classify different numerical abilities into four functional categories: Basic (the ability to identify and understand numbers), Computational (the ability to apply simple manipulations to numbers, e.g. addition, averages, etc), Analytical (making sense of information by using inference, percentages, frequencies, etc), and Statistical (basic skills to understand probabilities, life expectancy, risk, and comparing information from different scales). These categories can serve as a useful structuring device when thinking about the skills that are tested by a particular numeracy measure. Below we discuss the methods that researchers have used to specifically measure numerical skills.

Much of the quantitative health information presented to the public involves communicating the risks associated with particular diseases or treatment options (Burkell, 2004). For example, patients may need to understand their chances of acquiring a post-op infection following different types of surgery, the overall survival rates for different treatments, or they may need to compare hospitals or physicians in terms of specific quality indicators. Most of this information comes in the form of explicit probabilities, relative frequencies, and proportions, and it is assumed that people can interpret these different measures to make an assessment of the likelihood of different outcomes. These relatively higher-level skills fall into the analytical and statistical categories delineated by Goldbeck et al. (2005). Because of the importance and common use of this type of outcome likelihood information, researchers have developed numeracy measures specifically designed to assess these skills.

Black et al. (1995) used one of the first crude measures of numerical ability that assessed the understanding of simple chance processes. Respondents were asked a relatively simple question about the number of times a fair coin would come up heads in 1000 tosses. The authors considered a participant numerate if they were able to answer this question correctly and they provided logically consistent responses to other questions about probability (e.g. estimating the probability of acquiring a disease as being greater than the probability of acquiring the disease and dying from the disease).

Schwartz, Woloshin, & Welch (1997) measured numeracy with three questions, which included a basic question assessing participants understanding of chance (i.e. how many heads would come up in 1000 tosses of a fair coin) and two questions asking the participants to convert a percentage to a proportion and a proportion to a percentage (i.e. the chance of winning a car is 1 in 1000, what is the percentage of winning tickets for the lottery?). This measure proved

popular among researchers, and several authors have developed expanded versions of the original 3-item measure. For example, another measure was developed by using the initial three items from the Schwartz et al. measure along with three additional questions. The additional questions assessed the ability of a patient to calculate the number of pills needed to equal a prescribed dose of a medication (Estrada et al., 1999; Estrada, Martin-Hryniewicz, Peek, Collins, & Byrd, 2004). Other researchers also adapted this numeracy measure by asking participants to calculate a tip on a dinner bill and to calculate one third of 100 (Parrott, Silk, Dorgan, Condit, & Harris, 2005).

Another important addition to the literature was the expanded numeracy measure created by Lipkus et al. (2001). They added eight questions to the items from the Schwartz et al. numeracy scale. The additional items were designed to assess a participant's ability to understand and compare risks (e.g. Which of the following numbers represents the biggest risk of getting a disease: 1%, 10%, or 5%?) and to move between decimal representations, proportions and fractions. Peters et al. (in press) have also used an expanded version of the Lipkus numeracy scale, introducing four more difficult items. Among other things, these additional items test the understanding of base rates, as well as the ability to make more complex likelihood calculations.

In addition to measuring numeracy by performance on quantitative tests, a subjective numeracy test has also been developed (Fagerlin et al., submitted). Instead of having participants solve numerical questions they are asked directly about their abilities and preferences in using numbers. The test has 8-items that measure perceived skill in working with numbers and preferences for information presentation format. The main advantages of the subjective numeracy test are the shorter administration time and the less stressful experience for the participant (Fagerlin et al., submitted).

Another interesting measure is the medical data interpretation test, which is specifically designed to measure the ability to compare numbers (Schwartz, Woloshin, & Welch, 2005). This measure “goes beyond testing simple literacy and numeracy and examines the ability to compare risks and put risk estimates in context (i.e. to see how specific data fit into broader health concerns and to know what additional information is necessary to give a medical statistic meaning” (pg. 290-291). This measure is interesting because it is testing particular comparative and inferential skills that are not tested by the measures discussed previously.

There are several additional measures of numeracy that have been developed in addition to the most frequently used numeracy measures discussed above. Weinfurt et al. (2003; 2005) used a single question to assess numeracy by asking participants the meaning of the statement “40% of cases like yours” in the context of a physician’s prognosis. Several measures have also been developed specifically for different types of patients. Apter (2006) developed a numeracy measure designed for Asthma patients. The four items were designed to assess a patient’s ability to calculate the number of pills to take daily, and their ability to interpret percentages and the risks of particular therapy related values. Other researchers have developed a numeracy measure specifically for diabetic patients (Montori, Leung, Thompson, Chung, Capes, & Smith, 2004), which measures a patient’s ability to determine whether a glucose value is at goal or not, count carbohydrates from a food label, and adjust insulin dose by a given percentage. In addition, Woloshin, Schwartz, & Welch (2005) have recently developed two scales measuring interest and confidence in using medical statistics. These measures are similar to the subjective numeracy test in that they measure perceptions and attitudes as opposed to performance on numerical problems.

Numeracy has been defined and measured in several different ways, both as a single construct and as a subcomponent of broader measures of functioning. When reading the literature on numeracy it is important to take note of the specific skills that are being tested by a particular numeracy measure. As Goldbeck et al. (2005) discuss, there is a hierarchy of numerical skills that affect reasoning and decision making behavior. In our review of literature, we focus on studies that conceptualize numeracy at a relatively high level – namely, the ability to understand and use probabilities, proportions, risks, and other notions of chance in decision making settings. This is the conceptualization of numerical skills that has been most often studied in the health domain. Note, however, that we do review the work of several researchers that conceptualize numeracy in different ways.

Numeracy findings in health and related domains

In this section we will review the evidence relating numeracy to judgment and decision making behavior.

Health and employment/financial outcomes.

Low numerical ability has been linked with poorer health, financial and employment outcomes. In one study, Estrada et al. (2004) assessed the relationship between literacy, numeracy and the ability of patients to follow dosing schedules for anticoagulation control. The study sample consisted of patients (n=143) attending an outpatient clinic for anticoagulation management. The authors measured numeracy with an expanded version of the Schwartz et al. (1997) scale, and also measured the amount of time that each patient spent in an appropriate therapeutic range. They found that those patients that were lower in numeracy had greater

variability in their measured control and spent larger amounts of time above their appropriate therapeutic range.

Parsons & Bynner (1997) used data from the National Child Development Study (NCDS) to assess the relationship between numerical skills and employment. They restricted their analysis to those people that were 37 years of age with good literacy skills, and compared the employment records of those high and low in numerical ability (n=1,714). The numeracy assessment tested relatively basic skills like calculating the change from a purchase, calculating the square feet in a room, and reading values from a graph. They found that for a variety of measures of employment, those lower in numeracy had fewer job opportunities and did not progress in jobs as much as high numerate participants. For example, unemployment was almost three times as likely for low numerate men and twice as likely for low numerate women.

Other researchers have investigated financial outcomes among the lower numerate. One group used the propensity to give focal answers to probability questions (i.e. the tendency to respond 0%, 50% or 100% when asked about the probability of an event) as a proxy for numeracy¹. They found that participants that tended to give focal answers to probability questions had lower overall wealth (Lillard & Willis, 2001). In a related finding from the English Longitudinal Study of Aging, participants with lower numerical skills were found to be less likely to hold more complex assets and pensions and had lower overall wealth (Banks & Oldfield, 2006). In this study they measured numerical ability with a six-item test that assessed the ability to calculate change from a purchase, convert a percentage into a frequency out of 1000, divide a quantity of money among 5 people, and calculate the earnings from the interest in a savings account.

¹ See Bruine de Bruin et al. (2000) for a discussion of number ability and the propensity to give focal responses.

In summary, these findings strongly suggest that lower numeracy is related to poorer health and worse financial and employment outcomes. In the next sections we review research that focuses on the relationship between numeracy and specific judgment and decision making behavior.

Perceptions of Risk and Benefit.

An understanding of the risks and benefits associated with particular choice options is important for many health and financial decisions. For example, many patients are expected to understand and weigh the risks and benefits of various treatment options shortly after being diagnosed with an illness. Employees are also expected to understand the risks and benefits of various investment options when planning for retirement. Below we review literature that suggests people lower in numerical ability have consistent biases in their perceptions of risk and benefit.

Many of the research studies examining risk perceptions were conducted in the context of breast cancer research. Black et al. (1995) asked women between the ages of 40-50 (n=145) several questions about the probability that they would develop or die of breast cancer in the next 10 years. They measured numeracy with a single question (i.e. the # of times a fair coin would come up heads in 1000 tosses). The entire sample overestimated their personal risk of breast cancer as compared to epidemiological data, and those lower in numeracy made even larger overestimations than those higher in numeracy.

Schwartz et al. (1997) also asked women (N=287) to estimate the risk of dying from breast cancer both with and without mammography screening. They presented the women with risk reduction information (due to mammography) in four different formats and calculated accuracy

by how well they adjusted their risk estimates in light of the new information. After adjusting for age, income, level of education, and the format of the information, participants higher in numeracy were better able to use the risk reduction data.

In another study, Woloshin and colleagues asked women (n=201) to estimate their 10-year risk of dying from breast cancer as a frequency out of 1000 (Woloshin, Schwartz, Black, & Welch, 1999). In addition, they asked the women to estimate how their risk compared to an average woman their age. Numeracy was measured with the 3-item scale used by Schwartz et al. (1997). Numeracy was not related to participant's comparison judgments, but participants lower in numeracy overestimated their risk of dying from breast cancer in the next 10 years. This study showed that participants lower in numeracy might still be able to make accurate risk comparisons, even though they are not able to make unbiased risk estimates.

Another group of researchers have also conducted studies in the breast cancer domain. In one study, a sample of women estimated their 5-year and lifetime risk of breast cancer and completed a numeracy scale adapted from Schwartz et al. (Davids, Schapira, McAuliffe, & Nattinger, 2004). Similar to the findings above, participants (n=254) as a whole overestimated their risk of breast cancer (as compared to epidemiological data), with those lower in numeracy making larger errors in their estimates than those higher in numeracy. In a separate report, these authors also showed that numeracy was related to the consistent use of frequency and percentage risk rating scales (Schapira, Davids, McAuliffe, & Nattinger, 2004). After controlling for age, health literacy, race and income only higher numeracy was shown to be predictive of using the percentage and frequency scales in a consistent manner (i.e. giving the same responses on both scales for the 5-year and lifetime breast cancer risk estimates).

There have also been a few studies that have not found a relationship between numeracy and breast cancer risk estimates. Dillard, McCaul, Kelso, & Klein (2006) investigated whether poor numeracy skills could account for the finding that women consistently overestimate their risk of breast cancer even after receiving epidemiological information about the risk. Numeracy was measured with the 3-item scale used by Schwartz et al. (1997). In this study (n=62), numeracy was not related to the participant's persistent overestimation of their risk. Another group of researchers asked a sample of Black and White women (n=207) to estimate their 5-year survival after a diagnosis of breast cancer, along with an estimate of the relative risk reduction due to screening mammography (Haggstrom & Schapira, 2006). Numeracy was measured with the 3-item scale used by Schwartz et al. (1997). They found no effect of numeracy after controlling for other demographic variables (i.e. race, family history of breast cancer, income, insurance type and level of education).

There have also been investigations of risks and benefits outside of the breast cancer domain. In a large survey of cancer patients (n=328), Weinfurt et al. (2003) asked participants to estimate the chance that they would be benefited by an experimental cancer treatment. Numeracy was measured with a single question (what does the following statement mean "40% of cases like yours"). Patients that did not correctly answer the numeracy question perceived a greater chance of benefit from the experimental therapy. In another study, participants were presented with several hypothetical scenarios that described a physician's estimate of the risk that a patient had cancer (Gurmankin, Baron, & Armstrong, 2004b). The authors asked participants to imagine that they were the patient described and to rate their risk of cancer. Numeracy was measured with a scale adapted from Lipkus et al. (2001). They found that patients lower in numeracy were more likely to overestimate their risk of cancer. In addition, Berger (2002) presented participants

(n=223) with brief news stories describing an increase in burglaries in the city in which the participants lived. The descriptions were supplemented with frequency information showing the burglary rates for the last eight years. Participants that scored higher on the numeracy test were less apprehensive about the increase in burglaries.

In conclusion, several studies have found that numerical ability is related to perceptions of risk and benefit. Participants lower in numerical ability tend to overestimate the risk of cancer and are less able to use risk reduction information to adjust their risk estimates. However, the results from Woloshin et al. (1999) suggest that the form of the risk question may be part of the problem. In other words, participants low in numerical ability might have trouble expressing their risk estimates on the scales generally used in this domain. Related to this is the finding that the low numerate have difficulty using frequency and percentage risk scales in a consistent manner. Future research should focus on disentangling response scale effects from the conceptual difficulties that the low numerate have with risk and benefit information.

Eliciting value judgments (utilities).

Much research has focused on measuring the value that patients place on different health states and health outcomes. Getting reliable and valid assessments of how patients value different health states is important for both modeling their decision making activity as well as optimizing healthcare delivery. The two primary methods for eliciting these utilities are the standard gamble and time-trade-off methods (Woloshin, Schwartz, Moncur, Gabriel, & Tosteson, 2001). Both of these methods require a participant to make repeated choices between different hypothetical health states until they are indifferent between options—namely, a choice in which they do not favor one health state over the other. For example, imagine that a patient

with health problems has a life expectancy of 20 years. The patient is faced with a series of choices between living in their current state of health for the remainder of their life, or living with perfect health for some shorter amount of time. A utility for the patient's current health state can be calculated at the point at which the patient finds both choices equally attractive (e.g. remainder of life with current health versus 10 years of perfect health = $10/20$... utility = .5). The methods used to elicit utilities from patients often require them to deal with probabilities and/or make trade-offs between states. Because of the quantitative nature of the task, some researchers have questioned the validity of the standard approach to eliciting utilities and have begun to look at the numerical abilities of patients.

Woloshin et al. (2001) used three different methods for eliciting utilities about the current health of a sample of women (n=96). The participants completed a standard gamble, a time trade-off, as well as a task in which they rated their current health on an analog scale. The authors measured numeracy with the same measure used by Schwartz et al. (1997). Low numerate participants showed a negative correlation between a question asking about current health and the utilities generated from the standard gamble and time trade-off tasks, indicating that these participants had difficulty with the utility elicitation tasks. The high numerate showed the expected positive correlation, and, interestingly, all participants showed a positive correlation with the analog rating scale. Similar studies have been conducted with head and neck cancer patients (Schwartz, McDowell, & Yueh, 2004). These findings also suggest that the standard methods for assessing utility may be biased in patients with limited numerical ability. In addition, similar conclusions were also reached using the subjective numeracy measure (Zikmund-Fisher, Smith, Ubel, & Fagerlin, in press).

These research studies suggest that patients lower in numeracy have difficulty with the standard procedures for assessing utilities. It is clear that low numerate patients have difficulties dealing with probabilities, but they also appear to have trouble making trade-offs. Making trade-offs and comparisons between hypothetical states involves additional reasoning skills not necessary when simply dealing with probabilities. Future work should investigate the interplay between the different skills that are needed to complete these tasks, which could lead to new methods of eliciting utilities that could be used with patients at all levels of numerical ability.

Information Presentation and Formatting.

Many researchers have tried to identify optimal methods of presenting numerical information to consumers (e.g. Journal of the National Cancer Institute Monographs, 1999; Fagerlin, Ubel, Smith, & Zikmund-Fisher, submitted). However, very little research has focused on how people with varying levels of numerical ability understand and use information presented in different formats. Fagerlin et al. (submitted) review some of the literature on presenting risk and benefit information, but for the most part can only speculate about how individuals varying in numerical ability would deal with different presentation methods. Below we discuss what is known about the relationship between information presentation and numeracy.

Peters et al. (2006) conducted several experiments that examined how individuals varying in numerical ability were able to understand and use probabilities expressed in different formats and to what extent they were affected by information framing.

Framing effects have proven to be a relatively robust phenomenon in psychological research (e.g. Kuehberger, 1998). For example, presenting the risk of a surgical operation as an 80% chance of survival versus a 20% chance of death will most likely change a patient's perceptions

of the surgery. There have been some previous attempts to examine the relationship between quantitative skills and framing effects, although these studies used more indirect methods of assessing numeracy. In one study conducted by Stanovich and West (1998), participants with lower total SAT scores were found to be more likely to show a framing effect for risky choices. In more recent study, Peters et al. (2006) conducted a single-attribute framing experiment and had participants (n=100) complete the numeracy measure developed by Lipkus et al. (2001). Participants were presented with a set of 5 exam scores belonging to different students. Half of the subjects received the negative frame (e.g. 20% incorrect), and the other half received the positive frame (e.g. 80% correct). Participants then rated the quality of each student's work. The authors found that the less numerate participants showed larger framing effects.

In a second study, Peters et al. (2006) examined whether numerical ability affected the perception of probability information. Participants (n=46) were asked to rate the risk associated with releasing a hypothetical mental health patient. One half of the participants read the scenario in the frequency form ("Of every 100 patients similar to Mr. Jones, 10 are estimated to commit an act of violence to others during the first several months after discharge") and the other half received the same information in percentage form ("Of every 100 patients similar to Mr. Jones, 10% are estimated to commit an act of violence to others during the first several months after discharge"). High numerate participants did not differ in their risk ratings between the two formats. Low numerate participants, however, rated Mr. Jones as being less of a risk when they were presented with the percentage format. The authors speculate that the low numerate, because of their limited numerical skills, have more difficulty transforming one representation to another ($10/100 = 10\%$), and were therefore differently affected by the format. The low numerate may have reported a higher levels of risk with the frequency format because in this

condition they generated more vivid images of the violent acts than in the percentage condition (Slovic, Finucane, Peters, & MacGregor, 2004).

A recent paper by Peters et al. (in press) also explored the relationship between numeracy and the format of numerical information. In each experiment, participants (n=303) were presented with different measures of hospital quality and asked to make an informed hospital choice. Numeracy was measured with an expanded version of the scale used by Lipkus et al. (2001). In the first study, participants saw a number of hospital quality indicators (e.g. % of time guidelines for heart attack care are followed), as well as a set of non-quality information (e.g. # of visiting hours per day) for three different hospitals. Information about cost was also provided. Participants were asked to choose a hospital that they would like to go to if they needed care, and they also responded to a series of questions about the information presented (e.g. Which hospital is most expensive for you?). In the first condition, all of the information was displayed in an unordered fashion. In the second condition, the same information was ordered (cost and quality information first) and the cost and quality information was highlighted. In the third condition, the non-quality information was deleted and only the cost and quality information was displayed. The authors found that the low numerate participants were helped significantly more by the altered presentation formats than the high numerate. The low numerate showed better comprehension in the ordered and quality-information-only conditions as compared to the unordered condition. The high numerate only showed better comprehension in the quality-information-only condition compared to the unordered condition.

In study 2, participants were told that they needed treatment for heart failure and were asked to choose between 15 different hospitals based on three pieces of information: cost, overall patient satisfaction, and death rate for heart failure patients. The formatting of the information

was varied by including black and white symbols or colored traffic light symbols to help participants better evaluate the goodness or badness of each piece of information. The results indicated that low and high numerate participants responded differently in the presence of the easier to evaluate symbols. The low numerate made more higher quality choices with the black and white symbols as compared to the colored traffic light symbols, while the reverse was true for the high numerate participants.

In a third hospital choice study, the authors found that low numerate participants were particularly sensitive to the verbal framing of the information. Low numerate participants showed greater comprehension when information was presented such that a higher number means better (the number of registered nurses per 100 patients) as compared to when a lower number means better (the number of registered nurses per 100 patients). Higher numerate participants were better able to deal with the information in both formats. Overall, these studies suggest that there are benefits to reducing the cognitive burden and increasing the salience of the important information in a display. This was particularly true for the low numerate participants.

Researchers have also investigated the relationship between numeracy and the ability to read graphs. Zikmund-Fisher et al. (in press) presented participants (n=155) with a survival graph that depicted the number of people given two different drugs that would be alive over a 50-year period. They then asked four questions about information displayed in the graph (e.g. What year is the difference in total survival between Pill A and Pill B was the largest?). They measured numeracy with the subjective numeracy measure. The ability to correctly interpret the survival graphs was strongly related to numeracy, with those higher in numeracy being better able to interpret the graphs.

In another study, researchers focused on trust and confidence in numerical information. Gurmankin, Baron, & Armstrong (2004a) conducted a web survey in which they presented subjects (n=115) with several hypothetical risk scenarios. The scenarios depicted a physician presenting an estimate of the risk that a patient had cancer in three different formats (verbal, numerical probability as percentage or numerical probability as fraction). Participants then rated their trust and comfort with the information, as well as whether they thought the physician distorted the level of risk. Numeracy was measured with a scale adapted from Lipkus et al. (2001). Overall, they found that participants were more trusting of the information in the numeric as compared to the verbal formats, although this effect interacted with numeracy. Even after adjusting for gender, age, and education, the results showed that those subjects with the lowest numeracy scores trusted the information in the verbal format more than the numeric, and those with the highest numeracy scores trusted the information in the numeric formats more than the verbal.

Sheridan and colleagues conducted two studies in which they assessed the relationship between numeracy and ability of participants to interpret risk-reduction information (Sheridan & Pignone, 2002; Sheridan, Pignone, & Lewis, 2003). Participants in both studies were presented with baseline risk information about a disease and then given risk-reduction information in one of four formats (relative risk reduction, absolute risk reduction, number needed to treat, or a combination of all methods). They measured numeracy with a measure adapted from Schwartz et al. (1997). In a sample of first year medical students (n=62) and a sample of patients from an internal medicine clinic (n=357), the authors found that participants lower in numeracy had more difficulty using the risk reduction information, and in particular, had trouble with the “number needed to treat” format.

Finally, Parrott et al. (2005) presented participants (n=206) with statistical evidence concerning the relationship between a particular gene and levels of LDL cholesterol. The statistical information was presented either in a strictly verbal form with percentage information, or a visual form that showed a bar graph of the mortality rates. They were interested in whether perceptions of the evidence differed between the formats, as well as whether numerical skills were related to these perceptions. They measured numeracy with four items adapted from the measure used by Schwartz et al. (1997). They did not find any relationships between numeracy and comprehension, perceptions of the quality of the evidence, or perceptions of the persuasiveness of the evidence. The authors do note, however, that the restricted range of numerical abilities may have contributed to the null effect.

In summary, low numerate participants tend to be worse at reading survival graphs, more susceptible to framing effects, more sensitive to the formatting of probability and risk reduction information, and tend to trust verbal more than numerical information. The low numerate can be helped by presenting information in a logically ordered format and displaying only the important information, presumably decreasing the cognitive burden on the participants. Additional research is needed to further elucidate the presentation formats that are most beneficial for individuals at all levels of numerical ability. Fagerlin et al. (submitted) discuss several different presentation issues that have not been investigated with relation to numerical ability. This could be a starting point for further research focusing on the effects of presentation format.

Numeracy and other decision making tasks.

There have also been some recent investigations focusing on numerical ability and economic decision making behavior. For example, Benjamin, Brown and Shapiro (2006) conducted

several experiments with Chilean High School students. In a series of tasks they assessed the students with respect to small-stakes risk aversion and short-run discounting, two preference anomalies that are thought to affect a wide range of personal economic activity. They used the scores on standardized tests (similar to the SAT) as a proxy for general cognitive ability. They found that those students with higher test scores, particularly higher mathematical ability, were less likely to show small-stakes risk aversion and short-run discounting.

It is possible that performance on many reasoning and decision making tasks are moderated by the numerical ability of the participants. Future research within behavioral economics and judgment and decision making should use more specific measures of numeracy and investigate the effect of numeracy in a wider array of reasoning and decision making tasks.

Numeracy and Affect.

In the final section we discuss recent research that connects numerical ability and affective processing. Peters and colleagues conducted two experiments that examined whether numerical ability was related to affective evaluations of numbers (Peters, et al., 2006). In one study, they used the jellybean task developed by Denes-Raj and Epstein (1994). Participants (n=46) were presented with two hypothetical bowls of jellybeans and were told that they would win \$5 if they picked a colored jellybean. One bowl had a total of 100 jellybeans with 9 colored jellybeans. This bowl was labeled “9% colored jellybeans”. The second bowl had a total of 10 jellybeans with only one colored jellybean and was labeled “10% colored jellybeans”. Participants were then asked to choose the bowl they would like to pick from. They also rated how they felt about the 9% chance associated with the first bowl (affect question), rated the precision of that feeling (“How clear a feeling do you have about the goodness or badness of Bowl A’s 9% chance of

winning?"), and finally completed the expanded numeracy measure based on Lipkus et al. (2001). Participants lower in numeracy were more likely to choose from the bowl with the lower chance of winning (9% versus 10%), and they also reported less precise feelings about the 9% chance. The authors speculate that because the low numerate were not able to draw meaning from the percentage information, they were drawn to the objectively worse bowl by an irrelevant source of affective information – namely, the bowl with greater the number of winning beans.

In a second study, these authors used a task developed by Slovic et al. (2004). Two groups of participants are asked to rate the attractiveness of a simple gamble. The first group was given the following: 7/36 chances to win \$9 and 29/36 chances to win nothing. The second group was given a similar gamble but with a small loss: 7/36 chances to win \$9 and 29/36 chances to lose 5¢. The initial findings from Slovic et al. (2004) were that participants rated the gamble with the small loss considerably higher than the gamble with no loss. Peters et al. (2006) had participants (n=171) complete this same task, but also had them complete measures of affect and affective precision, as well as the expanded version of the Lipkus et al. (2001) numeracy scale. They found that high numerate participants rated the bet with the small loss as more attractive than the bet with no loss, while low numerate participants rated the two bets as equally attractive. In this case the high numerate participants were actually making objectively worse judgments as compared to the low numerate. The authors explain this difference by pointing out that the high numerate are actually better able to deal with numbers and therefore draw more affective meaning from numbers. In fact, high numerate participants were shown to have more positive and more precise feelings toward the 7/36 chances of winning, as well as more positive feelings toward the \$9. High numerate participants had particularly strong positive feelings toward the \$9 when it was accompanied by the small 5¢ loss, suggesting that they were particularly

sensitive to the comparison of the small loss and the comparatively much larger gain. In this case, it is possible the ability of the high numerate decision makers to draw meaning from numbers and number comparisons actually led them astray.

Summary. This work suggests that people lower in numeracy do not draw as much affective meaning from numbers, and consequently, may be more influenced by irrelevant sources of affective information. Clearly more work needs to be done relating numeracy to affective and deliberative modes of thinking. We will return to the discussion of numeracy and dual-processing models of information processing later in the review.

Selected findings from the cognitive/developmental literature related to numeracy

In this section we present a selection of findings from cognitive neuroscience and cognitive/developmental psychology that can further our understanding of numeracy. However, what follows is not a comprehensive review of the numeracy-related work from these domains. Our goal is to present some contemporary research directions that focus on different aspects of numeracy than the higher-level (primarily applied) work that we have reviewed above.

Basic number processing.

Several lines of research have focused on building cognitive and neuroanatomical models of the basic number processing systems in the brain. Much of this recent surge is due to the rapid advancement of neuroimaging technology. Due to the infancy of the field, however, most of this research is aimed at very low-level number processing – for example, simply representing numbers in the brain, or judging number magnitude.

One very robust finding in experimental psychology is known as the “distance effect” (Butterworth, 1999). Imagine that I ask you to tell me which of the following two numbers is largest: 3 or 5? This is a relatively straightforward task. However, the time that it takes for you to make this judgment is directly related to the distance between the numbers presented. For example, you would respond faster with numbers that were further apart (e.g. 1 or 9?). This simple effect is known as the “distance effect”, and it has been replicated in adults, children, as well as other non-human animals (Dehaene, 1996). Dehaene and colleagues have used this simple paradigm to probe the basic processing systems in the brain responsible for number comparisons. For example, in one study they used additive-factors methodology along with electroencephalography (EEG) to examine the cognitive processing stages and important brain areas involved in number comparison (Dehaene, 1996). They identified several cognitive processing stages (i.e. word or digit identification, magnitude comparison, response selection, and error capture and correction) and identified several areas in the parietal lobe related to number processing. In more recent work, these researchers have developed a more sophisticated model of number processing in the parietal lobe using results from fMRI and other neuropsychological work (Dehaene, Piazza, Pinel, & Cohen, 2003). Although the specifics of the model are tentative, they identify three different circuits in the parietal lobe. One circuit is involved in the manipulation of numbers regardless of number notation (e.g. eight versus 8), another is involved in the manipulation of numbers in verbal form, and the third is an attentional circuit.

Further research at the intersection of developmental psychology, neuroscience, and neuropsychology has focused on identifying the basic number processing systems in the brain. Feigenson, Dehaene, & Spelke (2004) review evidence from several research domains, which

taken together suggest the presence of two basic number processing systems. One system appears to be responsible for representing large quantities in an approximate form, while the other is responsible for representing small quantities in a precise form. The researchers note that these very basic systems appear in adults, infants, and non-human animals. These systems are necessary for representing numerical quantities and serve as the foundation for higher-level numerical skills such as understanding fractions, proportions, square roots, etc.

We are only beginning to understand how the brain represents and operates on simple numerical quantities, let alone how it deals with more complex numerical concepts like probabilities. Particularly as these fields develop further, it will be increasingly important to combine what is known at this basic cognitive/anatomical level with work at higher-levels of analysis.

Probabilities, proportions, and fractions.

As noted above, we focused the bulk of our review on numeracy defined as the ability to use probabilities, proportions, fractions, etc. We used this definition because this is often the type of numerical information that must be used for decisions that involves some uncertainty about the future (e.g. electing to have a surgery or not, or understanding the probability of death associated with not having regular mammography screenings). In this section we present some selected findings from cognitive/developmental research that can further our understanding of why people have such difficulty with probabilities, proportions, and fractions.

The field of judgment and decision making has produced an extensive literature on how adults perceive and use probabilities and other chance related concepts. This work has been detailed in numerous volumes (e.g. Kahneman, Slovic, Tversky, 1982; Wright & Ayton, 1994),

and a review of this literature is far beyond the scope of this paper. As noted above, it would be very interesting to examine the moderating role of numerical ability with respect to the fundamental findings from this research area.

There has also been work on how children and adolescents understand and use probabilities and other chance related concepts. For the most part, this developmental research has been very similar to the adult work, in that the same experimental paradigms are often used with younger participants. For example, researchers have used standard tasks to examine the representative heuristic in children (Smith, 1998), children's violations of utility theory (Schlottmann, 2000, 2001), children's understanding of expected value (Liqi, Fuxi, & Gang, 2002), and the ability of children to reason in a Bayesian manner (Zhu & Gigerenzer, 2006).

However, there are particular findings that are strongly related to how people understand probabilities and related ratio concepts (proportions, fractions). As discussed extensively by Reyna & Brainerd (2007; 2008), both ratio biases and denominator neglect may explain, in part, why both children and adults tend to have such difficulty with problems involving ratio concepts. The ratio bias is the tendency for decision makers, when given a choice between two samples, to choose the sample with the larger numerator, even if the probability is slightly worse for the larger sample. The jellybean task is an example of ratio bias that we have already encountered in our review of the adult numeracy literature. In this task, participants are asked to pick a single jellybean from one of two different bowls, and if they pick a red bean they win a prize. One bowl contains more beans with a larger numerator (more winning beans), but has an objectively worse chance of winning than the bowl with fewer winners and a smaller sample overall. The ratio bias is the tendency for participants to choose the bowl with the larger numerator. In a review of the development of probability judgments, Reyna & Brainerd (1994) discuss several

experiments from developmental psychology that show that children also show ratio biases. In addition to overweighing the numerator, they also discuss findings in which children underweigh the denominator in dealing with probability problems. This denominator neglect makes it difficult for children to solve probability problems in which the denominators differ, and these same results have been found in adult participants (Reyna & Brainerd, 2007; 2008).

Reyna & Brainerd (2007; 2008) discuss several potential explanations for why individuals have such difficulty with ratio concepts, particularly probability and risk judgments. One line of research has focused on the formatting of the probability information. Specifically, probability information in a frequency format (e.g. 10 out of 100) is thought to be more intuitive, natural and easier to understand than the equivalent information in a standard probability form (see Gigerenzer 1994, 1995). This hypothesis has received some support, although other researchers have found that individuals perform equally well with standard probability and frequency formats (see discussion in Reyna & Brainerd, 2007; 2008).

A second proposed explanation is based on recent work on dual-process theories of information processing. Several theorists have proposed that information is processed in two distinct modes: the experiential or affective mode (also called system 1) supports automatic, associative, and relatively rapid processing, and the deliberative mode (also called system 2) supports conscious, reason/rule based processing and is relatively slow (e.g., Epstein, 1994; Sloman, 1996; Kahneman & Frederick, 2002; Stanovich & West, 2000). It is thought that all judgments and decisions are based on some combination of processing from these two systems, and an overreliance on one system in a particular context can result in biases in judgment and decisions. For example, ratio biases and denominator neglect can be explained by too much experiential processing or too little deliberative processing.

The last explanation discussed by Reyna and Brainerd (2007; 2008) is based on their fuzzy-trace theory (see Reyna & Brainerd, 1992). Fuzzy-trace theory is also a dual process approach in which information is processed and stored in two distinct ways: intuitive, relatively vague, and gist-based representations (basic semantic representation), and a separate detailed, verbatim, quantitative memory representation. Judgment biases, such as ratio biases and denominator neglect, are explained by the workings of these two systems of representating information. For example, there are some judgment problems in which fuzzy gist representations, although normally very functional and adaptive, can lead to judgment biases.

In summary, researchers have begun to develop theories of the cognitive processes that are responsible for the difficulty individuals have with probabilities and other ratio concepts. This work can provide a framework for further insights into these numerical problems, as well as provide predictions of how individuals varying in numerical ability are likely to behave in contexts in which probabilities and ratio concepts are fundamental.

Cognitive models of higher-level number processing.

Researchers have developed a relatively large body of work on the cognitive processes involved in various numerical computation tasks (e.g. Sokol, et al. 1991; Frensch & Geary, 1993), although relatively little is known about the cognitive processing involved in higher-level numerical decision making tasks. Kida & Smith (1995) provide one example of cognitive modeling intended to aid in the understanding of a higher-level decision making task. The authors present a model of the encoding and retrieval processes involved in representing numerical information in accounting contexts. They rely on research findings from human memory, social cognition and cognitive development, including concepts from fuzzy-trace

theory and aspects of evaluative/affective processing that are fundamental to the dual process theories discussed above.

For our purposes, the general approach is more important than the specifics of this particular cognitive process model. One approach to studying the effects of numeracy and decision making would be to specify the precise cognitive stages that are needed to accomplish the particular numerical task under study. Then further investigations can be initiated to assess the differences between those individuals that vary in numerical ability.

Summary and Discussion

Summary of primary findings relating numeracy to decision making

The current research findings relating decision making and numeracy in healthcare and related domains appears to be relatively isolated with little theoretical connection between studies. In brief, we find that individuals lower in numeracy are more likely to have poor health and financial outcomes, have difficulty judging risks and providing consistent assessments of utility, are worse at reading graphs, show larger framing effects, are sensitive to the formatting of probability information, trust verbal more than numerical information, are more likely to show small-stakes risk aversion and short-run discounting, and appear to not draw as much affective meaning from numbers. Overall, at least in the tasks used by researchers thus far, it is clearly worse for decision making performance to have poor numerical skills. In addition, there has also been some work on how best to present information to the lower numerate, which includes presenting information in a logically ordered format and displaying only the important information, presumably decreasing the cognitive burden on the participants.

There have been attempts, however, to relate the current numeracy findings to broader theories of information processing. For example, dual process theories could be an informative framework in which to interpret existing numeracy findings and to initiate future research. As discussed briefly above, many of the numeracy findings could be interpreted in the context of deliberative-experiential dual processing – namely, different mixtures of deliberative and experiential/affective processing by those high and low in numeracy. For example, Peters et al. (2006) make experimental predictions by discussing the deliberative and experiential processing of those high and low in numeracy. They note that participants higher in numeracy are better able to understand numbers in a deliberative fashion and are also able to draw relevant affective meaning from numbers. Participants lower in numeracy do not have the deliberative capacity to use the numbers, cannot draw affective meaning from them and, therefore, are more susceptible to irrelevant sources of information. Benjamin et al. (2006) also discuss their findings regarding small-stakes risk aversion and short-run discounting in terms of dual processing models.

Fuzzy-trace theory is another dual processing approach that can be useful in making predictions about the differences between the low and high numerate in judgment and decision making. As discussed above, several judgment biases closely related to understanding ratio concepts (probabilities, proportions, etc) can be understood in terms of fuzzy trace theory. Specifically, it may be that the low numerate are more likely to rely on gist processing, which in some cases results in poorer judgments and decisions. Although they focus on simpler numerical systems, Feigenson et al. (2004) also discuss a kind of dual processing system for numbers, in which one system is responsible for representing large, approximate quantities, and the other represents small quantities in exact form. There may be important connections between these different dual processing theories. The experiential (affective) system, gist processing, and the

more basic systems of representing approximate magnitude may be describing related cognitive and neural systems. For more discussion of the relationship between fuzzy trace theory and deliberative/experiential dual processing systems see Reyna & Brainerd (2007; 2008).

Gaps in our knowledge and Future Directions.

Below we discuss some gaps in our knowledge concerning numeracy and some possible future directions for the field. In brief, we discuss the need for a deeper understanding of the precise numerical skills that are used by decision makers and the corresponding need for more sophisticated measures of numeracy. Then researchers can more effectively investigate the relationships between numeracy and decision making, as well as develop effective information presentation techniques and educational interventions.

The need for a deeper understanding of numerical skills needed for decision making tasks.

Individuals engage in many different kinds of judgment and decision making tasks that involve numerical information. There is a set of numerical skills that are needed depending on the task and the goals of the individual, and these skills are not captured in our current conceptualizations of numeracy. For example, there are several skills that are needed for a patient to use probabilistic information concerning risks and benefits. First, patients must be able to access the information. This involves the ability to identify likelihood information from written text, tables and charts. Then patients need to manipulate and use the information by performing calculations, making comparisons, and drawing inferences. For example, a patient might be presented with survival rates for several different treatment options. The patient must be able to first identify the survival rates in the presented information, make some kind of

comparison between the different options (e.g. calculate a difference), and make an inference about the meaning of the comparison. The patient's decision will be the result of a reasoning process involving the survival rate information, the patient's needs and values, and a series of trade-offs. In addition, there is often the added complexity of a time delay between costs and rewards. Many of these decisions involve making concrete sacrifices now with the hope of long-term, probabilistic rewards in the future.

Current discussions of numeracy are often too broad and do not always specify the specific skills that are needed for numerical decision tasks. It is important to recognize that most judgments and decisions involving numerical information involve much more than just an understanding of probabilities or proportions. New conceptualizations of numeracy need to account for the hierarchy of skills that are needed for many of these judgments and decisions. One way toward this goal would be to focus on building models of the actual processes and skills used by decision makers in these tasks. With more precise models of the skills involved in various numerical judgments and decisions, we may be better able to specify the processing differences between the low and high numerate. For example, combining the theories of general information processing (i.e. the dual process theories discussed above) with a more sophisticated understanding of specific task requirements would allow researchers to make more precise predictions about the performance of individuals varying in numerical ability.

The measurement of numeracy.

With a more sophisticated understanding of the skills needed for numerical decision tasks, researchers would be able to develop better measures of numeracy. Some researchers have, in fact, developed numeracy measures that focus on the specific skills necessary for a given task

(e.g. the numeracy measures for diabetic and asthmatic patients discussed in the measurement section). The medical data interpretation test is another good example of a more sophisticated numeracy measure focusing not only on the understanding on probabilistic information, but also on how this information is used to make judgments and decisions.

In summary, more work is needed to identify the primary numerical skills that are most important for decision making outcomes, and correspondingly which skills should be captured when measuring numeracy.

Improving information presentation.

There has been some progress in developing information presentation formats that are effective for individuals varying in numerical ability, and we should continue to use findings from cognitive psychology and judgment and decision making on how best to present information to decision makers. However, with a more developed understanding of the specific processing differences between individuals varying in numerical ability we will be better able to develop more effective information presentation formats.

Final remarks

The study of numeracy and decision making behavior is an exciting and multi-disciplinary research area. This field has the potential to benefit large segments of the population by improving the decision making environment for people with lower levels of numeracy. However, the field as a whole is in its infancy and would greatly benefit from more theoretical development and improved measurement of numerical skills.

References

- Apter, A. J., Cheng J, Small D, Bennett IM, Albert C, Fein DG, George M, Van Horne S. (2006). Asthma numeracy skill and health literacy. *J Asthma*, 43(9), 705-710.
- Baker, D. W. (2006). The meaning and the measure of health literacy. *Journal of General Internal Medicine*, 21(8), 878-883.
- Baker, D. W., Parker, R. M., Williams, M. V., Clark, W., & Nurss, J. (1997). The relationship of patient reading ability to self-reported health and use of health services. *American Journal of Public Health*, 87(6), 1027-1030.
- Baker, D. W., Williams, M. V., Parker, R. M., Gazmararian, J. A., & Nurss, J. (1999). Development of a brief test to measure functional health literacy. *Patient Education & Counseling*, 38(1), 33-42.
- Baker, D. W., J. A. Gazmararian, M. V. Williams, T. Scott, R. M. Parker, D. Green, J. Ren, and J. Peel (2002). Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *American Journal Public Health*, 92(8), 1278-1283.
- Baker, D. W., J. A. Gazmararian, M. V. Williams, T. Scott, R. M. Parker, D. Green, J. Ren, and J. Peel (2004). Health literacy and use of outpatient physician services by Medicare managed care enrollees. *Journal of General Internal Medicine*, 19(3), 215-220.
- Banks, J. & Oldfield, Z. (2006). *Understanding pensions: Cognitive function, numerical ability and retirement saving* (WP 06/05). London: The Institute for Fiscal Studies.
- Beaman, A. L. (1991). An empirical comparison of meta-analytic and traditional reviews. *Personality & Social Psychology Bulletin*, 17, 252-257.
- Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2006). *Who is "behavioral"?* *Cognitive ability and anomalous preferences*. Cambridge, MA: Harvard University.

- Berger, C. R. (2002). Base-rate bingo - Ephemeral effects of population data on cognitive responses, apprehension, and perceived risk. *Communication Research*, 29(2), 99-124.
- Beverly, S. G., & Burkhalter, E. K. (2005). Improving the financial literacy and practices of youths. *Children & Schools*, 27(2), 121-124.
- Black, W. C., Nease, R. F. & Tosteson, A. N. (1995). Perceptions of Breast Cancer Risk and Screening Effectiveness in Women Younger Than 50 Years of Age. *J Natl Cancer Inst*, 87, 720-731.
- Burkell, J. (2004). What are the chances? Evaluating risk and benefit information in consumer health materials. *Journal of the Medical Library association*, 92(2), 200-208.
- Butterworth, B. (1999). *The Mathematical Brain*. Macmillan.
- Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. *Journal of Personality*, 48, 449-472.
- Cooper, H. M. & Hedges, L. V. (Eds.) (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine*, 19(4), 310-315.
- Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, 8(1), 47-68.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3/4/5/6), 487-506.

- Dillard, A. J., McCaul, K. D., Kelso, P. D., & Klein, W. M. P. (2006). Resisting good news: Reactions to breast cancer risk communication. *Health Communication, 19*(2), 115-123.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist, 49*, 709-724.
- Estrada, C., Barnes, V., Collins, C., & Byrd, J. C. (1999). Health literacy and numeracy. *JAMA: Journal of the American Medical Association, 282*(6), 527.
- 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*(2), 88-93.
- Fagerlin, A., Rovner, D., Stableford, S., Jentoft, C., Wei, J. T., & Holmes-Rovner, M. (2004). Patient education materials about the treatment of early-stage prostate cancer: A critical review. *Annals of Internal Medicine, 140*(9), 721-728.
- Fagerlin, A., Ubel, P. A., Smith, D. M. & Zikmund-Fisher, B. J. (submitted). Making numbers matter: Present and future research in risk communication.
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (submitted). Measuring numeracy without a math test: development of the subjective numeracy scale (SNS).
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307-314.
- Fox, J., Bartholomae, S., & Lee, J. (2005). Building the Case for Financial Education. *Journal of Consumer Affairs, 39*(1), 195-214.

- Frensch, P. A. & Geary, D. C. (1993). Effects of practice on component processes in complex mental addition. *Journal of experimental psychology: Learning, memory, and cognition*, 433-456.
- Gazmararian, J. A., M. V. Williams, J. Peel, and D.W. Baker (2003). Health Literacy and Knowledge of Chronic Disease. *Patient Education and Counseling*, 51 (3), 267-275.
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective probability*. (pp. 129-161). New York: Wiley.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, 102, 684-704.
- Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine*, 29(4), 375-376.
- Gramling, R., Irvin, J. E., Nash, J., Sciamanna, C., & Culpepper, L. (2004). Numeracy and medicine: key family physician attitudes about communicating probability with patients. *J Am Board Fam Pract*, 17(6), 473.
- Gurmankin, A. D., Baron, J., & Armstrong, K. (2004a). The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. *Medical Decision Making*, 24(3), 265-271.
- Gurmankin, A. D., Baron, J., & Armstrong, K. (2004b). Intended Message Versus Message Received in Hypothetical Physician Risk Communications: Exploring the Gap. *Risk Analysis*, 24(5), 1337-1347.

- Haggstrom, D. A., & Schapira, M. M. (2006). Black-white differences in risk perceptions of breast cancer survival and screening mammography benefit. *Journal of General Internal Medicine, 21*(4), 371-377.
- Hibbard, J. H., & Peters, E. (2003). Supporting informed consumer health care decisions: Data presentation approaches that facilitate the use of information in choice. *Annual Review of Public Health, 24*, 413-433.
- Hibbard, J. H., Slovic, P., Peters, E., Finucane, M. L., & Tusler, M. (2001). Is the informed-choice policy approach appropriate for Medicare beneficiaries? *Health Affairs, 20*(3), 199-203.
- Hilgert, M. A., Hogarth, J.M., & Beverly, S.G. (2003). Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin* (July), 309-322.
- Journal of the National Cancer Institute Monographs (1999). Cancer risk communication: What we know and what we need to learn. Volume, 25.
- Kahneman, D., Slovic, P. & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 49–81). New York: Cambridge University Press.
- Kida, T., & Smith, J. F. (1995). The encoding and retrieval of numerical data for decision making in accounting contexts: Model development. *Accounting, Organizations and Society, 20*(7/8), 585-610.

- Kirsch, I. S., Jungeblut, A., Jenkins, L., & Kolstad, A. (2002). *Adult literacy in America: A first look at the findings of the National Adult Literacy Survey* (3rd ed. Vol. 201). Washington, DC: National Center for Education, U.S. Department of Education.
- Kuehberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior & Human Decision Processes*, 75(1), 23-55.
- Kutner, M., Greenberg, E., Jin, Y., and Paulsen, C. (2006). *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy* (NCES 2006-483). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Lillard & Willis. (2001). *Cognition and Wealth: The importance of probabilistic thinking* [WP 2001-007]. Ann Arbor, MI: University of Michigan Retirement Research Center.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21, 37-44.
- Liqi, Z., Fuxi, F., & Gang, H. (2002). A study on children and adults judgement of expected value. *Acta Psychologica Sinica*, 34(5), 517-521.
- Lucey, T. A. (2005). Assessing the Reliability and Validity of the Jump\$Start Survey of Financial Literacy. *Journal of Family and Economic Issues*, 26(2), 283-294.
- Lusardi, A. & Mitchell, O. S. (2006). *Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth* (WP 2006-114). Ann Arbor, MI: University of Michigan Retirement Research Center.
- Montori, V., Leung, T., Thompson, C. Chung, J., Capes, S., Smith, S. (2004). Validation of a diabetes numeracy evaluation tool. *Diabetes*, 53(suppl. 2), A224-225.

- Montori, V. M., & Rothman, R. L. (2005). Weakness in Numbers: The Challenge of Numeracy in Health Care. *Journal of General Internal Medicine*, 20(11), 1071-1072.
- National Center for Education Statistics. *National Assessment of Adult Literacy (NAAL)*. Retrieved January, 10, 2007 from <http://nces.ed.gov/NAAL/index.asp?file=KeyFindings/Demographics/Overall.asp&PageId=16>
- Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*, 10(10), 537-541.
- Parrott, R., Silk, K., Dorgan, K., Condit, C., & Harris, T. (2005). Risk Comprehension and Judgments of Statistical Evidentiary Appeals: When a Picture is Not Worth a Thousand Words. *Human Communication Research*, 31(3), 423-452.
- Parsons, S., & Bynner, J. (1997). Numeracy and employment. *Education & Training*, 39(2), 43-51.
- Paulos, J. A. (Ed.). (1988). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Hill and Wang.
- Perry, V. G., & Morris, M. D. (2005). Who is in control? The role of self-perception, knowledge, and income in explaining consumer financial behavior. *Journal of Consumer Affairs*, 39(2), 299-313.
- Peters, E., Dieckmann, N. F., Dixon, A., Hibbard, J. H., Mertz, C. K., & Slovic, P. (in press). Less is more in presenting quality information to consumers.
- Peters, E., Hibbard, J. H., Slovic, P. & Dieckmann, N. F. (Submitted). Numeracy skill and the communication, comprehension, and use of risk and benefit information.

- Peters, E., Vastfjall, D., Slovic, P., Mertz, C., Mazzocco, K., & Dickert, S. (2006). Numeracy and Decision Making. *Psychological Science, 17*(5), 407-413.
- Reyna, V. F., & Brainerd, C. J. (1992). A fuzzy-trace theory of reasoning and remembering: paradoxes, patterns, and parallelism. In N. Hearst, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (pp. 235-259). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Reyna, V.F., & Brainerd, C.J. (1994). The origins of probability judgement: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability*. (pp. 239-272). New York: Wiley.
- 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.
- 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.
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin, 86*, 638-641.
- Rudd, R. E., Colton, T., & Schacht, R. (2000). *An overview of medical and public health literature addressing literacy issues: An annotated bibliography*. (NCSALL Reports #14). Cambridge, MA: The National Center for the Study of Adult Learning and Literacy.
- Schapira, M. M., Davids, S. L., McAuliffe, T. L., & Nattinger, A. B. (2004). Agreement between scales in the measurement of breast cancer risk perceptions. *Risk Analysis, 24*(3), 665-673.

- Schlottmann, A. (2000). Children's judgements of gambles: A disordinal violation of utility. *13(1)*, 77-89.
- Schlottmann, A. (2001). Children's probability intuitions: Understanding the expected value of complex gambles. *72(1)*, 103-122.
- Schwartz, L. M., Woloshin, S., & Welch, H. (2005). Can Patients Interpret Health Information? An Assessment of the Medical Data Interpretation Test. *Medical Decision Making*, *25(3)*, 290-300.
- Schwartz, L. M., Woloshin, S., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127(11)*, 966-972.
- Schwartz, S. R., McDowell, J., & Yueh, B. (2004). Numeracy and the shortcomings of utility assessment in head and neck cancer patients. *Head & Neck*, *26(5)*, 401-407.
- Scott, T. L., J. A. Gazmararian, M. V. Williams, and D. W. Baker (2002). Health literacy and preventive health care use among Medicare enrollees in a managed care organization. *Medical Care*, *40(5)*, 395-404.
- Sheridan, S. L., & Pignone, M. (2002). Numeracy and the medical students' ability to interpret data. *Effective Clinical Practice*, *5(1)*, 35-40.
- Sheridan, S. L., Pignone, M. P., & Lewis, C. L. (2003). A randomized comparison of patients' understanding of number needed to treat and other common risk reduction formats. *Journal of General Internal Medicine*, *18(11)*, 884-892.
- Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3-22.

- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, 24, 311-322.
- Smith, H. (1998). Misconceptions of chance: Developmental differences and similarities in use of the representativeness heuristic. *Psychological Reports*, 83(2), 703-707.
- Sokol, S. M., McClosky, M., Cohen, N. J. & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of experimental psychology: Learning, memory, and cognition*, 355-376.
- Speros, C. (2005). Health literacy: Concept analysis. *JAN Journal of Advanced Nursing*, 50(6), 633-640.
- Stanovich, K. E., & West, R. F. (1998). Individual differences in framing and conjunction effects. *Thinking and Reasoning*, 4, 289-317.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioural and Brain Sciences*, 23, 645–726.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *The American Economic Review*, 93(2), 174-179.
- Weinfurt, K. P., Castel, L. D., Li, Y., Sulmasy, D. P., Balshem, A. M., Benson, A. B., 3rd, et al. (2003). The correlation between patient characteristics and expectations of benefit from Phase I clinical trials. *Cancer*, 98(1), 166-175.
- Weinfurt, K. P., DePuy, V., Castel, L. D., Sulmasy, D. P., Schulman, K. A., & Meropol, N. J. (2005). Understanding of an aggregate probability statement by patients who are offered participation in Phase I clinical trials. *Cancer*, 103(1), 140-147.

- Wiess, B. D., Mays, M. Z., Martz, W., Castro, K. M., DeWalt, D. A. & Pigone, M. P. (2005). Quick assessment of literacy in primary care: the newest vital sign. *Ann Fam Med*, 3, 514-522.
- Williams, M. V., D. W. Baker, E. G. Honig, T. M. Lee, and A. Nowlan (1998a). Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest*, 114(4), 1008-1015.
- Williams, M. V., D. W. Baker, R. M. Parker, and J. R. Nurss (1998b). Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. *Archives of Internal Medicine*, 158(2), 166-172.
- Wolf, M. S., J. A. Gazmararian, and D. W. Baker (2005). Health literacy and functional health status among older adults. *Archives of Internal Medicine*, 165, 1946-1952.
- Woloshin, S., Schwartz, L. M., Black, W. C., & Welch, H. G. (1999). Women's perceptions of breast cancer risk: how you ask matters. *Medical Decision Making*, 19, 221-229.
- Woloshin, S., Schwartz, L. M., Moncur, M., Gabriel, S., & Tosteson, A. N. A. (2001). Assessing values for health: Numeracy matters. *Medical Decision Making*, 21(5), 382-390.
- Woloshin, S., Schwartz, L. M., & Welch, H. (2005). Patients and Medical Statistics: Interest, Confidence, and Ability. *Journal of General Internal Medicine*, 20(11), 996-1000.
- Woloshin, S., Schwartz, L. M., & Welch, H. G. (2004). The value of benefit data in direct-to-consumer drug ads. *Health Affairs (Supplement Web Exclusive)*, 23(3), W4234-W4245.
- Wright, G. & Ayton, P. (Eds.) (1994). *Subjective Probability*. West Sussex, England: John Wiley & Sons Ltd.
- Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, 98(3), 287-308.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (in press). Validation of the subjective numeracy scale (SNS): Effects of low numeracy on comprehension of risk communications and utility elicitation. *Medical Decision Making*.