Published in final edited form as:

J Empir Res Hum Res Ethics. 2009 December ; 4(4): 17–26. doi:10.1525/jer.2009.4.4.17.

The Role of Numeracy in Informed Consent for Surveys

Mick P. Couper and Eleanor Singer

University of Michigan (USA)

Abstract

If empirical estimates of disclosure risk are included in informed consent statements for surveys or other forms of research, participants must be able to understand the information provided. Using data from an online vignette-based experiment, this article explores the role that numeracy or quantitative literacy may play in comprehension of disclosure risk. Results suggest that less numerate persons are less sensitive to extreme differences in the disclosure risk described in the hypothetical vignettes.

Keywords

informed consent; numeracy; disclosure risk; survey participation

Two key ethical issues for survey researchers are obtaining informed consent and maintaining the confidentiality of responses. *Informed consent* implies that two requirements have been met: (1) that research participants have been informed about the essential elements of the research, including the risks and benefits of participation, and have understood the information; and (2) that they have given their consent to participate. Although known breaches of confidentiality in survey research are exceedingly rare, such breaches are the most likely source of risk or harm to survey respondents. Recent research has focused on quantification of the risk of disclosure of personal information from surveys and on ways to minimize such risk through statistical disclosure limitation (e.g., Raghunathan, Reiter, & Rubin, 2003; Reiter, 2009; Zarate & Zayatz, 2006).

The importance of comprehension in assuring informed consent is increasingly recognized in medical research (e.g., Bergenmar et al., 2008; Hochhauser, 2008; McNutt et al., 2008). With respect to social surveys, research on informed consent (see Singer, 1993 for a review) has largely focused on how the kind or amount of information disclosed to respondents affects their willingness to participate and on how assurances of confidentiality affect response rate and quality. Relatively little research has focused on issues of comprehension.

This paper focuses on the role of one aspect of comprehension—quantitative literacy, or numeracy—in understanding statements of disclosure risk in surveys. Assuming that communicating the potential risk of disclosure to respondents is ethically desirable, what is the best way of doing so? How well do respondents understand statements of quantitative risk? We explore what role, if any, numeracy plays in respondents' preference for disclosure

Address correspondence to: Mick Couper, Survey Research Center, Institute for Social Research, University of Michigan, P.O. Box 1248, Ann Arbor, MI 48106. Phone: 734- 647-4599; mcouper@umich.edu.

PLEASE DIRECT ALL REQUESTS FOR PERMISSIONS TO PHOTOCOPY OR REPRODUCE ARTICLE CONTENT THROUGH THE UNIVERSITY OF CALIFORNIA PRESS'S RIGHTS AND PERMISSIONS WEBSITE, [HTTP://WWW.UCPRESSJOURNALS.COM/REPRINTINFO.ASP](http://WWW.UCPRESSJOURNALS.COM/REPRINTINFO.ASP).

risk expressed in words or numbers, and in their sensitivity to disclosure risk manipulations in hypothetical survey descriptions.

Background

The idea of numeracy as a distinct component of literacy has gained traction in recent years, especially in the health field. According to an Institute of Medicine (2004) report, health literacy is defined as “the degree to which individuals can obtain, process, and understand the basic health information and services they need to make appropriate health decisions.” Components of literacy identified by the IOM include both oral and print literacy, and numeracy. There are many definitions of numeracy or quantitative literacy; for example, numeracy is defined by Rothman and colleagues (2008) as the “ability to understand and use numbers in daily life,” by Peters et al. (2006) as the “ability to process basic probability and numerical concepts,” and by Fagerlin et al. (2007) as “aptitude with probabilities, fractions, and ratios.” Our goal is not to develop or refine the meaning of numeracy, but rather to use existing measures of numeracy to explore issues related to disclosure risk and informed consent in surveys.

A number of different measures of numeracy have been developed (Rothman et al., 2008), ranging in scope from a few questions (Schwartz et al., 1997; Lipkus, Samsa, & Rimer, 2001) to up to ten minutes administration time (e.g., the numeracy component of the Test of Functional Health Literacy in Adults, or TOFHLA; Parker et al., 1995). Most of these have employed an objective numeracy test, in the form of mathematical problems to test understanding of frequencies, probabilities, and percentages. However, Fagerlin et al. (2007) argue that participants in research studies are “not especially receptive to taking an aptitude test.” They further note that objective numeracy measures may not work well on the telephone (given the higher cognitive load and working memory required) or on the Internet (given that participants could use calculators or consult others). Given these concerns, they developed an eight-item subjective numeracy scale (Fagerlin et al., 2007) with good internal consistency (Cronbach’s $\alpha = 0.82$) and relatively high correlation ($r = 0.53$) with an objective numeracy scale formed by combining the Schwartz et al. (1997) and Lipkus, Samsa, and Rimer (2001) measures. This scale has subsequently been validated in two web surveys and a paper survey among members of the general population (Zikmund-Fisher et al., 2007). In this paper, we use three items from the Schwartz et al. (1997) objective numeracy scale and the eight-item Fagerlin et al. (2007) subjective numeracy scale.

The research reported here is part of a broader research program aimed at estimating the probability (risk) of statistical disclosure in publicly available data files and developing the best way of communicating such risk to potential respondents. Our part of the research program was to craft informed consent statements using empirically-determined estimates of disclosure risk that would meet ethical requirements without unnecessarily arousing respondents’ concerns about the confidentiality of their personal information (see Couper et al., 2008, 2009). This paper explores the role that numeracy plays in respondents’ understanding of risk information and its effect on their willingness to participate in surveys.

We address several research questions with respect to numeracy. Specifically:

1. How do objective and subjective measures of numeracy developed for the health field perform in a survey setting, and particularly on the Web?
2. What are the demographic correlates of numeracy?
3. Is numeracy related to preferences for disclosure risk expressed in words or numbers?

4. What role, if any, does numeracy play in the effect of risk disclosure manipulations on willingness to participate in surveys?

The last two questions are specific to work on finding ways to best communicate the disclosure risk as part of the informed consent process. We expect that more numerate individuals (i.e., those with a better understanding of numbers) would be more comfortable with risk expressed in numbers, while less numerate persons would express a preference for risk described in words. The last research question specifies an interaction between numeracy and the level of risk described in the survey on stated willingness to participate. Our expectation is that more numerate persons would be more sensitive to the risk manipulation than less numerate persons.

Methods

Sample and Administration

As noted earlier, the analyses reported here form part of an experiment exploring the effect of varying the stated risk of disclosure on willingness to participate in hypothetical surveys described in a series of eight vignettes. Those results and the design of the broader study are reported elsewhere (Couper et al., 2009). Our focus here is on the numeracy questions that formed part of that study. The data are from a web survey administered by Market Strategies Inc., on a volunteer sample drawn from Survey Sampling International's Internet panel. We received 6400 completed questionnaires¹ out of a total of 217,542 invitations sent, for a "response rate" of 2.9 percent. Invitees were told it was a study of survey participation, and that they would see descriptions of different types of surveys and be asked whether or not they'd be likely to take part.

This is not a probability sample. Our focus, to use Kish's (1987) terms, is on randomization rather than representation. We view this as an experiment with a large and diverse group of volunteer subjects. Our respondents are more middle-aged (47.2% age 45–64, 8.1% age 65 or older) than the general population (33.6% age 45–64, 16.6% age 65+, according to census estimates for 2007²); they are also better educated (37.9% college graduates, 21.2% high school or less) compared to the general population (26.2% college graduates, 46.6% HS or less, according to NCES estimates for 2007³). We have more females (54.1%) than the general population (49.4%) and more non-Hispanic whites (78.9% in our survey, 66.0% in the population, based on 2007 census estimates). Given that our sample is based on Internet volunteers, we anticipate overestimating the levels of numeracy relative to the general population. The likely exclusion of those with the lowest levels of numeracy may also attenuate possible relationships between numeracy and our variables of interest. Thus, if anything, we are likely to underestimate the effect of numeracy.

Questionnaire

Following the eight fictional survey invitation vignettes (described in Couper et al., 2009), respondents were asked a variety of additional items, including questions designed to explore perceptions of risk and benefit as well as concerns about privacy and attitudes toward surveys (see also Couper et al., 2008). They were also asked the questions related to numeracy that form the basis of the analyses described below. The entire questionnaire took about 16 minutes to complete.

¹An additional 2470 started the survey but did not complete it. We exclude these cases from our analysis.

²http://www.census.gov/popest/archives/2000s/vintage_2007/

³http://nces.ed.gov/programs/digest/d07/tables/dt07_009.asp

The vignettes manipulated four factors: (1) the risk of disclosure (no mention of risk; no chance of disclosure; one in a million; and one in ten) and a description of the harm such disclosure might cause (tailored to the specific topic of the vignette); (2) the survey topic (two high-sensitivity topics [sexual behavior and personal finances] and two low-sensitivity topics [leisure activities and work]); (3) the size of the incentive for participation (\$10 or \$50); and (4) the mode (face-to-face or mail). Each subject was exposed to a set of eight vignettes, with each set containing risk descriptions at all four levels, one each for a sensitive and a nonsensitive topic. The sets were randomly assigned to subjects after they had agreed to participate in the web survey, and the order in which the vignettes were administered was also randomized within subjects. Willingness to participate (WTP) was measured by a single question, asked immediately after each vignette had been read:

On a scale from zero to ten, where zero means you would definitely not take part and ten means you would definitely take part, how likely is it that you would take part in this survey?

Following the set of vignettes and a series of general privacy and confidentiality questions, respondents were asked about their preference for risk described in words or numbers, followed by the subjective and objective numeracy measures.

PREFERENCE FOR WORDS VERSUS NUMBERS—Respondents were randomly assigned to one of two versions of the question, with one mentioning first numbers and then words (as in the example below) and the other mentioning words then numbers, in order to control for any order effects. The response options were ordered to match the question wording. One version of the question is presented here:

Suppose you had your choice of how the chance of somebody finding out your name and address, along with your answers to a survey, was described. Would you prefer to see it described in numbers (e.g., “[numeric risk]”) or in words (e.g., “[verbal risk]”)?

1. Strongly prefer numbers
2. Somewhat prefer numbers
3. Doesn't matter
4. Somewhat prefer words
5. Strongly prefer words
6. Other (specify)

Respondents were randomized to receive one of the following numeric risks:

1. 1/10 (one in ten)
2. 1/100 (one in a hundred)
3. 1/1000 (one in a thousand)
4. 1/100,000 (one in a hundred thousand)
5. 1/1,000,000 (one in a million)
6. 0 (zero)

In similar fashion, they were randomized to one of the following verbal risk descriptions:

1. A very large chance
2. A large chance

3. A small chance
4. A very small chance
5. An extremely small chance
6. No chance

Following these items, respondents were asked a series of attitude items on privacy and confidentiality before the balance of the numeracy items below were asked.

SUBJECTIVE AND OBJECTIVE NUMERACY SCALES—The three items from the Schwartz et al. (1997) objective numeracy scale are as follows:

- a. Imagine that we have a pair of dice (for example, from a board game or a casino craps table). Imagine we now roll one of them 1,000 times. Out of 1,000 rolls, how many times do you think it would come up even (numbers 2, 4, or 6)?
_____ times
- b. In the Big Bucks Lottery, the chance of winning a \$10.00 prize is 1%. What is your best guess about how many people would win a \$10.00 prize if 1,000 people each buy a single ticket to Big Bucks?
_____ people
- c. In the Acme Publishing Sweepstakes, the chance of winning a car is 1 in 1,000. What percent of tickets to Acme Publishing Sweepstakes win a car?
_____ percentu

The scale score is simply the sum of correct responses to the three items.

The subjective numeracy scale developed by Fagerlin et al. (2007) has eight items in two subscales, one measuring preference for display of numeric information and the other measuring self-reported cognitive abilities. The preference items are as follows:

For each of the following questions, please choose the number from 1 (not at all helpful) to 6 (extremely helpful) that comes closest to your answer:

- a. When reading the newspaper, how helpful do you find tables and graphs that are parts of the story?
- b. When people tell you the chance of something happening, do you prefer that they use words (e.g., “it rarely happens”) or numbers (e.g., “there’s a 1% chance”)?
- c. When you hear a weather forecast, do you prefer predictions using percentages (e.g., “there will be a 20% chance of rain today”) or predictions using only words (e.g., “there’s a small chance of rain today”)?
- d. How often do you find numerical information to be useful?

Item (c) is reverse scored. The items measuring self-reported cognitive ability are as follows:

For each of the following, please choose the number from 1 (not at all good) to 6 (extremely good) that comes closest to your answer:

- a. How good are you at working with fractions?
- b. How good are you at working with percentages?
- c. How good are you at calculating a 15% tip?

- d. How good are you at figuring out how much a shirt will cost if it's 25% off?

The scale score is the mean of responses to the eight items (with the third preference item reversed).

The order of the objective and subjective numeracy questions was randomized, such that about half the respondents first got the objective numeracy questions followed by the subjective numeracy items, while the balance got them in the reverse order. The survey then ended with a series of demographic questions and debriefing items.

The study was approved by the University of Michigan Behavioral Sciences IRB.

Results

Objective and Subjective Numeracy Scores

Following Schwartz et al. (1997), all missing data on the three-item objective numeracy scale (ONS) are treated as incorrect. The score is then simply the number of correct answers (range 0–3). The overall mean is 1.46 (S.D. = 1.04), with one-fifth (21.7%) of respondents getting none of the items correct and a similar proportion (19.8%) getting all three correct. Looking at the individual items, 57.8% got the first one (a) correct, 61.6% got (b) correct and 26.7% got (c) correct. In terms of missing data, 96.5% of respondents provided an answer to all three questions. However, a number of these responses were “don’t know,” “no idea,” and the like. Further, given that these questions elicited open-ended responses, considerable effort was needed to classify the answers as correct or incorrect.

Even though we know of no national estimates, the ONS has been used in a variety of settings, so we can see how well our group of Internet volunteers performs relative to these other groups. The Schwartz et al. (1997) study that first reported use of these items administered them by mail to a group of 500 female veterans from a New England registry. In their sample, 30% of women had no correct answers, while 16% had three correct answers. In a convenience sample of 140 community-dwelling older adults (age 50+) in Ontario, Donelle, Hoffman-Goetz, and Arocha (2007) also reported 16% getting all three items correct. Lipkus, Samsa, and Rimer (2009) administered the three-item scale to a sample of 463 volunteers recruited by mail. They found that 24% got none of the three correct, while 18% got all three correct.

Even though Internet users (and especially those who volunteer for opt-in panels) have higher levels of education than the general population and a minimum level of literacy is required for completion of an Internet survey, we see relatively low levels of numeracy, consistent with other convenience samples. This is true even though respondents are not under the same time pressures as in other modes of data collection (especially interviewer administration) and could in theory look up the answers or consult others for the correct answers. Our focus is not on estimating numeracy levels of the general population, but rather on exploring how variation in performance on the ONS relates to other variables of interest. For our purpose, we have sufficient variation in scores on the ONS.

As noted, the subjective numeracy scale (SNS) consists of two subscales. Following Fagerlin et al. (2007), all respondents missing more than half of the items (three or more items for the subscales, five or more for the full scale) were treated as missing. Doing so, we lose 25 cases on the first subscale (preference), 23 on the second (ability) and 18 on the full SNS. All in all, 96.8% of respondents gave answers to all eight items on the scale, and a further 2.7% answered seven of the eight items. The scores on each subscale are simply the means of the non-missing responses. The two subscales are correlated ($r = 0.431$) so we focus on the full SNS. Cronbach's alpha for the eight-item scale is 0.821, which matches the

reliability ($\alpha = 0.82$) reported by Fagerlin et al. (2007). The eight-item SNS is in turn correlated ($r = 0.430$) with ONS, slightly lower than the $r = 0.53$ reported by Fagerlin et al. (2007) between the SNS and an eleven-item objective numeracy scale which included the three Schwartz et al. (1997) items. Given this correlation, we standardized both scales (mean = 0, S.D. = 1) to facilitate combining them into an overall measure.

The order in which the ONS and SNS items were presented is significant ($p < .001$); asking the subjective questions first results in a higher mean for both subjective and objective scales. Further, the correlation between ONS and SNS is higher (0.447) when SNS is asked first than when ONS is asked first (0.405). To control for this, we simply added an indicator to the models to control for the order of scale presentation.

Demographic Correlates of Numeracy

We turn next to an examination of demographic correlates of numeracy. A model regressing the standardized combined (ONS+SNS) numeracy measure on selected demographic variables and panel characteristics is presented in Table 1.

About one-fifth of the variation in numeracy scores is explained by this set of predictors. The correlates of numeracy match those reported in other studies: numeracy declines with age, but increases with education level; women, African-Americans, and Hispanics have lower levels of numeracy. Those who have completed more online surveys in the past month have higher numeracy scores, suggesting that they are not just answering haphazardly to complete the survey. Finally the order in which the objective and subjective numeracy scales are presented remains significant in the multivariate model. This model suggests that the combined numeracy measure behaves as expected and can be used as a predictor in subsequent analyses.

Preference for Words versus Numbers

Our next question focuses on the role, if any, that numeracy may play in stated preference for having disclosure risk information presented in words or numbers. We combined the two versions of the preference question into a single measure, creating a 5-point scale with 1 being “strongly prefer words” and 5 being “strongly prefer numbers.” Overall, 46.6% somewhat or strongly preferred risk described in numbers, with 20.6% expressing a preference for words, and the balance (32.8%) expressing no preference. The order of presentation of response options had a significant effect on this preference ($\chi^2 = 142.2$, d.f = 4, $p < .0001$), with numbers being preferred more when presented first (evidence of a primacy effect). We thus include a variable in subsequent models to control for the response order.

Table 2 presents two linear regression models, the first the baseline model with only the demographic variables and the second a model with the combined numeracy scale added.

Model 1 explains a modest amount of variation (about 4%) in the preference for risk disclosure in words or numbers. Age, education, gender (all $p < .001$) and race ($p = .029$) are all significant predictors of preference. When numeracy is added in Model 2, the proportion of variance explained (R^2) increases to about 12%. The effect of age is diminished, and education, gender, and race are no longer significant. The coefficient for numeracy is positive, suggesting that higher levels of numeracy are associated with greater preference for risk presented using numbers rather than words. Similar results are found when specifying the model as an ordered logit, whether using the full 5-point preference scale or a collapsed 3-point version.

We also examined the effects of the separate numeracy measures on preferences for risk disclosure in words or numbers. Not surprisingly, given that both focus on preferences, the four-item preference subscale of the SNS performs better ($R^2 = 0.184$) than either the subjective cognitive ability SNS subscale ($R^2 = 0.066$) or the objective numeracy scale ($R^2 = 0.065$). The relationship of the other variables to preferences for numbers or words remains largely unchanged.

Effect of Numeracy on Susceptibility to Risk Manipulation

Our final analysis explores an interaction between numeracy and disclosure risk manipulation on willingness to participate in the hypothetical survey. Our expectation is that those with lower levels of numeracy are less susceptible to the risk manipulation, as the risk descriptions have less meaning for them than for those with higher levels of numeracy. Table 3 presents a summary of the F-tests and significance of the variables included in two linear regression models. The first model regresses the stated willingness to participate (WTP) in the survey described in the first vignette on a series of predictors, and is estimated using ordinary least squares (OLS) regression. The second model is a linear mixed model regressing WTP for each of the eight vignettes seen by each respondent on the same set of predictors, and is estimated using SAS PROC MIXED and IVEware (Raghunathan et al., 2001, 2005) to account for the repeated measures within persons. Two control variables are added to the second model, the first to account for learning effects across the eight vignettes (vignette order) and the second an indicator for those who gave the same answer to all eight vignettes, on the assumption that they did not pay attention to the differences among the vignettes. Removing these cases does not change the substantive conclusions. Similarly, inclusion of the two randomization variables (order of subjective and objective numeracy, and order of words versus numbers) has no effect on the models, and the variables are excluded for the sake of parsimony.

With the exception of sensitivity ($\eta^2 = 0.12$ for the first vignette), all of the effects can be regarded as small, using Cohen's (1988) criteria. However, our focus is on the effect of the numeracy scale and the preference item on WTP. When looking only at the first vignette, we find a significant effect ($p < .05$) for the preference item but not for the main effect of the numeracy scale ($p = .09$) or its interaction with disclosure risk ($p = .07$). A combined test of the addition of these two variables (numeracy and the interaction with risk) shows them to add significantly to the model ($F = 38.31$, d.f. = 4, 6355, $p < .001$), although the increase in R^2 is minimal (less than .01). However, when respondents are exposed to all eight vignettes (and hence all four risk manipulations), numeracy has a significant effect on the relationship between risk disclosure and willingness to participate. Adding the demographic controls used in the earlier models does not change this relationship. Separate analyses using only SNS or ONS also yield similar results.

To illustrate this interaction more clearly, we divided the numeracy scores into quartiles and collapsed the middle two quartiles into a single group, then reran the model for all eight vignettes. In Figure 1, we plot the least squares means (adjusted for other variables in the model) for this interaction.

The "low" numeracy group contains those in the lowest quartile of the combined standardized scale, while the high numeracy group contains those in the highest quartile, and the "medium" numeracy group represents those in the second and third quartiles combined. We can see from Figure 1 that there is a main effect of numeracy, with those exhibiting lower levels of numeracy being less willing to participate across all levels of risk. However, we see that those in the highest numeracy category are much less likely to participate in the high risk (1 in 10) survey than in the low risk (1 in 1,000,000) survey, suggesting that they are much more sensitive to the risk manipulation. Those in the lowest quartile on the

numeracy scale show a smaller difference in willingness to participate between a high risk and a low risk of disclosure. The difference in stated willingness to participate in the one-in-a-million condition between those with high numeracy (top quartile) and low numeracy (bottom quartile) is on the order of half a point on the 11-point scale.

Discussion

Our first research question addressed the relative performance of existing numeracy measures in a survey setting.

Evaluation of Existing Measures of Numeracy

We found missing data (blank responses) to be slightly higher on the ONS than on the SNS, and extra effort is required to code the open-ended responses to the ONS. The indirect (subjective) approach to measuring numeracy may be less daunting to respondents and less susceptible to mode effects. Using the SNS only, rather than the ONS or the combined scale, yields results similar to those reported here. Since the two scales yield equivalent results, they can be used interchangeably for the kinds of purposes used here. For finer-grained measurement of quantitative literacy or numeracy, a larger number of ONS measures may be preferred.

Despite the selection bias of the volunteer Internet panel in terms of better educated and literate persons, relatively low levels of performance were seen for the objective numeracy scale. Furthermore, considerable variation in scores was seen for both ONS and SNS scales, suggesting they are useful tools in discriminating among individuals with different levels of ability and comfort working with numbers.

Correlates of Numeracy

We replicated the existing literature with respect to the demographic correlates of numeracy, which suggests that older persons, those with less education, minorities, and women exhibit lower levels of numeracy. Adding the numeracy measure to models containing these demographic variables reduced many of the effects of demographics on preferences for disclosure risk expressed in words or numbers to non-significance. In other words, many of the demographic differences in preference for words or numbers can be accounted for by numeracy.

Numeracy as a Predictor of Preferred Description of Risk in Informed Consent

Numeracy is a significant predictor of preference, with more numerate persons having a greater preference for numerical descriptions of risk and less numerate individuals having a greater preference for verbal descriptions of risk.

Numeracy as a Predictor of Willingness to Participate in Surveys

Finally, we find a significant effect of numeracy on willingness to participate in a survey when respondents are exposed to multiple vignettes. Overall, those with lower levels of numeracy have lower levels of willingness to participate in the hypothetical surveys. There is also a significant interaction between numeracy and the risk manipulation, with the effect of the extreme risk conditions (1 in 10 versus 1 in 1,000,000) larger for those with higher numeracy scores, suggesting that less numerate individuals are less prone to notice or be affected by the risk described. Both of these findings suggest that an attempt to describe the size of the disclosure risk numerically in informed consent statements may be counterproductive, although we have not tested this directly.

Practical Implications

Our research indicates that numeracy plays a role in respondents' reactions to descriptions of disclosure risk in survey introductions. Since many potential survey respondents in the U.S. have low levels of literacy and low levels of numeracy, it is important to find ways to communicate this information in such a way that it accurately conveys the possible risks of participation while not discouraging people from participating. Since those potential respondents in our study with the lowest numeracy also have the lowest levels of willingness to participate at the lower levels of risk common in survey settings, our findings suggest that describing the risk in words rather than numbers may best serve to reassure a broad spectrum of potential respondents about the confidentiality of their data. We are currently experimenting with several alternative versions of such assurances.

Limitations

These results are only suggestive, however. Our sample is likely to be skewed toward those with higher levels of literacy and numeracy, and we should be cautious about generalizing to those with minimal numeracy skills. Further, our findings on the effects of numeracy and risk disclosure on willingness to participate are based on a series of hypothetical vignettes. They need to be tested in real survey settings, measuring actual participation. In addition, our experiment did not vary whether the risk was described in words or numbers, and we based our conclusions on stated preferences rather than actual behavior. Interestingly enough, a much larger percentage of our sample (46.6%) preferred to have disclosure risk described in numbers rather than words (20.6%), even though their level of objective numeracy was not very high and lower numeracy was associated with less sensitivity to large differences in disclosure risk.

Despite these limitations, our research suggests that the importance of numeracy may go beyond its role in complex risk-benefit tradeoffs in medical decision-making. More research is needed on its role in the informed consent process in social surveys as well as biomedical research.

Best Practices

Surveys and other research activities are increasingly providing detailed information to participants about the risks involved in participating in research. Our results suggest that many participants may have low levels of numeracy or quantitative literacy, and that an inability to fully understand the levels of risk described threatens the *informed* part of the consent process. This suggests that the more we try to offer precise descriptions of the generally very low risks of disclosure resulting from participation in surveys, the more difficulty some participants may have in understanding these risks and the more likely they may be to refuse altogether. Research on estimating the disclosure risks should be accompanied by efforts to better communicate such risks to potential respondents. Our results further suggest that less numerate individuals have a stronger preference for verbal—as opposed to numeric—descriptions of risk. Given the relatively low levels of numeracy in the general population, this may suggest that verbal descriptions of the risks involved in participation may be more appropriate. For example, we are currently experimenting with informed consent statements that contrast a survey organization's experience in conducting research without evidence of anyone's having been harmed as a result of disclosure with those that attempt to convey a numeric estimate of risk.

Research Agenda

Our findings are based on vignettes describing hypothetical surveys, using extremes on the risk continuum, in an online survey of volunteers. More research is needed to understand the

role that numeracy plays in decisions about participation in surveys, and to find ways to best communicate the actual risk involved in such participation in ways that potential survey respondents can understand. Specifically, more research is needed on whether numerical or verbal descriptions of risks better communicate to a broad range of participants the risks associated with survey participation. Such research should be done under “real-world” conditions with probability samples of respondents.

Educational Implications

Both researchers and ethics committee members should pay more attention to empirical evidence relevant to the ethical decisions they make. The way in which risk information is communicated to respondents may influence their decision to participate in the research, and ethics committees as well as researchers should be sensitive to the possibility that requiring one wording rather than another may lead to uninformed refusal rather than informed consent.

Acknowledgments

We thank NICHD (Grant #P01 HD045753-01) for support. We thank Market Strategies, Inc., for programming and implementing the web survey, and John Van Hoewyk for assistance with analysis.

References

- Cohen, J. Statistical power analysis for the behavioral sciences. 2. Philadelphia, PA: Lawrence Erlbaum Associates; 1988.
- Couper MP, Singer E, Conrad FG, Groves RM. Risk of disclosure, perceptions of risk, and concerns about privacy and confidentiality as factors in survey participation. *Journal of Official Statistics* 2008;24(2):1–22.
- Couper MP, Singer E, Conrad FG, Groves RM. Experimental studies of disclosure risk, disclosure harm, topic sensitivity, and survey participation. Draft paper under review. 2009
- Bergensmar M, Molin C, Wilking N, Brandberg Y. Knowledge and understanding among cancer patients consenting to participate in clinical trials. *European Journal of Cancer* 2008;44(17):2627–2633. [PubMed: 18818068]
- Donelle L, Hoffman-Goetz L, Arocha J. Assessing health numeracy among community-dwelling older adults. *Journal of Health Communication* 2007;12(7):651–665. [PubMed: 17934942]
- Fagerlin A, Zikmund-Fisher B, Ubel P, Jankovik A, Derry H, Smith D. Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making* 2007;27:672–680. [PubMed: 17641137]
- Hochhauser M. Consent comprehension in the 21st century: What is missing? *Drug Information Journal* 2008;42(4):375–384.
- Kish, L. Statistical design for research. New York, NY: John Wiley; 1987.
- Lipkus I, Samsa G, Rimer B. General performance on a numeracy scale among highly educated samples. *Medical Decision Making* 2001;21:37–44. [PubMed: 11206945]
- McNutt L, Waltermaurer E, Bednarczyk R, Carlson B, Kotval J, McCauley J, Campbell J, Ford D. Are we misjudging how well informed consent forms are read? *Journal of Empirical Research on Human Research Ethics* 2008;3(1):89–97. [PubMed: 19385786]
- Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: A new instrument for measuring patients’ literacy skills. *Journal of General Internal Medicine* 1995;10:537–541. [PubMed: 8576769]
- Raghunathan T, Lepkowski J, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology* 2001;27(1): 85–95.

- Raghunathan, T.; Lepkowski, J.; Van Hoewyk, J.; Solenberger, P. University of Michigan, Institute for Social Research; 2005. IVEware: A software for the analysis of complex survey data with or without multiple imputations. (www.isr.umich.edu/src/smp/ive)
- Raghunathan T, Reiter J, Rubin D. Multiple imputation for disclosure limitation. *Journal of Official Statistics* 2003;19:1–16.
- Reiter J. Using multiple imputation to integrate and disseminate confidential microdata. *International Statistical Review* 2009;77(2):179–195.
- Rothman R, Montori V, Cherrington A, Pignone M. Perspective: The role of numeracy in health care. *Journal of Health Communication* 2008;13:583–595. [PubMed: 18726814]
- Schwartz L, Woloshin S, Black W, Welch HG. The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine* 1997;127(11):966–972. [PubMed: 9412301]
- Singer E. Informed consent: Consequences for response rate and response quality in social surveys. *American Sociological Review* 1978;43:144–162. [PubMed: 655499]
- Singer E. Informed consent in surveys: A review of the empirical literature. *Journal of Official Statistics* 1993;9:361–375.
- Singer E, Van Hoewyk J, Neugebauer R. Attitudes and behavior: The impact of privacy and confidentiality concerns on participation in the 2000 census. *Public Opinion Quarterly* 2003;65:368–384.
- Zarate A, Zayatz L. Essentials of the disclosure review process: A federal perspective. *Journal of Empirical Research on Human Research Ethics* 2006;1(3):51–62. [PubMed: 19385823]
- Zikmund-Fisher B, Ubel P, Smith D, Fagerlin A. Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communication and utility elicitations. *Medical Decision Making* 2007;27:672–680. [PubMed: 17641137]

Biographies

Mick P. Couper is Research Professor in the Survey Research Center of the Institute for Social Research at the University of Michigan, and in the Joint Program in Survey Methodology. His research focuses on issues relating to survey nonresponse and to the use of technology in survey data collection. He is the author of *Designing Effective Web Surveys*, co-author of *Nonresponse in Household Interview Surveys* (with Robert M. Groves), and *Survey Methodology* (with Robert M. Groves and others). He is also a co-editor of *Computer-Assisted Survey Information Collection* and *Methods for Testing and Evaluating Survey Questionnaires* (with Stanley Presser and others).

Eleanor Singer is Research Professor Emerita in the Survey Research Center of the Institute for Social Research at the University of Michigan. Her research focuses on motivation for survey participation and has touched on many important ethical issues in the conduct of surveys, such as informed consent, incentives, and privacy and confidentiality. She was a member of the National Academies panels that produced *Protecting Participants and Facilitating Social and Behavioral Science Research* (2003) and *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics* (1993), and she chaired the panel whose report, *Expanding Access to Research Data*, appeared in 2006. She is most recently a co-author of *Survey Methodology* (with Robert M. Groves and others) and a co-editor of *Methods for Testing and Evaluating Survey Questionnaires* (with Stanley Presser and others) and edited a special issue of *Public Opinion Quarterly* on nonresponse bias, published in 2006.

FIG. 1.
Interaction of Categorical Numeracy Score and Risk Manipulation on WTP: All Eight Vignettes.

TABLE 1

Linear Regression of Standardized Combined Numeracy Measure (High Score = Greater Numeracy) on Selected Demographic and Panel Characteristics.

Predictor	Coefficient	Std. error
Intercept	0.47***	(0.063)
Age	−0.0063***	(0.00082)
Education		
HS or less	−0.91***	(0.032)
Some college	−0.43***	(0.026)
College graduate	—	—
Race		
White	0.11*	(0.045)
Black	−0.34***	(0.054)
Other	—	—
Hispanic origin		
Yes	−0.25***	(0.049)
No	—	—
Gender		
Male	0.36***	(0.024)
Female	—	—
Number of online surveys in past month		
<10	−0.17***	(0.028)
10–19	−0.052	(0.030)
20+	—	—
Order of numeracy scales		
ONS then SNS	−0.22***	(0.023)
SNS then ONS	—	—
R ²	0.202	
(n)	(6344)	

*
 $p < .05$,

**
 $p < .01$,

 $p < .001$

TABLE 2

Linear Regression of Preference for Words or Numbers (High Score = Greater Preference for Numbers) on Selected Demographic and Panel Characteristics and Numeracy.

Predictors	Model 1		Model 2	
	Coefficient	Std. error	Coefficient	Std. error
Intercept	3.57***	(0.085)	3.43***	(0.081)
Age	−0.0045***	(0.0011)	−0.0024*	(0.0011)
Education				
HS or less	−0.40***	(0.045)	−0.014	(0.046)
Some college	−0.23***	(0.037)	−0.052	(0.036)
College graduate	—	—	—	—
Race				
White	0.079	(0.064)	0.020	(0.061)
Black	−0.045	(0.077)	0.099	(0.074)
Other	—	—	—	—
Hispanic origin				
Yes	−0.046	(0.070)	0.062	(0.067)
No	—	—	—	—
Gender				
Male	0.098**	(0.034)	−0.056	(0.033)
Female	—	—	—	—
Order of numeracy scales				
ONS then SNS	−0.041	(0.033)	0.052	(0.032)
SNS then ONS	—	—	—	—
Order of words & numbers				
Words first	0.38***	(0.033)	0.38***	(0.032)
Numbers first	—	—	—	—
Numeracy			0.42***	(0.017)
R ²	0.039		0.124	
(n)	(6344)		(6344)	

*
p < .05,

**
p < .01,

p < .001

TABLE 3

The Effect of Numeracy, Preference for Numbers, and Key Experimental Manipulations on Willingness to Participate; Summary Results of Multivariable Linear Regression Models.

	Numerator d.f.	First vignette	All vignettes
Risk (1= one in ten, 2 = one in a million, 3 = no chance, 4 = no mention)	3	$F = 13.5, p < .0001$	$F = 112.36, p < .0001$
Sensitivity (1 = high, 0 = low)	1	$F = 844.53, p < .0001$	$F = 14671.6, p < .0001$
Prime (1 = privacy, 0 = computers)	1	$F = 38.47, p < .0001$	$F = 41.78, p < .0001$
Incentive (1 = high, 0 = low)	1	$F = 8.25, p = .0041$	$F = 12.99, p = .0003$
Mode (1 = FTF, 0 = mail)	1	$F = 7.34, p = .0068$	$F = 17.15, p < .0001$
Vignette number	1		$F = 97.68, p < .0001$
Same score	1		$F = 510.62, p < .0001$
Preference for words or numbers	1	$F = 7.26, p = .0071$	$F = 0.62, p = .43$
Numeracy	1	$F = 2.97, p = .085$	$F = 41.6, p < .0001$
Risk*numeracy	3	$F = 2.33, p = .073$	$F = 24.42, p < .0001$
Denominator d.f.		6355	44,302