Proficiency of FPPI and objective numeracy in assessing breast cancer risk estimation☆

Audrey M. Weil a,⁎, Christopher R. Wolfe a, Valerie F. Reyna b, Colin L. Widmer a, Elizabeth M. Cedillos-Wynott a, Priscila G. Brust-Renk b

a Miami University, Oxford, OH, United States
b Cornell University, Ithaca, NY, United States

A R T I C L E  I N F O
Article history:
Received 18 December 2014
Received in revised form 19 August 2015
Accepted 21 August 2015

Keywords:
Numeracy
Fuzzy-trace theory
Risk
Probability estimation
Medical decision-making

A B S T R A C T
Two studies examined the effectiveness of the Fuzzy Processing Preference Index, (FPPI) an individual differences measure of base rate neglect/respect, and an objective numeracy scale in predicting subjective probabilities of the likelihood of breast cancer, BRCA mutations, and the conditional probability of breast cancer given BRCA mutations in medical risk scenarios. FPPI and objective numeracy independently predicted estimate accuracy for breast cancer and genetic mutation risk. Surprisingly, objective numeracy positively correlated with overestimating conditional probabilities across the board, as well as BRCA mutations and breast cancer risk for high-risk scenarios. FPPI was strongest in predictions for high-risk scenarios, but did not predict conditional probability estimates. FPPI uniquely predicts risk estimation accuracy controlling for objective numeracy suggesting the two measures assess distinct cognitive processes. We conclude that FPPI and other numeracy measures may be profitably used together, and FPPI appears better than traditional numeracy measures in some medical decision-making contexts.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Subjective probability plays a vital role in our everyday decision making process (Kahneman & Tversky, 1974). Many choices we make, conclusions we reach, and actions we take are influenced by how likely we think an event might be. For instance, a woman who believes she has a high probability of developing breast cancer may be more likely to seek more frequent mammograms, and a woman who believes she is likely to carry a mutation in the BRCA1 or BRCA2 genes may be more likely to undergo genetic testing. Yet communicating numerical risk is challenging for health care providers (Brust-Renk, Royster, & Reyna, 2013). The goal of this research is to investigate factors that contribute to more accurate estimates of probabilities involving breast cancer risk.

One of the most pertinent ways in which subjective probability estimation impacts our everyday lives is when we estimate risk (Kahneman & Tversky, 1979; Weber, 1994). Assessing how dangerous a situation is or weighing the pros and cons of a choice can have life-altering consequences, particularly when the person making the decision is not well-informed. This is especially true in medical decision making where inadequate information or an inability to understand that information could lead to potentially fatal decisions (Gigerenzer & Edwards, 2003; Reyna, Nelson, Han, & Pignone, in press). Digital technologies can provide people with an estimate of their personal risk of medical illness. For instance the Breast Cancer Risk Assessment Tool on the National Cancer Institute’s (NCI) website is an interactive tool based on the Gail Model which estimates an individual’s risk of invasive breast cancer (Breast Cancer Risk Assessment Tool., 2015). This is a valuable resource, but people often fail to understand what that risk means in practical terms. Such tools are often best used in concert with medical advice or tools that help the patient understand what the bottom-line meaning of the risk is.

The importance of accurately estimating risk in a medical context is aptly demonstrated by the difficulties people encounter when assessing breast cancer risk. Breast cancer is a serious issue that affects many people today. The National Cancer Institute estimates that approximately 1 in 8 women will develop breast cancer at some point during their lives (Breast Cancer Risk in American Women, 2014), and in just the past year 40,000 women and 430 men have died from breast cancer in the United States alone (Breast Cancer, 2014). Due to recent medical advances and the availability of the internet, a lot of information about breast cancer is widely available to the public, though it is typically not presented in a format understandable to most people (Brust-
Renck et al., 2013). Despite the benefits that information about breast cancer might bring to everyday people, many people lack the ability to fully understand the information they are given. For instance, understanding their own risk for breast cancer often involves complicated processes such as interpreting base-rates, joint-probabilities (Wolfe & Reyna, 2010a, 2010b), or conditional probabilities (Wolfe, Fisher, & Reyna, 2012; Wolfe, Fisher, Reyna, & Hu, 2012) as well as comparing risks and understanding fractions, percentages, decimals, and frequencies (Reyna, Nelson, Han, & Dieckmann, 2009). Thus, the ability to understand numerical information and relationships between numbers is key in estimating medical risk.

Recent research on individual differences in judgment and decision making (Reyna et al., 2009) suggests that some people are better able than others to interpret the numerical information necessary to estimate their risk of breast cancer. The probability judgments these people make often have better correspondence to objective values (or best estimates of those values) and coherence (internal consistency) than judgments of people who are less numerate. Identifying these people and understanding what gives them this advantage has immediate implications for improving how medical information is presented, and how we can help people make informed choices about risk (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008; Reyna & Farley, 2006).

Our ability to understand and use numerical information is known as numeracy (Ancker & Kaufman, 2007; Nelson et al., 2008). Although numeracy plays an important role in estimating medical risk, national surveys estimate that about half of the population of the United States has no more than a rudimentary ability to deal with quantitative information (Reyna & Brainerd, 2007). In the medical world, patients with low numeracy are especially prone to framing and formatting effects (Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007; Peters, Hart, & Fraenkel, 2011), and overestimating their own risk of cancer (Schwartz, Wooslohin, Black, & Welch, 1997; Davids, Schapira, McUlliffe, & Natttinger, 2004). This high perception of cancer risk can then, in turn, lead to higher screening rates which may be generally beneficial so long as false positives are understood and managed appropriately (Champion, 1991; McCaul, Branstetter, Schroeder, & Glasgow, 1996; Jirojwong & Maclennan, 2003; Nelson, Huffman, Fu, & Harris, 2005). Additionally, incorrect beliefs about cancer risk can in turn lead to damaging behaviors such as worse self-care (Reyna et al., 2009; Wolf, Gazmararian, & Baker, 2005). Thus, it is critical for both a person's physical and emotional wellbeing to understand how to accurately estimate individual cancer risk.

Aside from the standard numerical pitfalls in risk assessment, there are often even more complicated numerical relationships involved in medical decision making. For example, understanding conditional probabilities, such as the chance of getting breast cancer given a genetic BRCA mutation, can further complicate the decision making process. Research has shown that even individuals with high levels of numeracy can struggle with these types of difficult concepts (Portnoy, Roter, & Erby, 2010; Peters, McCaul, Stefanek, & Nelson, 2006; Reyna et al., 2009; Wolfe, 1995; Wolfe, Fisher, Reyna, & Hu, 2012; Wolfe, Fisher, & Reyna, 2012). These difficulties further emphasize the need to help patients understand complicated information given to them in order to make informed medical choices.

Numeracy has historically been defined and measured in several different ways. However it has been argued that numeracy scales have been developed without adequate theoretical grounding (Liberalli, Reyna, Furlan, Stein, & Pardo, 2012; Reyna & Brust-Renck, 2014). When measuring numeracy in relation to how it predicts medical risk assessment, the expanded numeracy scale (Greene, Peters, Mertz, & Hibbard, 2008; Hibbard, Peters, Dixon, & Tusler, 2007; Peters et al., 2007) presents a good example of a numeracy scale that is both valid and reliable. This scale was originally adapted from the numeracy scale created by Lipkus, Samsa, and Rimer (2001), and performance on this scale has been linked to better comprehension of treatment options, hospital choices, and other health-related decisions (Peters et al., 2007). It is important to differentiate measures of numeracy like this with measures of education and intelligence. Although these variables have been shown to be related (e.g., Reyna & Brainerd, 2007), research assessing numeracy in highly educated individuals still found significant deficits in numerical decision making (Lipkus et al., 2001; Woloshin, Schwartz, Moncur, Gabriel, & Tosteson, 2000), suggesting numeracy is a unique factor in accurate medical decision making.

A more recent measure of one aspect of numeracy is the fuzzy processing preference index (FPPI; Wolfe, Fisher, & Reyna, 2013). The FPPI measures an individual’s ability to integrate quantitative base-rates and qualitative verbal information in order to estimate subjective probabilities. Wolfe and Fisher (2013) demonstrated that those participants who tended to incorporate quantitative base-rates into their judgments were more accurate in their probability estimations. Data from both laboratory experiments and web-based studies indicate that the FPPI has good psychometric properties, with Cronbach’s Alpha ranging between .91 and .96 in several experiments, indicating reliability, and validity suggested by significant correlations with “Rule Based” Process Dissociation Procedure scores; the number of conjunction fallacies in joint probability estimation; and logic index scores on syllogistic reasoning tasks (Wolfe & Fisher, 2013).

Development of the FPPI was guided by fuzzy trace theory (FTT), a dual-processing theory of judgment and decision making (Reyna, 2008; Reyna, 2012; Wilhelms & Reyna, 2013). A central tenet of FTT is that when we process information we simultaneously encode multiple traces of that information along a continuum ranging from broad, meaning-based gist traces, to detailed but superficial verbatim traces (Reyna, 2008). According to FTT people have a fuzzy processing preference in that they prefer to reason with the simplest gist representation of the information possible (Wolfe, 1995; Brainerd & Reyna, 2001; Reyna, 2012). Interestingly, research has shown that instead of an inferior system, gist processing is in many ways adaptive and a prototype of mature, expert reasoning (Reyna & Casillas, 2009; Reyna & Farley, 2006; Reyna, 1996; Reyna & Lloyd, 2006). Critically, FTT claims that successful judgment and decision making depend on forming an appropriate understanding of the gist of the information and knowing what verbatim details are important to incorporate (Reyna, 2008).

When assessing the quality of probability judgments there are two different components that contribute to the value of the judgment: correspondence and coherence. Correspondence refers to the empirical accuracy of the judgment made in relation to objective probabilities, which can be further broken down into measures of accuracy (Yates, Lee, Shintotsuka, Patalano, & Sieck, 1998) and calibration (Keren, 1991). Coherence refers to internal consistency among judgments made by the same individual and logical fallacies rather than empirical accuracy (Hammond, 2000; Wolfe, Fisher, & Reyna, 2012). Much of our previous research has focused on coherence (Fisher & Wolfe, 2011; Wolfe, Fisher, & Reyna, 2012; Wolfe, Fisher, Reyna, & Hu, 2012; Wolfe & Reyna, 2010a, 2010b) and the present work assesses probability estimation accuracy – the correspondence between subjective and objective probabilities of breast cancer risk.

The present research was embedded in two larger studies aimed at developing and testing an Intelligent Tutoring System (ITS) designed to educate people on breast cancer and genetic risk. This particular study, however, focused on the relation between different individual difference measures of numeracy and participants’ ability to estimate risk. We looked at the predictive power of both an objective numeracy scale and FPPI on accuracy in assessing the risk of breast cancer, genetic mutation, and the conditional probability of breast cancer given a mutation in 12 scenarios vetted by a medical expert. In order to measure our dependent variables, participants interacted with an ITS called BRCA Gist or one of two control conditions (see Wolfe, Reyna, Brust-Renck et al., 2014; Wolfe, Reyna, Widmer, et al., 2014 for a review of these conditions and their effectiveness).

Our first hypothesis was that high levels of objective numeracy would predict more accurate probability estimates, but that high FPPI
scores would be the best predictor of performance. We predicted that FPPI would be the better predictor due to its more narrow focus on base-rates, which can be integral to risk estimation. Our second hypothesis was that participants with high FPPI scores would be better able to use pertinent knowledge, such as the effect a BRCA mutation can have on cancer risk, to solve conditional probabilities.

2. Materials and methods

2.1. Participants

A total of 401 women participated in our two studies. Two-hundred and twenty of these participants were part of a laboratory study, and consisted of undergraduate students from a university in the Midwestern United States and a university in the Eastern United States (Wolfe, Reyna, Brust-Renck, et al., 2014; Wolfe, Reyna, Widmer, et al., 2014). Students were drawn primarily from Introductory Psychology courses, and received course credit for their participation. The remaining 181 participants were part of a field study in which participants were drawn from either a community in central New York, or through web sites including www.hispanic.com, www.facebook.com, and www.avonfoundation.org. Of these participants, 25% reported having advanced degrees, 47% had bachelor's degrees, 18% reported completing some college, 8% reported as having completed high school, and 2% did not complete high school. The majority of these participants (69%) self-identified as Caucasian, 1% as African-American, 4% as Asian-American, 6% as mixed ethnicity, 6% as other, and 14% chose not to answer. These participants were compensated for their time with a $50 gift certificate.

2.2. FPPI

Wolfe and Fisher (2013) developed the Fuzzy Processing Preference Index (FPPI) as a way to assess individual differences in participants' tendency to integrate quantitative base rates with qualitative text. They accomplished this by having participants answer 19 base-rate questions. The level to which the participant respected the base-rate was used to calculate their final score, with higher scores indicating a higher base-rate respect. For example, one FPPI item reads, “At Cloverdale High School 10% of the seniors go on to college. Bob is a senior at Cloverdale High. He gets mostly As and Bs in school and is well liked by his teachers. What is the probability that Bob will go to college?” (Wolfe & Fisher, 2013, p. 9). In studies developing the FPPI, Wolfe and Fisher (2013) found that in the absence of any base rate information, most participants (.891) identify Bob as college bound. The FPPI formula weighs the extent to which individuals rely solely on this gist of the qualitative verbal information, meaning that their estimates are close to .891, and the extent to which they are able to integrate quantitative base rate information, resulting in estimates that Bob will go to college closer to .10. In addition to the 19 base rate questions, 4 M-scale (matching strategy scale) questions were included to control for a simple matching heuristic. In these M-scale questions, the structure of the question was the same as the 19 base rate questions, but the text itself made the likelihood of the event either extremely high or extremely low. For example, “In Little Rock, Arkansas only 10% of the High School soccer referees are women. Sam has been a High School soccer referee for three years. Sam will not be refereeing this year because Sam is pregnant. What is the probability that Sam is a woman?” Estimates of exactly 10% are taken as evidence of the matching strategy. Thus these questions serve as a check that the participants are reading the question, and not employing a simple matching strategy. FPPI items are presented in the appendices of Wolfe and Fisher (2013).

2.3. Objective numeracy

For our measure of objective numeracy we used the expanded numeracy scale developed by Peters et al. (2007). This scale builds off of the numeracy scale created by Lipkus et al. (2001), but contains four additional items which assess familiarity and skill with ratios and class-inclusion relations. It consists of 15 questions that measure participants' ability to assess risk, estimate probability, convert metrics, and calculate general arithmetical computations. This scale's Cronbach's alpha is consistently around .83. In addition, its validity can be seen by its high correlation with a test of healthy literacy, the S-TOFHLA, and the fact that participants given high numeracy scores made better numerical decisions in the medical domain.

2.4. Risk assessment tools

Participants were presented with 12 different scenarios in which a woman was described with features that put her at no risk of breast cancer, a medium risk, or a high risk (Fisher et al., 2013; Wolfe, Reyna, Brust-Renck, et al., 2014; Wolfe, Reyna, Widmer, et al., 2014). These risk categories were based on scores created by the Pedigree Assessment Tool (PAT) which generates scores from 0 to 10 to represent a woman's risk of breast cancer based on various risk factors such as having an immediate familiar member with breast cancer (Hoskins, Zwaagstra, & Ranz, 2006). The scenarios we used were based on PAT scores of 0, 3–5, and 8–10 respectively (see the Appendix A for examples of low, medium and high risk scenarios). Participants categorized each scenario into a low, medium, or high risk group, and then decided if the woman should undergo testing. Paragraph length and reading difficulty were equated across trials with a word range of 56–60 per paragraph, and a Flesch Reading Ease score ranging from 56.9–62.9. Flesch-Kincaid Grade Level scores ranged from 7.3–7.9.

2.5. Procedure

2.5.1. Laboratory study

Participants were tested in groups of 1–5 on separate computers. They were randomly assigned to either a control condition or 1 of 5 versions of the experimental condition. These experimental conditions were collapsed for the purposes of this study, and will be referred to as the BRCA Gist condition. BRCA Gist was designed to help participants understand complicated factors involved in assessing breast cancer risk, genetic mutation, and the pros and cons of genetic testing (Widmer et al., in press; Wolfe et al., 2013; Wolfe, Reyna, Brust-Renck, et al., 2014; Wolfe, Reyna, Widmer, et al., 2014). The control was a different ITS that was about as effortful and time consuming as BRCA Gist, but its content was entirely on diet and exercise. After interacting with the ITS participants answered questions that included the two individual difference measures (FPPI and objective numeracy), as well as the 12 risk assessment scenarios used as the dependent measure for this analysis. At the end of the experiment, participants were debriefed and given course credit.

2.5.2. Field study

In the field study, participants were tested from the comfort of their own home. At the start of the experiment, participants were randomly assigned to one of three conditions. The experimental condition (BRCA Gist) and the control conditions were the same in the field and the laboratory study. However, this study included a second condition in which participants were presented with static versions of pages taken from the NCi’s website. In this condition, the content of the information was the same as the content in BRCA Gist, but was presented as plain text instead of through an ITS grounded in FTT. The timing of this study matched that of the other study, and the same dependent measures were included. At the end of the experiment, participants were debriefed and given $50.

2.6. Data analysis plan

We used an ANOVA to compare FPPI scores across the two locations, and then across experimental conditions to ensure there weren’t any
significant differences. These analyses were repeated for objective numeracy scores. We then ran a correlational analysis for FPPI scores and objective numeracy scores for both the laboratory and field study to assess their relationship. To ensure that both FPPI and objective numeracy scores independently contributed to evaluating risk assessment, another correlational analysis was done while controlling for shared variance between FPPI and objective numeracy in both the laboratory and field studies, and additionally between FPPI, objective numeracy, and education in the laboratory study. In previous research a G-Power analysis provided evidence that an N of 180 with a Cohen’s delta = .52 with power = .80 at alpha = .05 was sufficient to detect differences among groups. This provided the basis of determining the sample size for the present study.

3. Results

There were no significant differences between the two locations for either the FPPI or objective numeracy scores, F < 1. Additionally, there were no significant differences in FPPI scores between experimental conditions in the laboratory study, F < 1, or the field study, F < 1. Objective numeracy scores similarly did not differ between conditions in either the laboratory study, F < 1, or the field study, F < 1. Thus, FPPI and objective numeracy scores were collapsed across location and experimental condition for all subsequent analyses.

The scores for FPPI in the laboratory study and field study had means of .41 (SD = .22) and .50 (SD = .27), respectively with a possible maximum score of 1. As expected, the majority of participants in both the field study and the laboratory study scored closer to the gist end of the spectrum, replicating previous studies’ findings that most people prefer to use gist processing (Weil, 2014; Wolfe & Fisher, 2013). Objective numeracy had a mean of .66 (SD = .25) in the laboratory study and a mean of .59 (SD = .30) in the field study. These scores were more evenly distributed in both studies.

The relationship between FPPI and objective numeracy differed between the two studies. In the laboratory study, FPPI and objective numeracy were not significantly related with a correlation of r(222) = .191, p = .097. In the field study, however, FPPI and objective numeracy had a stronger, positive correlation r(179) = .300, p < .001. We also ran several analyses to test the validity of FPPI and objective numeracy as predictors for accurate risk assessment. As can be seen from the partial correlations in Table 1, the effects of FPPI and objective numeracy are still present when controlling for each other, and are not simply proxies for one another. For example, in the field study we can see in Table 1 that while controlling for numeracy we still see a significant correlation of −.154 between FPPI and risk estimates for low breast cancer risk scenarios, suggesting that FPPI is contributing variance to risk assessment beyond what numeracy can contribute.

In the field study there was an interaction between experimental condition and FPPI scores for the conditional probabilities. When considering only participants in the control condition, higher FPPI scores were associated with lower accuracy in estimating the risk of breast cancer given a genetic mutation in the low, F(2119) = 4.8820, p = .009, medium, F(2119) = 3.692, p = .028, and high risk scenarios, F(2119) = 3.505, p = .033 when compared to participants with lower FPPI scores.

Objective numeracy proved to be a strong predictor of accurate risk assessment in both studies (see Table 2 for correlations). In the laboratory study, objective numeracy specifically predicted more accurate estimates of breast cancer risk in the low risk group, and estimates of genetic mutation risk in the low, medium, and high risk groups. In the field study, objective numeracy predicted more accurate risk assessment for breast cancer and genetic mutation in the low and medium risk groups. Contrary to expectations, objective numeracy was positively correlated with risk assessment for scenarios depicting high breast cancer risk and high genetic mutation risk. If we look at Table 2 we can see that in the high breast cancer and high genetic risk scenarios numeracy was positively correlated with risk assessment with .209 and .430 respectively, both significant at the p < .01 level. This suggests that individuals with higher objective numeracy were more likely to somewhat overestimate risk for these scenarios. This relationship between objective numeracy and BRCA mutation risk estimation in the high risk scenario was in the opposite direction as the one we see in the laboratory study which may suggest limits to validity and reliability in some circumstances. Finally, objective numeracy positively correlated with risk estimates for the conditional probability of breast cancer given a genetic mutation in low, medium, and high risk groups in the laboratory study, again suggesting individuals with higher numeracy were overestimating these risks to some extent. This pattern was replicated in the field study.

Much like objective numeracy, FPPI was also a strong predictor of more accurate risk assessment in both studies. In the laboratory study, high FPPI scores predicted better risk estimates for the low and medium breast cancer risk groups. It also predicted higher accuracy for low, medium, and high genetic mutation risk estimates. Similarly, in the field study, FPPI predicted more accurate estimates of breast cancer and genetic mutation risk for low, medium, and high risk scenarios. Unlike objective numeracy, all of the significant correlations between FPPI scores and probability estimates were in the predicted direction.

### Table 1
Partial correlations between FPPI, numeracy, and education for each scenario type.

<table>
<thead>
<tr>
<th>Scenario risk level</th>
<th>Estimated probability</th>
<th>Low risk</th>
<th>Medium risk</th>
<th>High risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Breast cancer</td>
<td>BRCA mutation</td>
<td>Cancer given BRCA</td>
<td>Breast cancer</td>
</tr>
<tr>
<td>Laboratory study</td>
<td>FPPI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeracy</td>
<td>−.151</td>
<td>−.103</td>
<td>−.019</td>
<td>−.265**</td>
</tr>
<tr>
<td>Education</td>
<td>−.207</td>
<td>−.180</td>
<td>.052</td>
<td>−.311**</td>
</tr>
<tr>
<td>FPPI</td>
<td>−.243**</td>
<td>−.343**</td>
<td>.203</td>
<td>−.180</td>
</tr>
<tr>
<td>Education</td>
<td>−.166</td>
<td>−.264**</td>
<td>.217**</td>
<td>−.174</td>
</tr>
<tr>
<td>Field study</td>
<td>FPPI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeracy</td>
<td>−.154</td>
<td>−.170</td>
<td>.011</td>
<td>−.147</td>
</tr>
<tr>
<td>FPPI</td>
<td>−.263**</td>
<td>−.351**</td>
<td>.232**</td>
<td>−.086</td>
</tr>
</tbody>
</table>

Note: Variables denoted with * are the controlled for variables. *p < .05. **p < .01.
4. Discussion

The goal of the present research was to assess whether objective numeracy and FPPI predict participants’ ability to estimate risk of breast cancer and BRCA gene mutations. Consistent with Wolfe and Fisher (2013), most of the participants in both the laboratory and field study demonstrated a fuzzy processing preference in that they preferred to reason using gist representations. Put another way, most participants relied much more heavily on the verbal text when estimating subjective probabilities while placing less weight on the quantitative base-rate provided. In accordance with prior work (Peters et al., 2007; Wolfe & Fisher, 2013), both objective numeracy and FPPI predicted the accuracy of risk estimates for the majority of risk estimation scenarios. We also found that higher objective numeracy successfully predicted that participants would exhibit lower accuracy for the conditional probability of breast cancer given a genetic mutation. FPPI was unsuccessful in predicting performance for the conditional probability scenarios. However, when considering just those participants in the control group, in which participants were not given relevant information on the impact a genetic mutation can have on breast cancer, FPPI successfully predicted less accurate estimates of conditional probabilities. Thus, the ability to integrate base-rates mattered when people lacked instruction in how to integrate and estimate conditional probabilities, but, when they received such instruction, only general numerical ability mattered.

From our results we can conclude that while both FPPI and objective numeracy predict risk assessment accuracy, they are distinct constructs. Additionally, this study and previous work have shown that FPPI is independent from measures that correlate with intelligence such as education and working memory capacity (Weil, 2014; Woloshin et al., 2000). Although FPPI and objective numeracy contribute unique variance to risk assessment, the relationship between underlying cognitive processes remains less than clear. FPPI is a more tightly focused instrument whereas numeracy is a broad construct subsuming a wider range of skills, aptitudes, and experiences.

While FPPI and objective numeracy were largely similar in their ability to predict accurate risk assessments, FPPI fared better in regards to high risk scenarios and conditional probability. In the field study for both high breast cancer and high genetic mutation risk scenarios, high objective numeracy actually predicted worse risk assessment (i.e. the correlations were positive rather than negative). This implies that those participants with good numerical ability, as measured by the objective numeracy scale, are worse than participants with poor numerical ability when assessing high risk situations. One interpretation suggests that objective numeracy, with its over reliance on verbatim accuracy (Liberali et al., 2012) fails to capture vital aspects of numeracy, such as number sense or order of magnitude estimation. The factors that are captured by this measure are just a piece of the puzzle of numeracy, and needs to be augmented by other pieces such as subjective numeracy. This leads to inappropriate conclusions under some conditions. Objective research taking a more comprehensive and theoretically-motivated approach to numeracy is needed to resolve these issues. Another interpretation is that low numeracy individuals generally “guess high” when making estimates of cancer risk and BRCA mutations. This tendency is generally maladaptive but happens to serve them well when confronted with high risk scenarios. In either case the FPPI appears to be the more appropriate measure in this particular context.

One finding of particular interest was the interaction between FPPI and the experimental condition, particularly the difference between control and experimental conditions. Considering that in the BRCA Gist and NCI conditions participants with higher FPPI scores tended to rate risk much lower than participants with low FPPI scores, it is reasonable to infer that participants in the BRCA Gist and NCI conditions were estimating the risk of breast cancer given a genetic mutation as higher than they would have without relevant information given that those in the control condition did not. Of course one of the confusing things about conditional probabilities is that under conditions of independence (which may be inappropriate in real life scenarios) the likelihood of the “given” entity should have no bearing on the probability of another event given that entity. For example, if the probability of breast cancer given a BRCA 1 mutation is assessed as, say, .85 that conditional remains .85 however likely or unlikely it is for different individuals to have a BRCA 1 mutation (holding other factors constant). In conditional probabilities the relationship between the two probabilities differs quite a bit from joint probabilities, which could explain why participants’ ability to respect base-rates as measured by FPPI scores was not successful in predicting better performance for conditional probabilities.

FPPI scores for participants in the control group followed the same pattern of estimation that we see in other risk estimation scenarios (i.e. higher FPPI scores predict lower, though less accurate estimates). While the actual risk of breast cancer and the risk of a BRCA mutation are often much lower than participants expect, the conditional probability of getting breast cancer given a genetic mutation is much more likely. Therefore, what might have been a more accurate risk estimation strategy in the breast cancer and genetic mutation scenarios is less accurate when estimating conditional probabilities. From this pattern, we can infer that participants with a high FPPI score who were not instructed on the impact a BRCA mutation can have on breast cancer had the ability to assess the conditional risk accurately, but simply lacked the appropriate knowledge. Without the proper knowledge in situations involving risk assessment, simply having a high level of numeracy is not enough to predict sound judgments. This highlights the importance of conveying relevant information in an understandable format to people making decisions involving risk (Brust-Renck et al., 2013).

Future work should focus on disentangling FPPI and objective numeracy in regards to underlying cognitive processes and the mental representations they act upon. It appears that FPPI is a better estimate of some aspects of numerical literacy than the objective numeracy scale. Fuzzy-Trace Theory (Reyna, 2008) suggests that “gist numeracy” may be more important than objective numeracy in estimating risk, and the need for an individual differences measure of gist numeracy is keen. Efforts are currently underway to create a suitable instrument for assessing gist numeracy (Brust-Renck, Corbin, Weldon, Setton, &
Reyna, 2014). Finally, more work is needed to ensure that people who would score on the lower end for FPPI or objective numeracy scales receive appropriate interventions to help them make sound medical decisions based on informed judgments. Interventions to help may be focused on understanding bottom-line meaning of information as opposed to the objective parameters, or on framing the risk in terms of categorical risk.

4.1. Conclusions

Numeracy has been conceived as an individual’s basic computational skills such as multiplication, estimation, statistical literacy, graph interpretation, and facility with different representations of the same information (Ancker & Kaufman, 2007) although most tests focus on understanding ratios (Reyna & Brainerd, 2008). Although the ability to integrate quantitative base rates with qualitative verbal information clearly fits the definition, this ability has not been assessed with traditional measures of numeracy. The FPPI reliably assesses this ability and predicts medical risk accuracy even when controlling for objective numeracy.

We examined the effectiveness of the FPPI compared to and combined with a measure of numeracy, and the impact numeracy and fuzzy processing preference have on subjective probabilities estimates of breast cancer risk. This research replicates previous studies showing the benefit of numeracy for judgment and decision making in medical contexts (Peters et al., 2007; Reyna et al., 2009). By and large, FPPI held its own in comparison to the objective numeracy scale, and even performed better in cases of estimating conditional probabilities, and high breast cancer and genetic mutation risk. FPPI results are interpretable for both practical purposes and in a theoretical framework of Fuzzy-Trace Theory (Reyna, 2008). This was not consistently the case with objective numeracy, which sometimes produced surprising predictions that high numeracy would be associated with less accurate estimates. It is particularly noteworthy that the FPPI and objective numeracy each predict unique variance. Thus, both of our studies suggest that the FPPI and other numeracy measures may be profitably used in concert with one another. Using the FPPI in future research on risk assessment is clearly warranted.

Appendix A. Example risk scenarios

High risk:

“Claire is an unattatched 35 year-old New Yorker. She has a vegan diet and is an avid jogger. Her family is of Scottish-Irish heritage. Recently, her 51-year-old uncle Sean was diagnosed with cancer of the breast. Claire has several siblings and to the best of her knowledge, her uncle Sean is the only family member with breast cancer.”

Medium Risk:

“Janet is a 35-year-old Denver woman of English and Scottish ancestry. Janet is prone to kidney stones and her urologist has her on a diet low in red meat. Her 61-year-old cousin was recently diagnosed with breast cancer. No one else in the family has cancer, but her mother has diabetes. Janet herself battles obesity.”

Low Risk:

“Alegria is a 47 year-old Mexican American. She lives in Phoenix, Arizona with her husband. Alegria is a heavy cigarette smoker. Her best friend is a 10-year breast cancer survivor, but this has not been enough to get her to give up her two pack a day cigarette habit. No one in her family has ever had breast cancer.”

References
