24. *Subjective Numeracy*

KERK F. KEE,
Chapman University
& YUHUA (JAKE) LIANG,
Chapman University

Numeracy is the ability to understand numerical information. This specific ability is critical to comprehending health information as well as weighing the risks and benefits of medical treatment options. Numerical information is often used in decision aids designed to empower patients in shared decision-making, especially when they are asked to consent to a medical procedure that involves calculated risks and benefits. Patients may evaluate quantitative data such as survival likelihood, treatment success rates, complication rates, and comparative procedural statistics, among others.

However, difficulty in understanding numerical risk information can lead to treatment decisions that do not align with patients' goals. Such a misalignment can possibly lead to a lower level of perceived provider satisfaction. Physicians and patients both succeed in sharing the decision-making process only when patients can fully comprehend the medical information they receive, and then apply it to choose the treatment option that is most compatible with their personal goals and/or values.

Numeracy presents a systemic challenge for effective health communication, especially for many Americans with low numeracy. About "22% of Americans scored in the lowest 2 levels for quantitative literacy, a performance level that corresponds to having the ability to solve only single-operation arithmetic problems (e.g., what is the difference in cost between 2 items?)" (Fagerlin et al., 2007, p. 672). Furthermore, Lipkus and colleagues (2001) informed us that even populations with more formal education reported low numeracy.

Numeracy has received much research attention in health communication, medical decision-making, and public health research over the last couple of decades. Prior studies found that individuals with low numeracy are less likely to recall risk information presented in pictographs or text (Zikmund-Fisher, Smith, Ubel, & Fagerlin, 2007), misinterpret risk information in the health context (Sheridan, Pignone, & Lewis, 2003), perceive inaccurately the risk of developing cancer (Donelle, Arocha, & Hoffman-Goetz, 2008), under utilize cancer screening for early detection (Ciampa, Osborn, Peterson, & Rothman, 2010), over estimate mammography's reduction of the risk of developing breast cancer (Schwartz, Woloshin, Black, & Welch, 1997), adhere to recommended health behaviors among patients living with chronic disease (Ciampa et al., 2010), seek health information online to a lesser degree (Rakovski et al., 2012), and report lower patient-provider satisfaction (Ciampa et al., 2010). This body of literature clearly documents the detrimental effects of low numeracy on health decisions, although these studies conceptualize numeracy differently.

Previous researchers have advanced conceptual definitions for numeracy, including a person's *ability* to understand risks expressed in numbers, such as frequencies, probabilities, and percentages (Lipkus et al. 2001); *aptitude* with ratios, probabilities, and fractions (Fagerlin et al. 2007); and *difficulty* in using numbers in daily life (Ciampa et al. 2010). Framing numeracy as a difficulty may be attributed to the observation that ratio and probability concepts are often not intuitive for many individuals (Nelson, Moser, & Han, 2013). Golbeck and colleagues (2005) define numeracy as "the degree to which individuals have the capacity to access, process, interpret, communicate, and act on numerical, quantitative, graphical, biostatistical, and probabilistic health information needed to make effective health decisions" (p. 375). Morris and colleagues (2013) conceptualized numeracy as a dimension of health literacy, while Rakovski and colleagues (2012) suggest that numeracy and health literacy are empirically two distinct constructs. Based on existing research, we summarize numeracy as *the ability to understand numeric or quantitative information*. For readers interested in a more exhaustive review of numeracy, we refer them to Reyna and colleagues (2009).

Measuring Subjective Numeracy

Measurement of numeracy has followed objective and subjective approaches. These approaches both measure the underlying construct of numeracy. The objective measure relies on actual mathematical tests, tapping into individuals' abilities to reason with numbers. The objective measure has been validated and well received (see Lipkus et al., 2001). However, the current chapter focuses solely on the *Subjective Numeracy Scale* (SNS) due to certain advantages over the objective measure. We compare and contrast these advantages in the following sections. Moreover, one should note that the 8-item SNS measure (Fagerlin et al., 2007) differs from a number of past studies that have ostensibly used subjective numeracy. Some previous studies used two items from the Health Information National Trends Survey (e.g., Ciampa et al., 2010; Nelson et al., 2013; Smith, Wolf, & Wagner, 2010) to measure subjective numeracy. However, 2-item measures prevent tests for structural properties using factor analytic approaches; therefore, we limit our discussion to the 8-item SNS measure from Fagerlin et al.

The SNS measure has specific advantages to overcome certain challenges for practitioners, making it more efficient to administer across populations. First, research participants generally do not enjoy taking an aptitude test. Validated measures of objective numeracy require actual mathematical calculations. This requirement may induce a negative reaction to the research study and/or decrease study completion rates. For longitudinal studies, this reaction may also lower the retention rate of voluntary participants. Second, when administered through an online survey, researchers have little assurance that participants are performing the mathematical calculations without assistance, such as using a computing device, or asking a friend for help. Third, when administered through the telephone, the simultaneous tasks of remembering all the components of the question and performing the calculations can stimulate cognitive overload, social desirability effects, and perceived demand characteristics, resulting in an inaccurate assessment of telephone participants' numeracy. In response to these challenges, Fagerlin and colleagues (2007) developed the SNS to measure numeracy in a more efficient and less aversive fashion.

The SNS consists of two dimensions. The first four items are designed to measure individuals' beliefs about their own cognitive ability and skill in performing different mathematical operations, such as fraction and percentage calculations. The last four items measure individuals' general preferences for the presentation of numerical and probabilistic information, such as through graphs.

Table 1. Fagerlin et al.'s Subjective Numeracy Scale.

Cognitive Abilities (1=*not at all good*, 6=*extremely good*)	
1.	How good are you at working with fractions?
2.	How good are you at working with percentages?
3.	How good are you at calculating at 15% tip?
4.	How good are you at figuring out how much a shirt will cost if it is 25% off?

Preference for Display of Numeric Information	
1.	When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? (1=*not at all*, 6=*extremely*)
2.	When people tell you about the chance of something happening, do you prefer that they use *words* ("it rarely happens") or *numbers* ("there's a 1% chance"?) (1=*always prefer words*, 6=*always prefer numbers*)
3.	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 is a small chance of rain today")? (1=*always prefer percentages*, 6=*always prefer words*; reverse coded)
4.	How *often* do you find numerical information to be useful? (1=*never*, 6=*very often*)

SNS's development and validity may be examined in light of the objective measure. The objective measure tested participants' actual ability to make decisions based on numeric assessment (e.g., what is the best risk of getting a disease: 1 in 100, 1 in 1000, or 1 in 10?). The SNS measure paralleled the objective measures without using a numeric assessment. Fagerlin and colleagues (2007) found that the SNS correlated significantly with the objective measure developed by Lipkus and colleagues (2001) (r = 0.63–0.68). Furthermore, the SNS was completed in less time (24 seconds per item v. 31 seconds per item, $p < 0.05$), less stressful (1.62 v. 2.69, $p < 0.01$) and less frustrating (1.92 v. 2.88, $p < 0.01$). More participants (50%) who completed the SNS versus 8% who completed the objective scale volunteered to participate in another study (Odds Ratio = 11.00, 95% CI = 2.14–56.65).

In terms of validity, SNS has not received a full range of efforts to validate. Specifically, the scale validation and follow up studies appeared to aim to establish the SNS's criterion validity rather than assessing the overall construct validity. Although the SNS has received empirical support in terms of its correlation with the objective measure, and it has several distinct advantages for data collection as discussed in the previous paragraph, we briefly review SNS's content, criterion-related, and construct validity (Kerlinger & Lee, 1999) in the following paragraphs.

Content validity refers to the extent to which a measure covers the full range of the subject matter. The SNS began with 42 items. The final scale was reduced to 8-items. Content validity decreases with the number of items; still, the final 8 item measure includes questions related to fractions, percentages, hypothetical applications of numeracy (e.g., tipping and discount), and preference for numeric information (e.g., "how often do you find numeric informational to be useful"). This wide range generally supports SNS's content validity, although multiple items per content area would further strengthen content validity.

Criterion-related validity is the extent to which a measure compares to other variables that assess the same construct. Two subtypes of this validity include *predictive validity* and *concurrent validity*. Predictive validity refers to the ability of the measure to predict an intended outcome. Predictive validity is often established by deploying longitudinal designs, which has not been used with the SNS at the writing of this chapter. In terms of concurrent validity, there is evidence in the original validation efforts because of the measure's correlation with the objective measure (r = .68). However, data from national representative samples in the United States and Germany (Galesic & Garcia-Retamero, 2010) did not yield relevant statistics to demonstrate isomorphism between the objective and subjective measures.

Construct validity refers to the extent to which a measure solely assesses the underlying construct. The SNS was developed out of criterion-related validity to approximate objective numeracy; validation efforts have not followed traditional rigors of establishing construct validity. Traditional approaches to establish construct validity involve (1) confirmatory factor analyses to establish the structural validity and dimensionality (such as parallelism), (2) test of convergence and discrimination with the use of nomothetic networks, (3) or multi-trait multi-method approaches. The SNS has limited empirical support in construct validity. The original validation study did not provide evidence to identify whether the 8-item measure conformed to the internal structure of a unidimensional measurement model. At the time of this writing, we are unaware of any such analysis that has been conducted. Researchers who apply SNS should keep the discussed aspects of validity in mind.

Use of Existing Measures and Future Studies

A number of studies have utilized the SNS since original development, often as an individual difference measure for statistical control or as a moderator for other variables. One encouraging area of work involved validating the numeracy construct using a Rasch Analysis (Weller et al., 2013). This approach to measurement follows Item Response Theory, which involves developing question items of varying difficulty, presenting these items to test respondents sequentially, and calculating their numeracy based on the difficulty of the final test items and their performance. Weller et al.'s (2013) approach resulted in a measure that correlated substantially with the objective measure (r = .55), and equally with SNS (r = .55). This correlation provides some indirect evidence that SNS may be as useful as objective measures in assessing numeracy, assuming that the Rasch Analysis resulted in the most valid measure of numeracy.

Practically, health care providers should not assume that patients can understand the statistics and numerical information given to them. In fact, based on patients' numeracy level, providers would be well advised to take additional time to explain the quantitative information related to prescriptions and treatments in order to reduce the likelihood of unnecessary future visits due to preventable misunderstandings. Furthermore, providers can use different versions of a health message with individuals, depending on numeracy levels. For example, with low numeracy patients, a provider can use graphical displays and/or analogies (Galesic & Garcia-Retamero, 2010). By helping patients of all numeracy levels to make fully informed decisions, effective health communication can close the gap in medical interventions that are intended to improve the quality of care for patients.

Recommended Readings

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision-making, 27*, 672–680.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine, 170*, 462–468.

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine, 29*, 375–376.

Nelson, W. L., Moser, R. P., & Han, P. K. (2013). Exploring objective and subjective numeracy at a population level: Findings from the 2007 Health Information National Trends Survey (HINTS). *Journal of Health Communication, 18*, 192–205.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision-making, 27*, 663–671.

References

Ciampa, P. J., Osborn, C. Y., Peterson, N. B., & Rothman, R. L. (2010). Patient numeracy, perceptions of provider communication, and colorectal cancer screening utilization. *Journal of Health Communication, 15*, 157–168. doi: 10.1080/10810730.2010.522699.

Donelle, L., Arocha, J., & Hoffman-Goetz, L. (2008). Health literacy and numeracy: Key factors in cancer risk comprehension. *Chronic Diseases in Canada, 29*, 1–8.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision-making, 27*, 672–680. doi: 10.1177/0272989X07304449.

Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine, 170*, 462–468. doi: 10.1001/archinternmed.2009.481.

Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine, 29*, 375–376. http://dx.doi.org/10.1016/j.amepre.2005.06.012

Kerlinger, F. N., & Lee, H. B. (1999). *Foundations of behavioral research.* New York: Wadsworth.

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. doi: 10.1177/0272989X0102100105.

Morris, N. S., Field, T. S., Wagner, J. L., Cutrona, S. L., Roblin, D. W., Gaglio, B.,… Mazor, K. M. (2013). The association between health literacy and cancer-related attitudes, behaviors, and knowledge. *Journal of Health Communication, 18*(sup1), 223–241. doi: 10.1080/10810730.2013.825667.

Nelson, W. L., Moser, R. P., & Han, P. K. (2013). Exploring objective and subjective numeracy at a population level: Findings from the 2007 Health Information National Trends Survey (HINTS). *Journal of Health Communication, 18*, 192–205. doi: 10.1080/10810730.2012.688450.

Rakovski, C., Sparks, L., Robinson, J. D., Kee, K. F., Bevan, J. L., & Agne, R. (2012). A regression-based study using jackknife replicates of HINTS III data: Predictors of the efficacy of health information seeking. *Journal of Communication in Healthcare, 5*, 163–170. http://dx.doi.org/10.1179/1753807612Y.0000000014

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

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine, 127*, 966–972. doi: 10.7326/0003-4819-127-11-199712010-00003.

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*, 884–892. doi: 10.1046/j.1525-1497.2003.21102.x.

Smith, S. G., Wolf, M. S., & von Wagner, C. (2010). Socioeconomic status, statistical confidence, and patient–provider communication: An analysis of the Health Information National Trends Survey (HINTS 2007). *Journal of Health Communication, 15*, 169–185. doi: 10.1080/10810730.2010.522690.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision-making*, *26*, 198–212. doi: 10.1002/bdm.1751.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision-making*, *27*, 663–671. doi: 10.1177/0272989X07303824.