



Published in final edited form as:

*Med Care*. 2014 April ; 52(4): 354–361. doi:10.1097/MLR.0000000000000089.

## Psychometric Evaluation of the Patient Satisfaction With Logistical Aspects of Navigation (PSN-L) Scale Using Item Response Theory

**Adam Carle, PhD,**

Cincinnati Children's Hospital Medical Center & University of Cincinnati

**Pascal Jean-Pierre, PhD,**

University of Notre Dame

**Paul Winters, MS,**

University of Rochester Medical Center

**Patricia Valverde, MP,**

University of Colorado Denver

**Kristen Wells, PhD,**

University of South Florida

**Melissa Simon, MD,**

Northwestern University

**Peter Raich, MD,**

Denver Health and Hospital Authority

**Steven Patierno, Ph.D.,**

George Washington University

**Mira Katz, PhD,**

Ohio State University

**Karen Freund, MD,**

Tufts University School of Medicine

**Donald Dudley, MD, and**

University of Texas Health Science Center

**Kevin Fiscella, MD**

University of Rochester Medical Center

### Abstract

**Background**—Patient navigation; the provision of logistical, educational and emotional support needed to help patients “navigate around” barriers to high quality cancer treatment offers promise. No patient reported outcome measures currently exist that assess patient navigation from the

patient's perspective. We use a partial independence item response theory (PI-IRT) model to report on the psychometric properties of the Patient Satisfaction with Navigation, Logistical (PSN-L) measure developed for this purpose.

**Methods**—We used data from an ethnically diverse sample ( $n = 1,873$ ) from the National Cancer Institute Patient Research Program. We included individuals with the presence of an abnormal breast, cervical, colorectal or prostate cancer finding.

**Results**—The PI-IRT model fit well. Results indicated that scores derived from responses provide extremely precise and reliable measurement between  $-2.5$  SD below and  $2$  SD above the mean and acceptably precise and reliable measurement across nearly the entire range.

**Conclusions**—Our findings provide evidence in support the PSN-L. Scale users should utilize one of the two described methods to create scores.

### Keywords

Patient navigation; patient reported outcomes; reliability; partial independence item response theory

---

## BACKGROUND

Obtaining diagnostic testing and appropriate treatment following a suspicious cancer finding can be overwhelming,<sup>1,2</sup> This occurs partly because provider and patient-related barriers can hinder access and timely intervention.<sup>3,4,5</sup> Patient navigation; the provision of logistical, educational and emotional support needed to help patients to “navigate around” these barriers offers promise.<sup>6</sup> A dearth of validated patient reported outcomes (PRO) measures relevant to navigation has hindered patient navigation research.<sup>7</sup> To address this, investigators for the National Cancer Institute sponsored Patient Navigation Research Program (PNRP) collaborated to develop new PRO navigation measures.<sup>6</sup> Studies have supported these measures' reliability, validity, and use.<sup>5,8</sup>

Satisfaction with care is a meaningful outcome to patients. Satisfaction reflects the patient's judgment of quality of care received relative to expectations.<sup>9</sup> It typically includes two broad domains of perceived quality: effectiveness and interpersonal.<sup>10</sup> Patients often make decisions regarding whether to continue or change sources of care based on their satisfaction. Payers have also started to base payments on patients' satisfaction and experience with care. Satisfaction with navigators is a key measure for assessing whether navigators have addressed patients' needs from the patients' perspectives.

Yet, to our knowledge, no PROs exist that assess satisfaction with the perceived effectiveness of patient navigation and the logistical aspects of getting care. Scores from such a measure could inform navigation program evaluation and navigation training. Findings could offer insights into common patient barriers to cancer care among patients receiving navigation and how effectively these barriers are addressed. This, in turn, would improve the design and targeting of patient navigation.<sup>11</sup> In this paper, we use an item response theory (IRT) model to report on the psychometric properties of a measure

developed assessing satisfaction with navigation's logistical aspects: the Patient Satisfaction with Navigation, Logistical (PSN-L).

## METHODS

### Participants

We used data from the National Cancer Institute Patient Research Program. Previous research details the PNRP's methods.<sup>6</sup> Briefly, the PNRP involved a cooperative agreement between NCI and eight sites (with a ninth site by the American Cancer Society) to assess the benefits of patient navigation among largely minority and low-income patients with abnormal cancer findings. Inclusion criteria were the presence of an abnormal breast, cervical, colorectal, or prostate cancer finding, or a new confirmed diagnosis of any one of these. The PNRP excluded patients if they had any prior history of treatment for cancer or had received navigation previously. All patients completed IRB approved informed consent. We included all patients ( $n = 1,873$ ; 84% female) with an abnormal screen who answered at least one PSN-L item. 55.1% of patients had breast cancer diagnoses, 26% cervical; 14.1% prostate, 4.6% colorectal, and 0.3% had multiple-concurrent cancer diagnoses. 36% of the patients reported their race/ethnicity as Black/African American, 33% as Hispanic or Latino, 29% as White, and 1% reported their race/ethnicity as other.

**Scale Development**—A multidisciplinary team of PNRP investigators developed the PSN-L.<sup>8</sup> The team included researchers and clinicians with training and experience in clinical psychology, psychometrics, medical anthropology, health services research, clinical oncology, and primary care. The group began with a literature review-based conceptual model of patient navigation.<sup>11</sup> This model included four components: 1) overcoming health system barriers (e.g. insurance), 2) providing health education about cancer across the continuum from prevention to treatment, 3) addressing patient barriers, and 4) providing psychosocial support. Based on these categories, the group developed content areas. Examples include: *financial barriers* (e.g. insurance), *structural barriers* (e.g., transportation and child care), *educational and communicational barriers* (e.g., obtaining needed information) and *emotional support*. The group then used an iterative process that included multiple iterations to generate (and winnow) items for each area. The development team selected a three-point response option (detailed below) and an option for “not a problem” based on their effort to balance administrative burden and precision.

Trained research assistants orally administered the 26 PSN-L items to all participants within three months following an abnormal finding or cancer diagnosis. Depending on language preference, measures were administered in English or Spanish. The scale was introduced to patients by the following: “For each problem, indicate whether you were very satisfied (very happy), a little satisfied (happy for the most part), or not satisfied (not happy) with the help you received from your navigator(s). Some things may not apply to you. If I ask about something that was never a problem for you, just say so (not a problem) and we will skip it.” Items were “scored” 0 for “not a problem”, 1 for “not satisfied”, 2 for “a little satisfied”, and 3 for “very satisfied”.

**Analytical Plan**—Determining the PSN-L’s psychometric properties using common psychometric techniques (“classical” or “modern”) presented a problem. This occurred because the response options for each PSN-L item include a response option indicating “not a problem”. Few participants reported experiencing the entire range of problems. Thus, “not a problem” functionally results in missing data for that item. A “0” score (“not a problem”) for an item results in the patient skipping that item. While methods exist for addressing missing item responses (e.g., imputation), these methods do not address the situation that occurs with the PSN-L where one observes systematic missingness. One could conduct analyses using the responses of only patients who’d experienced each of the problems. However, even with an extremely large sample, few patients would experience the entire range of problems. And, even if one could identify a sample of patients who had experienced all problems, these patients would most likely represent a systematically different group of patients than those experiencing some but not all problems. Thus, it would be inappropriate to conduct psychometric analyses on the subgroup of patients with complete data. Though some other potential approaches exist, Reardon and Raudenbush<sup>12</sup> articulately describe the fatal weaknesses of these approaches.

Given these problems, Reardon and Raudenbush developed a partial independence IRT model (PI-IRT). A Rasch-type model, the PI-IRT delivers all of the information commonly used IRT models offer, but allows for the problem that results when patients legitimately answer some but not all items. Thus, we used the PI-IRT.<sup>12</sup> As described below, the approach re-encodes item responses to create “risk sets.” By re-encoding, one can use data from all patients who answered at least one item and estimate a patient’s level of satisfaction using the model. Space constraints preclude a detailed statistical discussion. However, given the PI-IRT’s relatively recent introduction, we briefly describe key aspects of it.

### **The Partial Independence Item Response Theory Model**

In PI-IRT, one creates a set of dichotomous items for each original item. The new dichotomous set captures the original polytomous item’s information.<sup>12</sup> In the PSN-L’s case, each polytomous item becomes a set of 3 items. The first item encodes whether the patient experienced the situation and therefore have satisfaction to report (i.e., satisfaction to report=1/no satisfaction to report=0). The second addresses whether the patient was at least a little satisfied (yes at least a little satisfied = 1 / no not at least a little satisfied = 0). The third item addresses whether the patient was very satisfied (yes very satisfied =1 / no not very satisfied =0). This leads to a total of 78 new “items.”

By re-encoding the responses one creates “risk sets” (using Reardon and Raudenbush’s<sup>12</sup> language) for each item in a set. Consider the “making medical appointments” item. All patients were “at risk” for needing help making medical appointments. However, only a subset of patients will need help making appointments and will have some type of satisfaction to report. Only these patients are “at risk” for having at least a little satisfaction. By re-encoding each item this way, one can use data from all patients who answered at least one PSN-L item, regardless of how many problems they experienced.

The PI-IRT analyses deliver a set of parameters that describe the psychometric properties of the items and scale. These include a set of conditional “threshold” parameters for each item.

They indicate how much satisfaction is present before a patient is likely to respond positively to an “item,” given that they answered any previous “items.” The PI-IRT also delivers marginal thresholds and estimates of item and total information (analogous to reliability), all of which we describe below. For all analyses we used GLLAMM<sup>21,22</sup> in Stata<sup>23</sup>, and followed Reardon and Raudenbush’s<sup>12</sup> steps.

Below, we present only the marginal threshold parameters, because one can most easily interpret the marginal parameters.<sup>12</sup> Unlike the conditional parameters, the interpretation of the marginal parameters does not depend on the responses to one or more previous items. This allows for direct interpretation. One interprets the marginal parameters relative to the latent mean of 0 (fixed for identification) and the estimated SD. They reflect the amount of satisfaction indicated by a positive response to a given category regardless of “risk”. For example, suppose the estimated SD equaled 2.17. Thus, if a patient endorsed that they feel Very Satisfied with the employer issues support their navigator gave them (and the marginal threshold equaled 2.660) this suggests that the patient would be at least about 2.660 standard deviations above the mean in satisfaction.

### Evaluating Model Fit

The validity (and thus interpretability) of parameters that result from the PI-IRT model depend on the extent to which the model fits the data.<sup>12</sup> Given the PI-IRT model’s differences from other latent variable measurement models, one cannot use these models’ fit indices to evaluate the PI-IRT. This limits the ability to evaluate unidimensionality and other assumptions in the same manner one would use for other latent variable measurement models.<sup>13</sup> However, Reardon and Raudenbush propose two methods based on the ability of the model parameters to predict responses.<sup>12</sup> Method 1 depends on predicting the conditional and marginal probability of a response for each person, averaging the individual probabilities across the sample, and then comparing the predicted conditional and marginal probabilities to the observed conditional and marginal probabilities respectively. Method 2 depends on integrating the estimated marginal probability over the estimated distribution of Satisfaction. This allows one to generate another estimate of the conditional and marginal probabilities. Then, one compares the predicted conditional and marginal probabilities to the observed conditional and marginal probabilities respectively. One uses both methods to evaluate fit. To the extent that the methods indicate good fit, one has more confidence that the data do not violate the model’s assumptions dramatically. However, as Liu and Verkuilen<sup>13</sup> note, this requires subjective judgment and limits the PI-IRT model.

## RESULTS

### Model Fit

Space constraints limit our ability to present results for each item (available upon request). However, in summary, the differences between predicted and observed tend to be quite small and negative (conditional: mean = -0.002, SD = 0.002; marginal: mean < -0.001, SD = 0.006) for Method 1. The negative value indicates the model tends to slightly underestimate observed conditional and marginal response probabilities. For Method 2, again, the differences between predicted and observed tend to be small and negative

(conditional: mean =  $-0.04$ , SD =  $0.03$ ; marginal: mean =  $-0.08$ , SD =  $0.02$ ). These findings demonstrate that the model fits the data well, providing support for the model's assumptions (e.g., dimensionality) and the resulting parameters and scores estimated using the model.

## Model Results

Table 1 summarizes the estimated marginal threshold parameters. The marginal thresholds ranged from  $-2.21$  to  $3.34$ , with a mean of  $0.15$  and SD of  $1.4$ .

Using the model's parameters, one can generate figures that visually display how response probabilities change with increasing levels of satisfaction. Figure 1 illustrates the estimated marginal item characteristic curves (M-ICCs) for two items: "Satisfaction with Emotional Support" and "Satisfaction with Work Issues." With respect to support, Figure 1 shows that patients become more likely than not to indicate they are at least a little satisfied (LS) with the navigator's help once their satisfaction with navigator generally is greater than one SD below the mean. For Work Issues, one sees a similar M-ICC. Figure 1 shows that it takes considerably more satisfaction with the navigator before patients are more likely than not to respond that they are very satisfied (VS) with the navigator's help with either emotional or work issues. For either of these items, patients are likely to be at least 1.5 SDs above mean satisfaction levels if they have endorsed they are very satisfied. Space constraints preclude presenting the entire set of M-ICCs (available upon request).

Figure 2 presents the distribution of estimated Satisfaction (bars in the figure) in the sample, as well as the model-based population distribution (solid line in the figure). The population distribution is based on a mean of 0 and the model estimated standard deviation (SD) of 2.17. Because the bars closely follow the solid line, Figure 2 shows that the sample and population distributions have similar shapes and that few specific levels of satisfaction are far over or under represented in the sample.

Figure 3's left panel depicts the estimated precision of latent Satisfaction scores. By reading values on the left-side y-axis, one can see the standard error of the estimate (a measure of precision) in the metric of the latent Satisfaction scores. By reading values on the right-side y-axis, one can see the standard error of the estimate expressed in terms of the standard deviation (SD) of latent Satisfaction scores. Low standard error estimate values correspond to relatively more precise (i.e., reliable) measurement. Larger values correspond to less precise (i.e., less reliable) measurement. Figure 3 shows that the PSN-L provides quite precise measurement between levels of about  $-2.5$  and  $2$ . And, although it provides less precise measurement at values less than  $-2.5$  and greater than  $2$ , it is only at the extremes that the standard error of measurement becomes larger than 1. This indicates that the PSN-L provides its most precise and reliable measurement between  $-2.5$  SD below the mean and 2 SD above the mean but that it provides acceptably precise and reliable measurement across nearly the entire range.

Figure 3's right panel visually displays the total information the PSN-L provides. Information corresponds to reliability. Unlike classical test theory, which gives a single estimate of reliability (e.g., Cronbach's alpha)<sup>14</sup> that unrealistically<sup>15</sup> applies across the entire spectrum of Satisfaction, IRT allows reliability to vary across trait levels. One would

expect that a measure would be more reliable at different levels of the trait, as some items will likely measure levels of the latent trait better than others. The monotonic peak shows that the PSN-L provides the most reliable measurement at mean levels of Satisfaction. This level of information corresponds to a marginal Cronbach's alpha of approximately 0.86,<sup>16</sup> which is considered a good level of reliability.<sup>17</sup> Additionally, the figure shows that the PSN-L provides acceptably good reliability<sup>18</sup> between levels of satisfaction 1 SD below and above the mean (alpha approximately 0.80). Even at 1.5 SD above and below the mean, responses to the PSN-L still appear to offer sufficient reliability (alpha approximately 0.70). Beyond 2 SD above and 2.5 below the mean, reliability begins to deteriorate rapidly. Nonetheless, across much of the Satisfaction spectrum, the PSN-L provides acceptable reliability. For readers less familiar with IRT's approach to reliability, please refer to Embretson and Reise.<sup>15</sup>

Table 2 presents the summed-score to EAP translation table.

## DISCUSSION

In this study, we sought to describe the psychometric properties of the PSN-L, a new scale developed to measure a patient's satisfaction with a medical navigator's help with logistical issues related to care. We used a partial independence item response theory model (PI-IRT) to do this. The scale appears to measure satisfaction with a navigator most precisely and reliably from 2 and a half standard deviations below the mean to about 2 standard deviations above the mean. This range included 75% of the patients in our sample. Although it provides the most reliable measurement in this range, it provides acceptable reliability across a wide range of satisfaction. Our work suggests that stakeholders can use responses to the PSN-L scale to measure navigator satisfaction.

The PSN-L scale offers a practical way for patient navigation programs to assess patients' navigation experiences. It includes items relevant to the most common logistical aspects of navigation and could thus be used to identify areas targeted for improvement. One could also use it to improve the targeting of patient navigation based on a patient's expected need for navigation. It could also be used to track the effects of additional navigation training or compare patient ratings across navigators. It could also be used in research studies seeking to understand the pathways through which navigation succeeds in achieving its goals. The PSN-L focuses on logistical aspects of navigation. The PSN-Interpersonal (described elsewhere)<sup>19</sup> focuses on the interpersonal aspects of navigation. Used together, the two scales (which do not include overlapping items) provide a reasonably complete assessment of navigation from the patient's perspective.

Given that the PI-IRT fit the data well and that satisfaction scores based upon the PI-IRT parameters provide reliable and internally valid estimates of patient satisfaction, the question becomes how should one score the PSN-L in the "real" world. As noted earlier, this is not a trivial question. Not all patients will experience all of the problems asked about on the PSN-L, which limits the utility of a simple sum approach. In addition, a simple sum approach directly implies that responses to different categories across different items indicate the same levels of Satisfaction regardless of the question. For example, this would mean that

indicating that one is very satisfied with the help one received being encouraged to talk to the doctor reflects the exact same amount of satisfaction one feels when one is very satisfied with help one received with problems related to work. And, finally, the simple sum approach does not allow a user to assess confidence in a patient's score. Generating scores based on the PI-IRT's parameters addresses all of these issues.

We propose one of two methods for deriving a score based on the PI-IRT model parameters. Scoring Method 1 is relatively simply and uses a patient's item responses and a PI-IRT-based table (Table 2) to derive a score in a latent "Satisfaction with Logistical Aspect of Navigation" metric. For Method 1, one first scores individual items responses in the following manner: not a problem = 0, not satisfied = 1, a little satisfied = 2, very satisfied = 3. One then takes the sum of the item responses across all 26 items. Using Table 2, one uses the summed score to estimate a patient's latent Satisfaction score. With the same method, one can assign a standard error to the patient's score (final column). In this way, "missing" takes a value of 0 and one can estimate a score for all individuals regardless of the number of problems for which the individual received help.

Scoring Method 2 uses a patient's item responses, the parameters from the PI-IRT model, and statistical software to estimate the "Navigator Satisfaction" score and its associated standard error directly. For example, using GLLAMM and Stata, one would input the thresholds and standard deviation described above into GLLAMM using matrices, treating the thresholds, mean, and SD as known. Then one would use GLLAMM's post-estimation commands to estimate scores for a patient or patients, using the observed responses converted to a series of dichotomous item responses as described above.

Both scoring methods provide a more reliable and valid estimate of a patient's satisfaction levels relative to a simple sum or other methods (e.g., single item methods). Both methods allow users to estimate confidence in the patient's score. The two methods correlate highly (0.98). And, both methods allow one to work with all individuals, regardless of how many problems an individual experienced. While Scoring Method 2 will provide the most precise Satisfaction estimate, we suspect that most users will use Scoring Method 1, given that it still provides a more precise estimate than other methods and that many users will have limited access to advanced statistical software.

## Limitations

Before concluding, we note some limits of our study. First, although we used wording based on principles of plain communication and pilot testing did not suggest problems with understanding, we did not conduct cognitive testing. Thus, we do not have data to support our use of happy as a synonym for satisfaction when introducing the questions to respondents nor can we be sure that patients understood other questions and concepts as expected. Second, we did not directly include patients during item development. Third, as Liu and Verkuilen note,<sup>13</sup> like any statistical method (simple sum or otherwise), the PI-IRT requires assumptions. These include a proportional odds assumption and (like other Rasch-type models), unidimensionality, and the assumption that all items share a common discrimination parameter. Though the results of our model provided good fit to the data, supporting the expectation that the data meet the model's assumptions (e.g.,



unidimensionality), the assumptions may be overly strict and lead to some misestimation. In addition, the possibility exists that multidimensional data can provide acceptable fit in a unidimensional model. Nevertheless, while the PI-IRT model does make assumptions, they are far more realistic than the assumptions underlying a simple score approach (e.g., that the items measure the construct perfectly without error, etc.). Finally, by definition, the PSN-L measures satisfaction from the patient's perspective. Given reliance on the patient's perspective only, it is possible that patients do not accurately or honestly report their experiences. Despite these limitations, we feel our study's strengths outweigh its weaknesses and that the PSN-L represents a useful complement to the PSN-I.

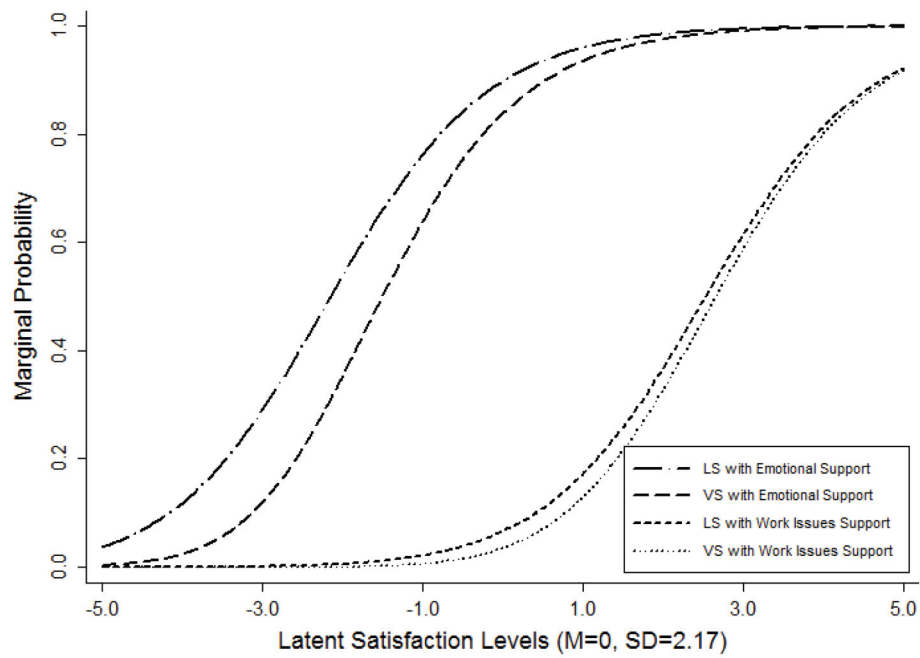
## Conclusions

The PSN-L addresses a gap in tools needed to assess patient navigation. Our psychometric findings support the PSN-L as a reliable and internally valid measure of satisfaction with a navigator. Scale users should utilize one of the two methods described in this paper to create a PSN-L score.

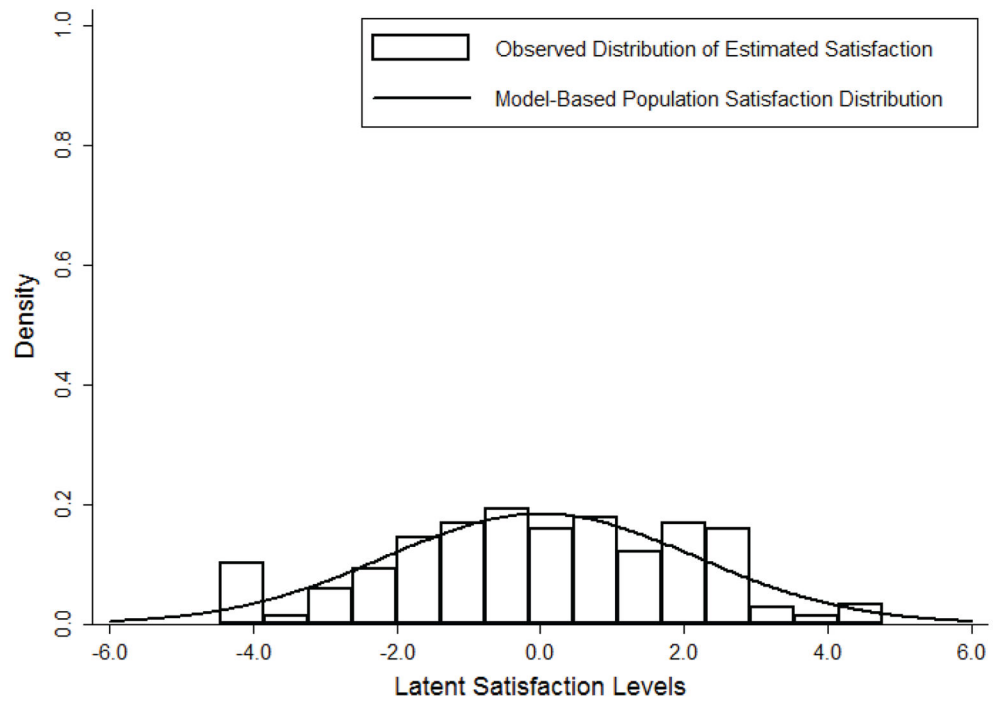
## References

- Hawley ST, Janz NK, Lillie SE, et al. Perceptions of care coordination in a population-based sample of diverse breast cancer patients. *Patient Education and Counseling*. 2010; 81:S34–S40. [PubMed: 21074963]
- Janz NK, Mujahid MS, Hawley ST, Griggs JJ, Hamilton AS, Katz SJ. Racial/ethnic differences in adequacy of information and support for women with breast cancer. *Cancer*. 2008; 113(5):1058–1067. [PubMed: 18618494]
- Bickell NA, Dee McEvoy M. Physicians' reasons for failing to deliver effective breast cancer care: a framework for underuse. *Medical Care*. 2003; 41(3):442. [PubMed: 12618647]
- Caplan LS, Helzlsouer KJ, Shapiro S, Wesley MN, Edwards BK. Reasons for delay in breast cancer diagnosis. *Preventive medicine*. 1996; 25(2):218–224. [PubMed: 8860288]
- Jean-Pierre P, Fiscella K, Freund KM, et al. Structural and reliability analysis of a patient satisfaction with cancer-related care measure. *Cancer*. 2011; 117(4):854–861. [PubMed: 20922802]
- Freund KM, Battaglia TA, Calhoun E, et al. National cancer institute patient navigation research program. *Cancer*. 2008; 113(12):3391–3399. [PubMed: 18951521]
- Fiscella K, Ransom S, Jean-Pierre P, et al. Patient-reported outcome measures suitable to assessment of patient navigation. *Cancer*. 2011; 117(S15):3601–3615.
- Jean-Pierre P, Fiscella K, Winters PC, et al. Psychometric development and reliability analysis of a patient satisfaction with interpersonal relationship with navigator measure: a multi-site patient navigation research program study. *Psycho-Oncology*. 2012; 21(9):986–992. [PubMed: 21681995]
- Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. *Social science & medicine*. 1997; 45(12):1829–1843. [PubMed: 9447632]
- Murakami G, Imanaka Y, Kobuse H, Lee J, Goto E. Patient perceived priorities between technical skills and interpersonal skills: their influence on correlates of patient satisfaction. *Journal of Evaluation in Clinical Practice*. 2010; 16(3):560–568. [PubMed: 20438604]
- Wells KJ, Battaglia TA, Dudley DJ, et al. Patient navigation: state of the art or is it science? *Cancer*. 2008; 113(8):1999–2010. [PubMed: 18780320]
- Reardon SF, Raudenbush SW. A Partial Independence Item Response Model for Surveys with Filter Questions. *Sociological Methodology*. 2006; 36(1):257–300.
- Liu Y, Verkuilen J. Item Response Modeling of Presence-Severity Items Application to Measurement of Patient-Reported Outcomes. *Applied Psychological Measurement*. 2013; 37(1): 58–75.

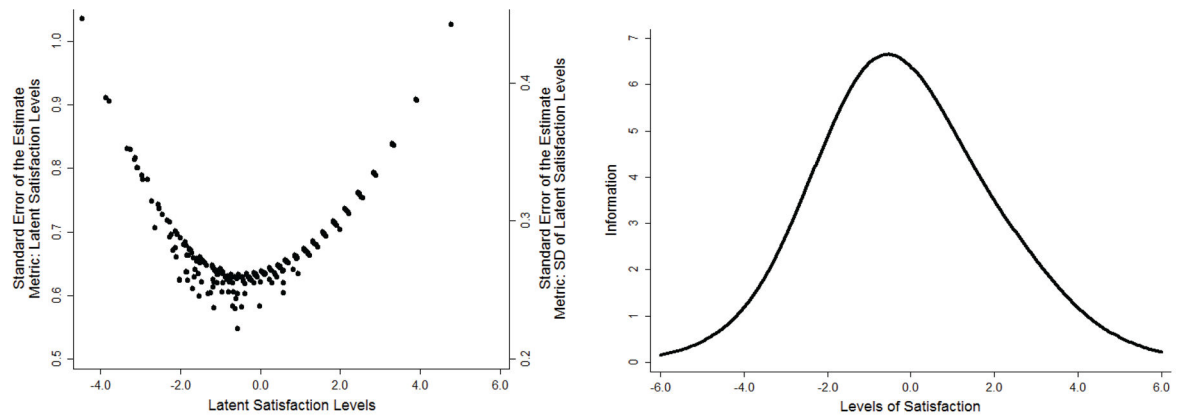
14. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951; 16(3):297–334.
15. Embretson, S.; Reise, S. *Item response theory for psychologists*. Mahwah, N.J: Lawrence Erlbaum; 2000.
16. Green BF, Bock D, Humphres RL, Linn MD. Technical Guidelines for Assessing Computerized Adaptive Tests. *Journal of Educational Measurement*. 1984; 21(4):347–360.
17. Crocker, L.; Algina, J. *Introduction to classical and modern test theory*. Holt, Rinehart and Winston; 6277 Sea Harbor Drive, Orlando, FL 32887: 1986. (\$44.75)
18. Nunnally, JC.; Bernstein, I. *Psychometric theory*. Vol. 2d. New York: McGraw-Hill; 1994.
19. Jean-Pierre P, Fiscella K, Winters PC, et al. Psychometric development and reliability analysis of a patient satisfaction with interpersonal relationship with navigator measure: a multi-site patient navigation research program study. *Psycho-Oncology*. 2011



**Figure 1.** Marginal Item Characteristic Curves for the Emotional Support and Satisfaction with Work Issues items (LS = “A little satisfied”; VS = “Very satisfied”).



**Figure 2.**  
Observed and model-based distribution of Satisfaction.



**Figure 3.** PSN-L Precision. Left panel: Assessing the precision of estimated Satisfaction using the SEM (higher SEM equals less precise measurement). Right Panel: Total information provided by the PSN-L across Satisfaction levels.

Table 1

Estimated Marginal Partial Independence Item Response Theory Thresholds (see the Results section for a discussion of how to interpret these parameters).

		Estimated Marginal Thresholds
1	Not a problem-Not Satisfied	-0.090
	Not satisfied-A Little Satisfied	-0.044
	A Little- Very Satisfied	0.159
2	Not a problem-Not Satisfied	-1.263
	Not satisfied-A Little Satisfied	-1.185
	A Little- Very Satisfied	-0.821
3	Not a problem-Not Satisfied	-0.337
	Not satisfied-A Little Satisfied	-0.220
	A Little- Very Satisfied	0.068
4	Not a problem-Not Satisfied	-0.131
	Not satisfied-A Little Satisfied	-0.005
	A Little- Very Satisfied	0.460
5	Not a problem-Not Satisfied	0.817
	Not satisfied-A Little Satisfied	0.864
	A Little- Very Satisfied	1.024
6	Not a problem-Not Satisfied	-1.171
	Not satisfied-A Little Satisfied	-1.061
	A Little- Very Satisfied	-0.527
7	Not a problem-Not Satisfied	-2.211
	Not satisfied-A Little Satisfied	-2.145
	A Little- Very Satisfied	-1.499
8	Not a problem-Not Satisfied	-1.957
	Not satisfied-A Little Satisfied	-1.894
	A Little- Very Satisfied	-1.414
9	Not a problem-Not Satisfied	-0.841
	Not satisfied-A Little Satisfied	-0.782

		Estimated Marginal Thresholds
	A Little-Very Satisfied	-0.271
	Not a problem-Not Satisfied	-1.850
10	Getting the health information you needed	-1.708
	A Little-Very Satisfied	-1.332
	Not a problem-Not Satisfied	-0.734
11	Making you more involved in decisions about your care	-0.617
	A Little-Very Satisfied	-0.152
	Not a problem-Not Satisfied	-0.029
12	Dealing with personal problems related to your health	0.037
	A Little-Very Satisfied	0.484
	Not a problem-Not Satisfied	2.515
13	Dealing with work or employer issues related to health care	2.535
	A Little-Very Satisfied	2.660
	Not a problem-Not Satisfied	-0.670
14	Understanding the medical tests you got	-0.538
	A Little-Very Satisfied	-0.106
	Not a problem-Not Satisfied	-0.971
15	Understanding your health issues	-0.842
	A Little-Very Satisfied	-0.383
	Not a problem-Not Satisfied	-1.772
16	Knowing who to call when you had a question	-1.632
	A Little-Very Satisfied	-1.233
	Not a problem-Not Satisfied	-1.547
17	Learning about community services available to you	-1.319
	A Little-Very Satisfied	-0.815
	Not a problem-Not Satisfied	2.576
18	Dealing with housing and landlord issues	2.594
	A Little-Very Satisfied	2.722
19	Dealing with paperwork	0.565

		Estimated Marginal Thresholds
	Not satisfied-A Little Satisfied	0.637
	A Little-Very Satisfied	0.836
	Not a problem-Not Satisfied	0.216
20	Understanding letters, reports, and health education materials	0.233
	A Little-Very Satisfied	0.515
	Not a problem-Not Satisfied	-0.316
21	Getting community services you are eligible for	-0.137
	A Little-Very Satisfied	0.196
	Not a problem-Not Satisfied	3.273
22	Getting child or eldercare so you can go to Dr's	3.285
	A Little-Very Satisfied	3.340
	Not a problem-Not Satisfied	0.753
23	Dealing with health insurance matters	0.803
	A Little-Very Satisfied	1.019
	Not a problem-Not Satisfied	0.774
24	Including family members in the care you received	0.804
	A Little-Very Satisfied	0.947
	Not a problem-Not Satisfied	2.040
25	Dealing with providers who do not speak your language	2.054
	A Little-Very Satisfied	2.204
	Not a problem-Not Satisfied	2.067
26	Overcoming barriers related to physical disability	2.086
	A Little-Very Satisfied	2.267



**Table 2**

## Values for Scoring Method 1

Score	Approximate Satisfaction	Approximate SD
0	-3.31	0.55
1	-3.09	0.51
2	-2.88	0.47
3	-2.7	0.44
4	-2.54	0.41
5	-2.4	0.39
6	-2.26	0.37
7	-2.14	0.35
8	-2.03	0.34
9	-1.92	0.33
10	-1.82	0.32
11	-1.73	0.31
12	-1.64	0.3
13	-1.55	0.29
14	-1.47	0.28
15	-1.39	0.28
16	-1.32	0.27
17	-1.25	0.27
18	-1.17	0.26
19	-1.11	0.26
20	-1.04	0.25
21	-0.97	0.25
22	-0.91	0.25
23	-0.85	0.24
24	-0.79	0.24
25	-0.72	0.24
26	-0.67	0.24
27	-0.61	0.23
28	-0.55	0.23
29	-0.49	0.23
30	-0.44	0.23
31	-0.38	0.23
32	-0.32	0.23
33	-0.27	0.23
34	-0.22	0.22
35	-0.16	0.22
36	-0.11	0.22

Score	Approximate Satisfaction	Approximate SD
37	-0.05	0.22
38	0	0.22
39	0.06	0.22
40	0.11	0.22
41	0.16	0.22
42	0.22	0.22
43	0.27	0.22
44	0.33	0.22
45	0.38	0.22
46	0.44	0.22
47	0.49	0.23
48	0.55	0.23
49	0.61	0.23
50	0.67	0.23
51	0.72	0.23
52	0.78	0.23
53	0.84	0.23
54	0.9	0.24
55	0.97	0.24
56	1.03	0.24
57	1.09	0.24
58	1.16	0.25
59	1.23	0.25
60	1.3	0.25
61	1.37	0.26
62	1.45	0.26
63	1.52	0.27
64	1.6	0.27
65	1.69	0.28
66	1.77	0.28
67	1.86	0.29
68	1.96	0.3
69	2.06	0.31
70	2.17	0.32
71	2.28	0.33
72	2.4	0.34
73	2.54	0.35
74	2.68	0.37
75	2.84	0.38

<b>Score</b>	<b>Approximate Satisfaction</b>	<b>Approximate SD</b>
76	3.02	0.39
77	3.22	0.39
78	3.43	0.37