Chapter 13

Service Satisfaction Scale

(SSS)

Background

Purpose
The Satisfaction with Services Scale (SSS) is a measure of youth and adult caregiver satisfaction with the treatment they receive. Service satisfaction has been defined at multiple levels. Broadly, it can be defined as “a health care recipient’s reaction to salient aspects of the context, process, and results of their…experience” (Pascoe, 1983, p. 189). A more specific definition found in the “discrepancy approach” defines satisfaction as “the emotional response to the judgment of difference between perceived services and consumers’ desired and expected services” (Gerkenesmeyer, Austin, & Miller, 2006, p.66). The SSS specifically asks youth and caregivers to report on whether they perceive their needs and preferences were met. Satisfaction with services represents one type of treatment outcome and thus can be thought of as an indicator of treatment process.

Theory
Satisfaction with services has long been seen as a potentially important nonclinical outcome. Many managed care companies consider satisfaction to be a helpful, high priority outcome (Lambert, Salzer, & Bickman, 1998). Likewise, some assert that satisfaction is a critical indicator of whether needs were met (Burnam, 1996). Despite this, research on consumer satisfaction suggests that it is at best mildly related to symptom change (e.g., Lambert et al., 1998; Luk et al., 2001; Lunnen & Ogles, 1998). Because consumer satisfaction is not predictive of clinical outcomes, the use of a lengthy measure of satisfaction with services was deemed unnecessary. However, because many managed care organizations mandate the use of service satisfaction indices, the Satisfaction with Services Scale was developed to assess service satisfaction in an efficient manner. The SSS is not intended to be used as a measure of clinical outcomes or treatment effectiveness, but is solely a measure of client and adult caregiver service satisfaction.

History of Development
The development of the SSS began with a review of the relevant literature on satisfaction with mental health services. At the most basic level, these measures assess whether youth and/or adult caregivers perceive their needs to have been met by treatment. Based on the discrepancy approach to satisfaction, whether preferences or desires for services are met is an additional component of satisfaction. Because of this, client and adult caregiver perception of their needs and preferences being met were targeted as important item content.
The scale tested was a five-item version that included four close-ended items (Likert-scale) and one open-ended item that asked respondents for any optional comments they may have about services. One of the close-ended items was dropped due to poor item psychometrics. After further testing and review of the literature, we identified an additional item that could strengthen the validity of the measure, an item about the degree to which youths or caregivers thought the services they received “took the right approach.” The current version of the PTPB includes this revised, five-item measure, with four close-ended items and the single open-ended item.

Structure

The SSS measures client and caregiver satisfaction with mental health services using five self-report items. Four items are rated on a four-point Likert-type scale ranging from one (No, definitely not) to four (Yes, definitely). In addition to the four Likert-scale items, there is a free-response item asking for other comments about the services received.

It should be noted that only four of the five items were evaluated as part of the psychometric study (see Chapter 2), with results reported below. The three items are: (1) did you get the kind of services you think you needed; (2) if a friend were in need of similar help, would you recommend our services to him or her; (3) if you were to seek help again, would you seek it from us; and (4) were the services you received the right approach for helping you. The free response item does not produce any quantitative output.

The SSS Total Score describes overall satisfaction with services. This total score is calculated by taking the average item total for the four Likert-type items. The psychometrics presented in this chapter are based on the complete psychometric study sample with the four Likert-type items noted above, and the analyses presented required 85% of the items to have valid answers. See Chapter 2 for more detail on the psychometric sample and test development procedures.

Administration

The SSS should be administered during the treatment and discharge phases and is to be completed by both the youth and the adult caregiver as indicated in Table 13.1. The suggested frequency of administration is every two months.

Table 13.1 Administration of SSS

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Treatment</th>
<th>Discharge</th>
<th>Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>A</td>
<td>C</td>
<td>Y</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>●</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Y = Youth (age 11-18); A = Adult Caregiver; C = Counselor
Recommended Frequency: Every two months
The suggested administration schedule of all the measures in the Peabody Treatment Progress Battery is presented in Appendix A. All PTPB measures with self-scoring forms can be found in Appendix B.

**Description**

**Basic Descriptives**

The descriptive statistics for the SSS appear in Table 13.2. They show fairly high satisfaction for the psychometric sample, but not so high as to restrict the variance by ceiling artifacts. Results suggest some variation in both adult caregiver and youth responses to each item.

**Table 13.2  Descriptive Statistics for SSS Summary Scores**

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSS Total Score – Adult Caregiver</td>
<td>383</td>
<td>3.59</td>
<td>0.51</td>
<td>-1.47</td>
<td>2.73</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>SSS Total Score – Youth</td>
<td>490</td>
<td>3.37</td>
<td>0.66</td>
<td>-1.30</td>
<td>1.87</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

**Quartiles**

High scores are those in the top quarter, with low scores in the bottom quarter as presented in Table 13.3. For youth, a score less than 3.00 is considered low, while for adult caregivers, a score less than 3.25 is considered low. For youth and adult caregivers, scores in the top range are blurred, with the highest score of 4.00 at the third quartile. This failure to distinguish upper scores shows a ceiling effect limitation, reflected in the comprehensive psychometric analysis presented in Table 13.6.

To aid interpretation, the quartiles were used to create low, medium, and high scores and percentile ranks based on comparison to the psychometric sample. This information is presented in the last section of this chapter. It should be noted that the restricted range of responses impacts both types of comparisons.

**Table 13.3  SSS Quartiles**

<table>
<thead>
<tr>
<th>Quartile</th>
<th>SSS Total Score – Youth</th>
<th>SSS Total Score – Adult Caregivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% Max</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>75% Q3</td>
<td>4.00</td>
<td>4.00</td>
</tr>
<tr>
<td>50% Median</td>
<td>3.50</td>
<td>4.75</td>
</tr>
<tr>
<td>25% Q1</td>
<td>3.00</td>
<td>3.25</td>
</tr>
<tr>
<td>0% Min</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Evidence of Reliability

Reliability Coefficients
The Cronbach’s alpha internal consistency reliability correlations are presented in Table 13.4. These alphas suggest a satisfactory degree of internal consistency for the total score for both youth and adult caregiver.

Table 13.4 Cronbach’s Alphas for the SSS

<table>
<thead>
<tr>
<th>Scale</th>
<th>Unstandardized Alpha</th>
<th>Standardized Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSS-Adult Caregiver</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>SSS-Youth</td>
<td>0.86</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Comprehensive Item Psychometrics
Table 13.5 presents the comprehensive item psychometrics. Shaded cells indicate that a criterion was out of the range of sought values, as described previously in Table 2.2 in chapter two. Only items with two or more shaded cells are considered problematic. The adult version shows typical problems with satisfaction scales. The average item scores are relatively close to the maximum value of the scale range indicating a ceiling effect. Most adult caregivers reported to be very satisfied with the services offered to them. Since this is true for all four items, the combination of item answers can be used as scale (high item-total correlations) but it is not a very efficient scale in differentiating clients with different levels of satisfaction. This scale can mainly be used to identify clients and their caregivers who are not very satisfied with the services they received.
<table>
<thead>
<tr>
<th>Version</th>
<th>Item</th>
<th>N</th>
<th>Mean</th>
<th>St Dev</th>
<th>Kurtosis</th>
<th>Item-Total</th>
<th>Std CFA Loadings</th>
<th>Measure</th>
<th>Infit</th>
<th>Outfit</th>
<th>Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youth</td>
<td>Services were right approach</td>
<td>490</td>
<td>3.32</td>
<td>0.77</td>
<td>1.01</td>
<td>0.72</td>
<td>0.79</td>
<td>0.3</td>
<td>0.91</td>
<td>0.88</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>Received needed services</td>
<td>490</td>
<td>3.35</td>
<td>0.75</td>
<td>1.42</td>
<td>0.7</td>
<td>0.77</td>
<td>0.12</td>
<td>0.97</td>
<td>0.95</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>Seek help from us again</td>
<td>490</td>
<td>3.41</td>
<td>0.8</td>
<td>1.58</td>
<td>0.73</td>
<td>0.8</td>
<td>-0.2</td>
<td>1</td>
<td>0.92</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>Would recommend our services</td>
<td>490</td>
<td>3.41</td>
<td>0.8</td>
<td>1.41</td>
<td>0.7</td>
<td>0.77</td>
<td>-0.22</td>
<td>1.11</td>
<td>1.06</td>
<td>0.86</td>
</tr>
<tr>
<td>Adult</td>
<td>Youth received right approach</td>
<td>383</td>
<td>3.49</td>
<td>0.63</td>
<td>1.29</td>
<td>0.69</td>
<td>0.72</td>
<td>0.91</td>
<td>0.94</td>
<td>0.84</td>
<td>1.11</td>
</tr>
<tr>
<td>Caregiver</td>
<td>Youth received needed services</td>
<td>383</td>
<td>3.51</td>
<td>0.59</td>
<td>0.69</td>
<td>0.69</td>
<td>0.71</td>
<td>0.71</td>
<td>0.91</td>
<td>0.84</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Would recommend our services</td>
<td>383</td>
<td>3.66</td>
<td>0.63</td>
<td>4.7</td>
<td>0.68</td>
<td>0.81</td>
<td>-0.59</td>
<td>1.23</td>
<td>1.12</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Seek help from us again</td>
<td>383</td>
<td>3.7</td>
<td>0.59</td>
<td>5.89</td>
<td>0.71</td>
<td>0.83</td>
<td>-1.04</td>
<td>1.06</td>
<td>0.82</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note: Items listed in descending order by item difficulty (Measure).
Standard Errors of Measurement
For the adult caregiver SSS Total Score, the standard error of measurement (SEM) is 0.20 points. With 95% confidence, we can say that the true score is between approximately ±2 SEMs, or 0.40 points on a one to four point scale. For the youth SSS Total Score, the standard error of measurement (SEM) is 0.25 points. With 95% confidence, we can say that the true score is between approximately ±2 SEMs, or 0.50 points on a one to four point scale.

Minimum Detectable Change
For youth, the minimum detectable change (MDC) threshold is 0.40 points with 75% confidence for the SSS Total Score, and gives us 75% confidence that a difference of more than 0.40 points is not due to chance. For adult caregivers, the MDC threshold is 0.32 points with 75% confidence for the total score, and gives us 75% confidence that a difference of more than 0.32 points is not due to chance. If the change is in a positive direction (i.e., increase in score value) it represents an increase in service satisfaction, while a change in the negative direction indicates a reduction in perceived satisfaction with services.

Test-Retest Reliability
Not available at this time.

Evidence of Validity

Scree Plot
The scree plot, Figure 13.1, suggests that both the youth and adult caregiver SSS are one-factor scales. Since the second eigenvalue is less than 1.0, efforts to construct a second factor would likely fail. The final factor structure was tested using confirmatory factor analysis (CFA).
A confirmatory factor analysis (CFA) was run with SAS CALIS in a mode emulating Bentler’s & Wu’s (1995) EQS. We ran a simple measurement model to see how well a one-factor model could explain the observed data. Results shown in Table 13.6 suggest that despite the good fit indices, the large residuals indicate the model is not good. This means that the common factor model is imprecise in its ability to explain the items’ correlation matrix. With a scale of only 4 items, the question of “factorial validity” is nearly moot, and these confirmatory analyses are reported only for consistency with other chapters. Standardized factor loadings ranged from 0.83 to 0.85 for the youth version, and 0.72 to 0.87 for the adult caregiver version.

**Table 13.6 Evaluation of the SSS Factor Structure**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Bentler CFI</th>
<th>Joreskog GFI</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSS-Youth One-Factor Model</td>
<td>0.93</td>
<td>0.93</td>
<td>0.05</td>
</tr>
<tr>
<td>SSS-Adult Caregiver One-Factor Model</td>
<td>0.82</td>
<td>0.84</td>
<td>0.09</td>
</tr>
</tbody>
</table>

For the CFI and GFI, values greater than 0.90 indicate good fit between a model and the data (Browne & Cudeck, 1993). For the SRMR, a value of below 0.08 shows a good fit (Hu & Bentler, 1999).

Data represented in this table reflect the four-item version of the SSS.
Scoring the SSS

Use Table 13.7 to calculate the SSS Total Score. The total score is the average item score. Enter the value for the responses in fields A-D and calculate fields E and F as instructed. There are no reverse coded items in the SSS. The self-scoring form is also available in Appendix B: Measures and Self-Scoring Forms.

Use the scoring form in the case where measures are fully completed (100% response rate). Otherwise, in cases with missing data, scoring can be done by computing the mean of completed items. Determining when too much missing data occurs for computing summary scores is at the discretion of the user. The analyses presented in this chapter required 85% of the items to have valid answers.

Table 13.7 SSS Self-Scoring Form

<table>
<thead>
<tr>
<th>Item</th>
<th>Values for Responses</th>
<th>Enter value for selected responses here and calculate scores as instructed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No, definitely not</td>
<td>No, not really</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Sum of A-D: E

E / 4: F

SSS Total Score = F

Interpretation

The SSS assesses youth and caregiver satisfaction with services. Total scores range from one to four. A higher score indicates greater perceived satisfaction. Interpretation of high scores is made difficult given the ceiling effects (lack of variance at the upper range) and blurring of the top range of medium scores and high scores. The average rating exceeds 3.5 on a 4-point scale, and a score of 4.0 is both ‘medium’ and ‘high’ (see Table 13.9 below). This problem of high scores blurring distinctions is common in satisfaction scales or any measure where many receive a near perfect score. Due to the importance of measuring service satisfaction for reporting agencies, the SSS was kept as a part of the PTPB despite the psychometric problems.
**Low, Medium, High Scores**

Based on the psychometric sample, if the total score falls below 3.25 for the adult caregiver, or 3.00 for the youth, service satisfaction is considered low. All scores equal to or above those cutoffs represent a medium to high level of service satisfaction (given the blurring of medium and high scores discussed previously). Low and medium / high scores are listed in Table 13.8.

**Table 13.8 SSS Low, Medium, and High Scores**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Low</th>
<th>Medium / High</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSS-Adult Caregiver</td>
<td>&lt; 3.25</td>
<td>3.25 - 4.00</td>
</tr>
<tr>
<td>SSS-Youth</td>
<td>&lt; 3.00</td>
<td>3.00 - 4.00</td>
</tr>
</tbody>
</table>

Note: Data represented in this table reflect the four-item version of the SSS. The medium and high categories were collapsed due to overlap between the quartiles.

**Percentile Ranks**

Percentile ranks appear in Table 13.9. For example, a youth total score of 3.00 is in the 34th percentile. This means that in the psychometric study, 34% of youth scored 3.00 or lower and 66% scored higher.

**Table 13.9 SSS Percentile Ranks for Total Scores**

<table>
<thead>
<tr>
<th>Youth</th>
<th></th>
<th>Adult Caregiver</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
<td>Percentile</td>
<td>Score</td>
<td>Percentile</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1.75</td>
<td>4</td>
<td>1.75</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2.25</td>
<td>8</td>
<td>2.25</td>
<td>3</td>
</tr>
<tr>
<td>2.5</td>
<td>10</td>
<td>2.5</td>
<td>4</td>
</tr>
<tr>
<td>2.75</td>
<td>16</td>
<td>2.75</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>3.25</td>
<td>43</td>
<td>3.25</td>
<td>27</td>
</tr>
<tr>
<td>3.5</td>
<td>57</td>
<td>3.5</td>
<td>42</td>
</tr>
<tr>
<td>3.75</td>
<td>68</td>
<td>3.75</td>
<td>55</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>4</td>
<td>100</td>
</tr>
</tbody>
</table>

Note: Data represented in this table reflect the three-item version of the SSS.
References


