



Published in final edited form as:

*Adm Policy Ment Health*. 2012 March ; 39(1-2): 71–77. doi:10.1007/s10488-012-0407-y.

## Development and Psychometric Evaluation of the Youth and Caregiver Service Satisfaction Scale

M. Michele Athay, M.S. and Leonard Bickman, Ph.D.  
Vanderbilt University

### Abstract

There is widespread need for the inclusion of service satisfaction measures in mental health services evaluation. The current paper introduces the Service Satisfaction Scale (SSS), a practical and freely available measure of global youth and adult caregiver service satisfaction. The development process as well as results from a comprehensive psychometric evaluation in a large sample of clinically referred youth ( $N = 490$ ) receiving home-based care, and their caregivers ( $N = 383$ ), are presented. Multiple models for psychometric analyses were used including classical test theory (CTT), item response theory (IRT), and confirmatory factor analysis (CFA). As expected, SSS total scores are negatively skewed but the measure displays otherwise adequate scale characteristics for both the youth and caregiver versions. Thus, the SSS is a brief and psychometrically sound instrument for measuring global satisfaction in home-based mental health service settings. It has several advantages compared to existing measures including brevity, parallel youth and caregiver forms, availability at no cost, and its development on a large sample of youth and caregivers with rigorous psychometric methodology.

### Keywords

Service Satisfaction; youth mental health; caregiver; psychometrics; SSS

---

Including measures of consumer satisfaction with mental health services has become an accepted feature of mental health services evaluation. Satisfaction is often viewed as an important indicator of service quality (Ayton, Mooney, Sillifant, Pows & Rasoon, 2007; Lambert, Salzer & Bickman, 1998) and consumer engagement with treatment (for example, see Hawkins, Baer & Kivlahan, 2008). Managed care companies are particularly interested in the demonstration of high satisfaction. In fact, service satisfaction has often been used to represent whether consumer needs were met (Bickman, 2000; Burnam, 1996). The focus on service satisfaction can also be seen in its inclusion for measurement in research, evaluation, and consumer studies.

However, despite the emphasis placed on service satisfaction, the relationship between service satisfaction and treatment outcomes (i.e. symptom change) has remained ambiguous. While some research suggests service satisfaction relates to treatment outcome (e.g. Fontana, Ford & Rosenheck, 2003), other research finds no such relationship (e.g. Garland, Aarons, Hawley, & Hough, 2003; Lunnen, Ogles & Pappas, 2008). Many have concluded that service satisfaction is at best slightly related to symptom change (e.g. Lambert et al., 1998; Luk, Staiger, Mathai, Wong, Birlson & Adler, 2001; Lunnen & Ogles, 1998). Thus, while service satisfaction should not be equated the importance of clinical outcomes, it is a

measure that is expected to be included in evaluating services. Broadly defining satisfaction as “a health care recipient’s reaction to salient aspects of the context, process and results of their...experience” (Pascoe, 1983, p. 189), service satisfaction is an indicator of the treatment process and not outcome. Service satisfaction may best be used to understand the overall experience of the interaction between providers and consumers and inform the process of making treatment more client-centered and acceptable to consumers.

Assessing and exploring service satisfaction for youth mental health treatment presents another challenge: whose service satisfaction should be measured? Should it be the youth receiving the treatment, the parent or caregiver, or both? Research shows only a small, if any, correspondence between caregiver and youth satisfaction (Copeland, Koeske & Greeno, 2004; Garland, Haine, Boxmeyer, 2007; Kaplan et al. 2001). This may be attributed to the developmental differences between adults and youth and the different aspects of care each focus on when assessing satisfaction (Aarons et al., 2010). Because they may not be related, attending to the perspectives of both the caregivers and the youth is important in obtaining a more complete view of service satisfaction.

## Measures of Service Satisfaction

Several measures have been developed to measure service satisfaction within youth mental health care (for example, see Biering, 2010 for a review). Many such measures are multidimensional and capture satisfaction with specific content. For example, The Satisfaction Scales contains items specifically targeting four different aspects of services: access and convenience, the youth’s treatment process and relationship to the therapist, parent and family services, as well as global satisfaction (Brannan, Sonnichsen, & Heflinger, 1996). Similarly, the Multidimensional Adolescent Satisfaction Scale (MASS) assesses youth satisfaction with four factors: counselor qualities, meeting needs, effectiveness, and counselor conflict (Garland, Saltzman, & Aarons, 2000). On the other hand, the Parent Satisfaction Survey (PSS) and corresponding Adolescent Satisfaction Survey (ASS) contains nine different service-specific content areas such as case management services, in-home services, intake and assessment services, etc. (Brannan & Heflinger, 1993, 1994). These measures all ask about services received in the previous six months. Multi-dimensional scales are recommended when an in-depth examination of satisfaction is desired. However, the use of brief surveys is arguably the most popular method of measuring satisfaction with the lowest cost and time burden.

Several shorter measures have been developed that assess only global satisfaction with services. These include the Satisfaction Scales of the Ohio Scales (Ogles, Melendez, Davis, & Lunnen, 2001), the Client Satisfaction Questionnaire (CSQ-8; Larson et al., 1979; Attkisson & Greenfield, 1994), and several measures developed that are based on the CSQ-8 including the Youth Satisfaction Questionnaire (YSQ; Stüntzner-Gibson, Koren, & DeChillo, 1995), and the Service Satisfaction Scale (SSS; Bickman et al., 2007). Originally developed to measure patient satisfaction among adult psychiatric patients, the CSQ-8 has demonstrated adequate psychometric properties including internal reliability (Copeland, Koeske, & Greeno, 2004). Closely tied to the CSQ-8, the brief 4-item SSS is the focus of the current paper.

## The Service Satisfaction Scale

The Service Satisfaction Scale (SSS) was originally developed for use with the Peabody Treatment Progress Battery (PTPB; Bickman et al., 2007), a freely available battery of measures used to assess treatment process and progress within youth mental health services. Prior literature review revealed that no current measures contained the qualities desired for use in the PTPB, namely a measure that: 1) is short, 2) is freely available, 3) provides a

global rating of service satisfaction, 4) was developed for use with youth (ages 11 – 18), 5) has established psychometric properties, and 6) was tested with large samples (i.e.  $N > 200$ ) of youth and caregivers. Although several measures met some of these criteria, none stood out as being a suitable choice for the current purpose by meeting all these criteria. For example, although analysis of the Ohio Scales Satisfaction Scale revealed adequate psychometric properties including test-retest reliability, the sample size included only 37 caregivers and 14 youth (Ogles et al., 2001). While the CSQ-8 is short, has established psychometric properties (Attkisson & Zwick, 1982), and is available in several languages, it was developed for use with adult psychiatric patients. Evidence is lacking for use with youth. In light of this, a new measure of service satisfaction with desired qualities listed above was developed.

The items for the SSS were adapted from the CSQ-8 and were originally used as the global satisfaction content area of the Satisfaction Scales (Brannan et al., 1996). Results of the study by Brannan et al. (1996) revealed that the 5-item global scale demonstrated adequate internal reliability for both the caregiver and youth versions over various mental health settings including outpatient, inpatient, group home, in-home, day treatment, therapeutic home, and case management ( $\alpha$  range 0.88 – 0.98). Thus, the development of the SSS began with these five items. After initial testing, one question was dropped due to poor item psychometrics and an additional open-ended question was included for any additional comments. The result was a five item SSS scale with 4 close-ended questions and 1 open-ended question. The current paper presents a comprehensive psychometric evaluation of the SSS used in a large sample of clinically referred youth (aged 11 – 18) and their caregivers. Multiple psychometric analysis models are utilized including classical test theory (CTT), item response theory (IRT), and confirmatory factor analysis (CFA), producing, arguably the most rigorous psychometric evaluations conducted of a global service satisfaction measure in this population.

## Method

### Participants

Participants were drawn from a larger study evaluating the effects of a measurement feedback system (Contextualized Feedback Systems; CFS<sup>tm</sup>) on youth outcomes. This sample represents 28 sites in 10 different states comprising part of a large national provider for home-based mental health services. Services include individual and family counseling in-home care, life skills training, substance abuse treatment, crisis intervention, intensive in-home services, and case management.

The sample for the current paper included all caregivers and youth who contributed data to the CFS evaluation study during the two and a half year data collection period. Inclusion in the sample required at least one completed SSS measure. Completion for this study required all four close-ended questions to be answered. If any item responses were missing, the total score was reported as missing. However, it is important to note that only six caregivers and eight youth had one or more missing item responses and were therefore excluded. Thus, the final sample included 490 youth and 383 caregivers. For those with more than one completed SSS measure, the first completed SSS was used. Youth included in this sample ( $N = 490$ ) were an average of 14.57 years old ( $SD = 1.84$ , range = 11 – 18), slightly more than half were male (55%) and they indicated their race as Caucasian (51%), African American (26%), more than one race (10.7%), or other races (12.3%). Caregivers in this sample ( $N = 383$ ) ranged in age from 24 to 77 (mean = 45.58,  $SD = 11.09$ ) and the majority were female (84.1%). Caregivers indicated their race as Caucasian (61.8%), African American (29.2%), or other races (9%). See Riemer, Athay, Bickman, Breda, Kelley & Vides de Andrade (2012) in this issue for more information.

## Measure

The SSS contains 4 close-ended questions and one open-ended question. The 4 close-ended questions are rated on a four point Likert-type scale. Respondents are asked how much each statement matches their opinion about services from 1 ( *No, Definitely Not* ) to 4 ( *Yes, definitely* ). The items are identical across the youth and caregiver forms with the exception of the wording related to whom the question is being asked. For example, the youth version asks “Did you get...” whereas the caregiver version asks “Did this youth get...”. The four 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 responses to these four items are averaged together to create a total score. The free response item does not produce any quantitative output.

## Procedures

Clinicians administered the measure to caregivers and youth (aged 11 – 18) at the end of each clinical session based on the measurement schedule included in CFS . Completed measures were sealed in an envelope and later entered into the computer by administrative assistants at each clinical site. Data were received de-identified after a rigorous data processing protocol (see Bickman et al., 2010). The Institutional Review Board of Vanderbilt University granted approval.

## Analyses

For the evaluation of the psychometric properties of the SSS, we used methods from classical test theory (CTT), confirmatory factor analysis (CFA), and item response theory (IRT), specifically Rasch analysis. These methods provide information concerning psychometric qualities of individual items as well as the overall scale. CTT and CFA analyses were conducted with SAS® version 9.2 software, IRT analyses utilized WINSTEPS 3.36.0 (Linacre, 2007). For more details, see Riemer et al., (2012) in this issue.

Within CTT, the characteristics of each SSS item is inspected through analysis of its distributional characteristics and relationship to the total scale score. Additionally, the total scale score is described with summary statistics and an indicator of the internal reliability (i.e., Cronbach s coefficient alpha). By observing the correlation between each item and the total scale score, items that are unrelated to the measure are identified by low correlations.

The SSS was developed as a unidimensional scale measuring a single construct. Therefore, all item responses are combined to create one total scale score representing the respondent s level of service satisfaction. The interpretations made from this total score are valid as long as the assumption that the measure is unidimensional remains true. In the current sample, CFA was used where all items load on only one latent variable to evaluate whether the data support this unidimensional model. This is inspected based on several fit indices including Bentler’s Comparative Fit Index (CFI), Joreskog Goodness of Fit Index (GFI), and Standardized Root Mean Square Residual (SRMR). Resulting fit indices are compared to commonly agreed upon standards (i.e. Bentler CFI & Jorekskog 0.90; SRMR 0.05) to determine whether the model is supported by the current data.

Although several different IRT models have been developed, a Rasch model is used in the current paper. More specifically, the rating scale model (RSM) with polytomously scored items (Andrich, 1978). Application of the RSM yields item difficulty ratings and item fit statistics (infit and outfit). Within IRT, item difficulties show where an item is most precise in estimating the level of service satisfaction (on a logit scale). Fit statistics indicate whether the items fit the proposed model. Items with fit statistics outside the range of 0.6 to 1.4

indicate that items are being responded to in unpredictable ways or an item is capturing noise (Wright & Linacre, 1994). Although the RSM is a 1-parameter logistic model, WINSTEPS 3.63.0 (Linacre, 2007) provides an estimate each item's discrimination, or its ability to differentiate persons with high and low service satisfaction. Items that are able to discriminate adequately will have values close to one.

## Results

The SSS total score is comprised of the averaged responses of items 1–4 of the SSS. A total of 490 clinically-referred youth and 383 of their adult caregivers completed the SSS. Because data were collected within 28 sites, youth and caregivers are nested within site. This introduced the possibility of clustering effects. Therefore, comparisons were conducted across sites to investigate this possibility. However, results revealed no significant differences in mean levels of service satisfaction scores for youths ( $F(27, 488) = 0.79, p = 0.80$ ) or caregivers ( $F(25, 341) = 1.10, p = 0.32$ ) based on site.

Service satisfaction of youth and caregivers demonstrated a small positive correlation ( $r = 0.31, p < .001, N = 364$ ). This is consistent with the finding from several other studies that found a weak-to-moderate correlation between these respondents (Garland et al., 2007; Lambert et al., 1998; Stüntzner-Gibson et al., 1995; Turchik, Karpenko, Ogles, Demireva & Probst, 2010). Total score and comprehensive item analysis for the youth and caregiver SSS are found in table 1. Scale scores for youths and caregivers demonstrate a satisfactory degree of internal consistency (Standardized Cronbach's alpha = .86 (youth), = 0.85 (caregiver)) and items display adequate item to total correlations (range 0.70–0.73 for youth and 0.68–0.71 for caregivers).

Individual item and scale score distributions are negatively skewed and the caregiver version displays high kurtosis. On a scale from 1 to 4, mean SSS scale scores were 3.37 (youth version) and 3.59 (caregiver version). The majority of youth and caregivers reported high satisfaction with services. This resulting ceiling effect indicates the scale has difficulty differentiating clients with differing levels of positive service satisfaction but is good at identifying those without good satisfaction.

In order to aid in interpretation, scores were classified as high, medium, and low according to the 25<sup>th</sup> and 75<sup>th</sup> percentiles in the psychometric sample. However, given the skewness of the data, the medium/high categories were combined. In this sample, youth scores less than 3.00 are considered low and caregiver scores less than 3.25 are considered low. Youth scoring 3.00 or higher and caregivers scoring 3.25 and higher are considered to have medium/high service satisfaction. Based on the internal reliability of the scale and the standard error of measurement, an index of Minimum Detectable Change (MDC) was calculated. This indicates, with 75% confidence, a change of 0.40 youth SSS points or 0.32 caregiver SSS points is not due to chance.

Confirmatory factor analysis indicated the proposed one-factor model fit the data for the youth SSS (Bentler CFI = 0.93; Joreskog GFI = 0.93; SRMR=0.05) but provided a poorer fit to the caregiver SSS (Bentler CFI = 0.82; Joreskog GFI = 0.84; SRMR=0.09). Standardized factor loadings ranged from 0.77 to 0.90 (youth) and 0.71–0.83 (Caregivers). For more information, see Bickman et al., 2010.

Application of a Rasch item response model yielded item difficulties ranging from –0.22–0.30 (youth SSS) and –1.04–0.91 (caregiver SSS) logits. These values indicate that the items of the SSS are located close together on the center of the latent continuum, suggesting the SSS is most reliable in measuring service satisfaction at a narrow range in the center of the continuum, something commonly found in IRT analysis of clinical measures (Reise &

Waller, 2009). According to infit and outfit statistics (see table 1), all items demonstrated adequate fit in each respective SSS model ranging from 0.88 to 1.11 for the youth version and ranging from 0.82 to 1.23 for the caregiver version. Additionally, according to IRT analyses, all items for each version of the SSS display adequate discrimination (i.e. close to 1.00) indicating their ability to differentiate those with high and low levels of service satisfaction.

## Discussion

As is seen with most other measures of service satisfaction, data were negatively skewed, with most respondents indicating they were highly satisfied with services (Brannan et al., 1996; Lebow, 1983; Stüntzner-Gibson et al., 1995). Therefore, as noted in the psychometric results, this indicates that the measure may be useful in identifying youths and caregivers who are dissatisfied with mental health services received but is not helpful in distinguishing among those who are positive about the services they received. But, it is possible that the phenomenon of service satisfaction truly is negatively skewed in nature and therefore the SSS results actually reflect the true nature of the construct. Therefore, it may well be the case that this instrument reflects true high satisfaction and it is not a problem with the measure per se. Alternatively, as has been pointed out by Aarons and colleagues (2010), this negative skew may also be a result of social desirability and not actual satisfaction. It should be noted, however, that the SSS is a very general indicator of satisfaction and may not reflect consumer perceptions of specific aspects of services. For example, questions focused on such characteristics as hours of operation or location may reveal a less skewed distribution of responses. In previous pilots of various items, however, we have not found an approach that produces a less skewed measure.

Despite the negative skew of the SSS items and total score, the items demonstrate adequate properties for use as a scale according to methods from CTT and IRT. This includes sufficient item-total correlations, model fit, and discrimination parameters. It is notable that no other studies were located that investigated the psychometric properties of service satisfaction item and scale properties utilizing methods from IRT. Combining these methods with CTT draws on the strengths of both approaches and leads to more comprehensive information about the items of a measure. For example, although CTT is simpler to use and is widely familiar to researchers, the resulting statistics are sample dependent and include arithmetic operations that assume variables are measured at an interval scale level. Unfortunately, this assumption of interval level scaling is not empirically proved for rating scale (i.e. Likert-type) items. IRT, on the other hand, provides detailed item-level statistics that are less sample dependent while also creating linear interval-level scales (Embretson, 1996).

Although CFA results of the youth SSS indicated the data provided a good fit to the model, the four items on the caregiver SSS had a less than desirable fit. Given that multivariate normality is an assumption of confirmatory techniques, this may be directly a result of highly skewed data. One technique for correcting for this is to remove items with significant skew (Ferguson & Cox, 1993). However, this is not a reasonable solution given the small number of items to begin with and that elimination of items with significant skew would target every item. Instead, we point to RSM results as evidence that a single, unidimensional factor is represented by the SSS items. As reported earlier, all item fit indices were within the desirable range, indicating that SSS items are measuring the same unidimensional latent trait. Additionally, we followed up with a principal components analysis of caregiver SSS and found that the majority of the variance was explained by the primary measure dimension (eigenvalue = 2.78). Thus, we feel confident that the caregiver SSS measures primarily a single factor.

The low correlation between youth and caregiver satisfaction emphasizes that youth and caregivers have different views about service satisfaction and that both voices are important in assessing the youth mental health services. These results indicate that one cannot assume the youth views on service satisfaction based on caregiver views. Further work is needed to determine what internal and external factors influence these views. Such information may be important for service providers in their ongoing efforts to provide client-oriented care, knowing that such care includes two different stakeholders: the youth and the caregiver.

There are, of course, several limitations associated with this large, naturalistic study. First, youth in the current study were aged 11–18 and received home-based mental health treatment. These findings may not generalize to younger ages or other treatment settings such as clinic-based or residential services. Second, the current sample included youth receiving home-based mental health care. Although the SSS is a global measure of satisfaction and may be applied easily to other settings, the current findings may not generalize to other treatment settings such as inpatient, residential, or clinic-based services.

In an increasing emphasis on consumer participation, the need to include measures of service satisfaction is likely to continue to increase. The current study presents practical measure of global service satisfaction for use with youth mental health services. The research team has expended considerable effort in making the SSS as short as possible without losing its positive psychometric properties. Overall, results of a comprehensive psychometric evaluation in a large sample of youth and caregivers receiving home-based mental health services indicate the SSS is a psychometrically sound instrument for use in this population. Compared to other satisfaction measures it has the advantages of being briefer and its development based on multiple methods for psychometric analysis on large samples of clinically referred youth and their respective caregivers.

## Acknowledgments

This research was supported by NIMH grants R01-MH068589 and 4264600201 awarded to Leonard Bickman.

## References

- Aarons GA, Covert J, Skriner LC, Green A, Marto D, Garland AF, Landsverk J. The eye of the beholder: Youths and parents differ on what matters in mental health services. *Administration and Policy in Mental Health*. 2010; 37:459–467. [PubMed: 20140489]
- Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978; 43:561–573.
- Attkisson, CC.; Greenfield, TK. Client Satisfaction Questionnaire-8 and Service Satisfaction Scale-30. In: Maruish, ME., editor. *The use of psychological testing for treatment planning and outcome assessment*. Hillsdale, NJ, England: Lawrence Erlbaum Associates; 1994. p. 402-420.
- Attkisson CC, Zwick R. The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning*. 1982; 5:223– 237. [PubMed: 10309980]
- Ayton A, Mooney M, Sillifant K, Powels J, Rasool H. The development of the child and adolescent versions of the Verona service satisfaction scale (CAMHSSS). *Soc Psych Epid*. 2007; 42:892–901.
- Bickman L. Are you satisfied with satisfaction? (Editorial). *Mental Health Service Research*. 2000; 2:125.
- Bickman, L.; Athay, MM.; Riemer, M.; Lambert, EW.; Kelley, SD.; Breda, C.; Tempesti, T.; Dew-Reeves, SE.; Brannan, AM.; Vides de Andrade, AR., editors. *Manual of the Peabody Treatment Progress Battery*. 2. Nashville, TN: Vanderbilt University; 2010. [Electronic version]<http://peabody.vanderbilt.edu/ptpb/>
- Bickman, L.; Reimer, M.; Lambert, EW.; Kelley, SD.; Breda, C.; Dew, S., et al. *Manual of the Peabody Treatment and Progress Battery (Electronic version)*. Nashville, TN: Vanderbilt University; 2007. <http://peabody.vanderbilt.edu/ptpb/>

- Biering P. Child and adolescent experience of and satisfaction with psychiatric care: A critical review of the research literature. *Journal of Psychiatric and Mental Health Nursing*. 2010; 17:65– 72. [PubMed: 20100307]
- Brannan, AM.; Heflinger, CA. Client satisfaction with mental health services. Paper presented at the 21st Annual Meeting of the American Public Health Association; San Francisco, CA. 1993 Oct.
- Brannan, AM.; Heflinger, CA. Parental satisfaction with their children's mental health service. Paper presented at Building on Family Strengths Conference; Portland, OR. 1994 Apr.
- Brannan AM, Sonnichsen SE, Heflinger CA. Measuring satisfaction with children s mental health services: Validity and reliability of satisfaction scales. *Evaluation and Program Planning*. 1996; 19(2):131–141.
- Burnam MA. Measuring outcomes of care for substance abuse and mental disorders. *New Directions in Mental Health Services*. 1996; 71:3–17.
- Copeland VC, Koeske G, Greeno CG. Child and mother client satisfaction questionnaire scores regarding mental health services: Race, age, and gender correlates. *Research on Social Work Practice*. 2004; 14(6):434–442.
- Embretson SE. The new rules of measurement. *Psychological Assessment*. 1996; 8:341–349.
- Fontana A, Ford JD, Rosenheck R. A multivariate model of patients satisfaction with treatment for posttraumatic stress disorder. *Journal of Traumatic Stress*. 2003; 16:93–106. [PubMed: 12602657]
- Ferguson E, Cox T. Exploratory factor analysis: A user s guide. *International Journal of Selection and Assessment*. 1993; 1:84– 94.
- Garland AF, Aarons GA, Hawley KM, Hough RH. Relationship of youth satisfaction with mental health services and changes in symptoms and functioning. *Psychiatric Services*. 2003; 54(11): 1544–1546. [PubMed: 14600318]
- Garland AF, Haine RA, Boxmeyer CL. Determinants of youth and parent satisfaction in usual care psychotherapy. *Evaluation and Program Planning*. 2007; 30:45–54. [PubMed: 17689312]
- Garland AF, Saltzman MD, Aarons GA. Adolescent satisfaction with mental health services: development of a multidimensional scale. *Evaluation and Program Planning*. 2000; 23:165– 175.
- Hawkins EJ, Baer JS, Kivlahan DR. Concurrent monitoring of psychological distress and satisfaction measures as predictors of addiction treatment retention. *Journal of Substance Abuse Treatment*. 2008; 35:207–216. [PubMed: 18082998]
- Kaplan S, Busner J, Chibnall J, et al. Consumer satisfaction at a child and adolescent state psychiatric hospital. *Psychiatric Services*. 2001; 52:202–206. [PubMed: 11157119]
- Lambert W, Salzar MS, Bickman L. Clinical outcome, consumer satisfaction, and ad hoc ratings of improvement in children s mental health. *Journal of Consulting and Clinical Psychology*. 1998; 66:270–279. [PubMed: 9583330]
- Larsen DL, Attkisson CC, Hargreaves WA, et al. Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*. 1979; 2:197– 207. [PubMed: 10245370]
- Lebow JL. Research assessing consumer satisfaction with mental health treatment. *Evaluation and Program Planning*. 1983; 6:211– 236. [PubMed: 10299619]
- Linacre, JM. WINSTEPS@2.26.0 [Computer Software]. 2007. Retrieved Jan 8, 2007, from <http://www.winsteps.com/index.htm>
- Luk ESL, Staiger P, Mathai J, Wong L, Birlerson P, Adler R. Evaluation of outcome in child and adolescent mental health services: Children with persistent conduct problems. *Clinical Child Psychology and Psychiatry*. 2001; 6(1):109–124.
- Lunnen KM, Ogles BM. A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology*. 1998; 66:400–410. [PubMed: 9583343]
- Lunnen KM, Ogles BM, Pappas LN. A multiperspective comparison of satisfaction, symptomatic change, perceived change and end-point functioning. *Professional Psychology: Research and Practice*. 2008; 39(2):145–152.
- Ogles BM, Melendez G, Davis DC, Lunnen KM. The Ohio Scales: Practical outcome assessment. *Journal of Child and Family Studies*. 2001; 10:199– 212.



- Pascoe GC. Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning*. 1983; 6:185–210. [PubMed: 10299618]
- Reise SP, Waller NG. Item response theory and clinical measurement. *Annual Review of Clinical Psychology*. 2009; 5:27–48.
- Riemer M, Athay MM, Bickman L, Breda C, Kelley SD, Vides de Andrade A. The Peabody Treatment Progress Battery: History and methods for developing a comprehensive measurement battery for youth mental health. *Administration and Policy in Mental Health and Mental Health Services Research*. 2012; 39:3–12. doi: 10.1007/s10488-012-0404-1. [PubMed: 22421933]
- Stüntzner-Gibson D, Koren PE, DeChillo N. The Youth Satisfaction Questionnaire: What kids think of services. *Families in Society*. 1995; 76:616–624.
- Turchik JA, Kerpenko V, Ogles BM, Demireva P, Probst DR. Parent and adolescent satisfaction with mental health services: Does it relate to youth diagnosis, age, gender, or treatment outcome? *Community Mental Health Journal*. 2010; 46:282–288. [PubMed: 20135350]
- Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measurement Transactions*. 1994; 8:370. Retrieved March 10, 2011 from [www.rasch.org/rmt/rmt83b.htm](http://www.rasch.org/rmt/rmt83b.htm).

Table 1

Item and Total Score Analysis of the SSS (Youth: N = 490; Caregiver: N = 383)

Version	Item	Mean	SD	Skew	Kurtosis	Corr.	CFA Loadings	Measure	Infit	Outfit	Discrim.
Youth	1	3.35	0.75	-1.19	1.42	0.70	0.77	0.12	0.97	0.95	1.03
	2	3.32	0.77	-1.10	1.01	0.72	0.79	0.30	0.91	0.88	1.08
	3	3.41	0.80	-1.38	1.41	0.70	0.77	-0.22	1.11	1.06	0.86
	4	3.41	0.80	-1.40	1.58	0.73	0.80	-0.20	1.00	0.92	1.06
	Total	3.37	0.66	-1.30	1.87	-	-	-	-	-	-
Caregiver	1	3.51	0.59	-0.93	0.69	0.69	0.71	0.71	0.91	0.84	1.12
	2	3.49	0.63	-1.10	1.29	0.69	0.72	0.91	0.94	0.84	1.11
	3	3.66	0.63	-2.07	4.70	0.68	0.81	-0.59	1.23	1.12	0.80
	4	3.70	0.59	-2.27	5.89	0.71	0.83	-1.04	1.06	0.82	1.00
	Total	3.59	0.51	-1.47	2.73	-	-	-	-	-	-

SD = Standard Deviation; Corr = Correlation with total; CFA = Confirmatory Factor Analysis; Measure = item difficulty; Discrim = Discrimination