Published in final edited form as:

Adm Policy Ment Health. 2012 March ; 39(0): 104–117. doi:10.1007/s10488-012-0414-z.

Measuring Youths' Perceptions of Counseling Impact: Description, Psychometric Evaluation, and Longitudinal Examination of the Youth Counseling Impact Scale v.2

Marcia A. Kearns, M.Ed., M.A.,
University of Missouri

M. Michele Athay, M.S., and
Vanderbilt University

Manuel Riemer, Ph.D.
Wilfrid Laurier University

Abstract

The Youth Counseling Impact Scale (YCIS) is an empirically validated treatment progress measure that assesses youths' perceptions of the short term effectiveness of therapy. Since its initial publication, the original 10-item measure has been shortened to ease measurement burden and revised to include a question about a youth's insight into his or her strengths. The current study describes the development of the revised YCIS (v.2) and evaluates its psychometric properties. Additionally, this study examines whether the YCIS (v.2) total score or subscale scores change over time and investigates whether there are gender or age differences for youths' perceptions of the impact of therapy. Results found the revised version obtained comparable information to that of the original measure, and that the revised version retained the factor structure of the original model with one primary general factor of Counseling Impact and two secondary factors (Insight and Change). Results also suggested that while the YCIS (v.2) total score and Change subscale score did not change linearly over the course of treatment, the Insight subscale score showed a small but significant linear increase over time. No significant differences in YCIS scores based on youth age or gender were found. The implication of these findings, the clinical and empirical utility of this measure, and its limitations are discussed.

Systematically measuring clients' treatment progress can aid clinicians in making real-time therapeutic decisions (Kazdin, 2008). In fact, research demonstrates that the use of systematically gathered data about clients' ongoing treatment progress can help clinicians improve treatment outcomes and prevent treatment failure (Harmon et al., 2007; Lambert et al., 2002). If a clinician discovers that a client is not making specific gains during treatment, the treatment approach can be adjusted.

Client progress in treatment is commonly assessed using outcome measures of symptoms and functioning. Several measures exist for this purpose such as the Outcome Questionnaire 45 (OQ-45) (see, Harmon, et al., 2007; Lambert, et al., 2002) or the Symptoms and Functioning Severity Scale (SFSS, Athay, Riemer, & Bickman, 2012). While these types of measures can be useful for assessing treatment outcomes, many sessions are often required before clinically significant symptom changes become apparent (see Hansen & Lambert, 2003 for a review). Therefore, these measures may miss smaller therapeutic gains that

clients perceive as important, such as gaining a new perspective or taking a first step toward behavior change. Clients' perceptions of such gains may indicate that treatment is helpful, which in turn may affect their overall treatment outcomes. Indeed, clients' early perceptions that therapy is helpful have been found to predict overall treatment outcomes (Addis & Jacobson, 1996). Similarly, therapeutic realizations, described as clients' in-session accomplishments such as clarifying problems, achieving insight, or gaining an increased capacity to cope (Kolden, 1996a; Kolden et al., 2006; Kolden & Howard, 1992; Kolden et al., 2000), are associated with clients' perceptions of having made progress within a session (Kolden, 1996a; Kolden et al., 2006; Kolden & Howard, 1992), which are in turn related to treatment outcomes (e.g. Kolden, 1996a; Kolden & Howard, 1992).

Therefore, in addition to outcome measures, a complementary way of evaluating client progress is by assessing "clients' internal reactions to sessions" (Stiles et al. 1994, p. 175) and the perceived impact of sessions on clients' thoughts, feelings, and behaviors using therapeutic impact measures (Riemer and Kearns 2010). Obtaining information about clients' perceptions of counseling sessions' more immediate effects (as opposed to long-term treatment outcomes) (Stiles, et al., 1994) can provide clinicians with information about session-specific and client-identified treatment gains (Elliott & Wexler, 1994). This information can help supplement information provided by outcome (e.g. symptoms and functioning) scales to provide a more complete picture of progress in therapy and aid in determining if/when the therapeutic plan needs to be altered.

Youths' perceptions of therapeutic impact can be assessed using the Youth Counseling Impact Scale (YCIS), a published therapeutic impact measure developed for use in individual treatment sessions with youths between the ages of 11 and 18 (Riemer & Kearns, 2010). The YCIS provides information about a youth's internal reactions following individual therapy sessions and the extent to which treatment-targeted changes occur in the weeks that follow a session. In the development and psychometric evaluation of the YCIS, it demonstrated sound psychometric properties and data confirmed the theoretical factor structure of the construct (Riemer & Kearns, 2010). Since this initial publication, however, the original 10-item measure was inspected for possible shortening in order to ease measurement burden and has been revised to include a question about a youth's insight into his or her strengths, to reflect the growing emphasis on using strengths-based approaches in youth psychotherapy (Maton, Schellenbach, Leadbeater, & Solarz, 2004). Furthermore, the cross-sectional nature of the sample used in the original study precluded an examination of YCIS scores over time. The current study, therefore, describes the development and psychometric evaluation of a shortened version (7-items; v.2) of the Youth Counseling Impact Scale and utilizes growth curve modeling to investigate the linear trajectory of youths' perceptions of counseling impact over the course of treatment.

Historical Overview

To date, the majority of therapeutic impact research is found in the adult literature, particularly in the area of session impact. These studies have primarily used one of two session-specific impact measures developed for adults: the Session Evaluation Questionnaire (SEQ, Stiles 1980) and the Session Impacts Scale (SIS, Elliott and Wexler 1994). Although these measures share some similarities, they represent different constructs (Mallinckrodt, 1994).

The SEQ rates overall session quality based on "Depth" and "Smoothness" and provides an overall rating of the client's post-session affective state (Stiles, 1980). Sessions are classified as Deep (powerful, effective) or Shallow (weak, worthless) and Smooth (relaxed, comfortable) or Rough (tense, distressing) (Stiles, 1980). The SEQ was developed for the

purpose of identifying sessions that were most impactful (what earlier theorists had referred to as “the good therapy hour”; Orlinsky & Howard, 1967) in order to identify the specific therapy exchanges associated with more beneficial individual sessions and overall positive therapy outcomes (Stiles, 1980).

The second of these two measures, the SIS, emerged from the “significant events” or “change events” research paradigm whereby investigators attempted to discover the specific experiences or in-session changes clients identified as being the most significant (Timulak, 2010). These important therapy moments were initially identified by having clients attend simulated therapy sessions and describe the in-session events they viewed as the most crucial (Elliott, 1985). Clients identified two classes of in-session impacts: helpful impacts and hindering impacts (Elliott, 1985). These two classes of impacts were used to form the basis of the SIS (Elliott & Wexler, 1994). Hindering impacts involve clients’ perceptions of having had a negative experience in a session and helpful impacts are in-session changes that clients describe as beneficial (Elliott, 1985; Elliott & Wexler, 1994). Helpful impacts can be further subdivided into task impacts (i.e. clients’ beliefs that they have made progress into their presenting problem during a session by experiencing insight into themselves or figuring out how to solve a problem) or relationship impacts (i.e. clients’ perceptions of having had a positive interpersonal interaction with the therapist during session) (Elliott & Wexler, 1994). Although SIS developers described task impacts as a single construct, other researchers have further subdivided it into impacts on understanding and impacts on problem-solving (Stiles, et al., 1994). Only a few studies have previously used the SIS in youth samples (e.g., Ji 2002; King et al. 2006) and the measure has not been validated or had norms created for use with youths. Given the overwhelming lack of available and validated therapeutic impact measures developed for use with youths, the YCIS was created (Riemer & Kearns, 2010).

Original Development of the YCIS

The YCIS was developed as part of a larger battery of treatment process and outcome measures (i.e. the Peabody Treatment Progress Battery (PTPB), Bickman et al., 2007) for a large national study examining the effects of providing clinicians with regular, systematic feedback about youths’ treatment progress. The battery includes both English and Spanish versions of all measures, which were all written at a fourth grade reading level. The goal of developing the YCIS was to produce a measure that assessed youths’ perceptions of the short-term beneficial impact of counseling—both the helpful immediate in-session changes and the positive impact of therapy sessions in the weeks following treatment. The immediate in-session changes have been found to relate to treatment outcomes in both adults (Kellett, Clarke, & Matthews, 2006) and youths (King, et al., 2006). The impact of sessions in weeks following, on the other hand, had not been addressed by other session impact measures. However, measuring this impact is important because youth treatments commonly involve teaching youths new ways to manage emotional and behavioral challenges and require practice of these new skills outside of sessions (see Garland, Hawley, Brookman-Frazee, & Hurlburt, 2008). Asking clients about the usefulness and/or results of these outside practice sessions, therefore, provides insight concerning what a youth is taking with him or her from the therapy session. Indeed, researchers have suggested that clients continue to process the events of the session during the time between sessions, not just while they are in session or immediately afterward (e.g. Mallinckrodt, 1994) and that evidence of the beneficial effect of treatment should not just be seen during sessions, but outside of therapy sessions as well (Orlinsky, Roonestad, & Willutzki, 2004). Therefore, the YCIS was designed to include questions assessing this type of impact.

The overall goal of developing the YCIS was to provide clinicians with a tool to measure the therapeutic impact of specific sessions or a sequence of sessions. With this tool, clinicians could try out specific therapeutic strategies within sessions and receive systematic feedback on whether the clients perceived those sessions as impactful/helpful. The measure was named the “Youth Counseling Impact Scale” (as opposed to a session impact scale) in order to reflect its broader scope in describing both in-session changes and between-session changes that occurred in the weeks following a session.

Items for the YCIS were initially derived from two places. First, items referring to internal changes that were made during sessions were adapted from the task impacts dimension of the SIS but were slightly reworded to be appropriate for youths. Items for the hindering dimension of the SIS were not used because the goal was to assess the positive impact of therapy (or lack thereof) and to have a short scale. Furthermore, the psychometric evidence for this subscale did not make a compelling case for inclusion (see Elliott & Wexler, 1994). The relationship task subscale of the SIS was also not used because the measurement battery already contained a measure of therapeutic alliance that captured information about interpersonal interactions with the therapist. Items that described changes that happened between sessions were produced based on interviews with clinicians and clinical supervisors who were employed by the service provider involved in the larger feedback study (Bickman et al., 2007, 2010). During these interviews, clinicians and supervisors were asked to generate a list of the most typical positive changes they would expect to occur as a consequence of a treatment session and the most important changes they would want to be alerted to in order to evaluate whether treatment was having a positive short term impact. The resulting items were then pilot tested in a sample of youths receiving mental health treatment, and the items that had the best psychometric properties and most strongly loaded onto a common factor were retained.

The original YCIS (now referred to with v.1), contained 10 items appropriately worded for youth (Riemer and Kearns 2010). The first five, modified from the task impacts dimension of the SIS, assessed “Insight,” which was defined as youths’ perceptions of experiencing in-session changes such as having increases in self-awareness, problem recognition, and understanding of their and others’ emotions (Riemer & Kearns, 2010). Insight subscale items included: “I learned something important about myself;” “I now understand better something about somebody else (like my parents, friends, or my brother or sister);” “I now understand my feelings better;” “I now understand better what my problems may be;” and “I now have a better idea about how I can deal with my problems.” Increases in insight have been found to be the most frequent type of helpful session event reported by both adult clients (Elliott, 1985; Elliott & Wexler, 1994) and adolescent clients (Dunne, Thompson, & Leitch, 2000). Its prevalence among client-nominated helpful therapy events suggests that it can be a useful way to measure treatment progress. In fact, in synthesizing the research literature, Eugster and Wampold defined patient progress as: “The patient reaching new understanding or insight, and having the experience of making progress in the session, through either cognitive operations such as gaining new knowledge, clarification of meaning, or making new connections, or emotional processes, such as experiencing new states or feelings” (1996, p. 1020).

The last five items in the YCIS (v.1) assessed “Change,” which was defined as youths’ perceptions of having made between-session changes in the 2 weeks prior to the current session such as trying out a new behavior, utilizing skills learned in counseling, or making improvements to their behavior (Riemer and Kearns 2010). The Change subscale included the following items: “I tried things my counselor suggested;” “I felt better about myself than before;” “I used things that I learned in counseling;” “I improved my behavior in my home;” and “I improved my behavior in school, at work, or other places like these outside of my

home.” Researchers have suggested that a true measure of outcome is one where clients have shown they have internalized treatment (Orlinsky, et al., 2004). The changes described in this subscale reflect concrete behaviors clinicians and supervisors reported youths would exhibit if they had internalized treatment by continuing to make progress between sessions.

The psychometric study of the YCIS (v.1) using cross-sectional data found that a hierarchical G-factor model with one primary general factor (Counseling Impact) and two secondary factors (Insight and Change), fit the data best (Riemer and Kearns 2010). It also found that ratings of counseling impact were related to, but distinct from, measures of therapeutic alliance, motivation for treatment, and satisfaction with services (Riemer and Kearns 2010). After initial development, the YCIS (v.1) was inspected for potential shortening given that it is one of ten clinically relevant measures included in the PTPB. This was desired for reduction of the measurement burden of the battery as a whole. Therefore, one aim of the current paper is to present the psychometric evaluation of the resulting, shortened measure (YCIS v.2). The second aim involves a longitudinal analysis of the YCIS (v.2). This is described below.

Therapeutic Impact Ratings Over Time

Little is known about how youths rate their perception of overall counseling impact, in-session increases in insight, or between-session changes over time. Research using the SIS in adults, however, has found that helpful impact ratings appear to show a small, linear increase over time (Elliott & Wexler, 1994; Reynolds et al., 1996). Research has also found differences in the trajectories of session impact scores between different therapy models and therapy durations that suggest that the construct may be capturing an important part of the therapeutic change process. For example, a study examining SIS ratings in 117 depressed clients, randomly assigned to receive 8 or 16 weeks of either cognitive behavioral therapy or psychodynamic-interpersonal therapy, found that clients receiving psychodynamic therapy rated earlier sessions lower on the problem solving component of the task dimension than clients receiving cognitive behavioral therapy (Reynolds, et al., 1996). They also found these ratings increased more steeply, for clients in the shorter treatments than those in the longer treatments. These results suggest that session impact ratings differ based on the theoretical models of change associated with the interventions used and that the change processes appear to be accelerated in treatments of shorter durations.

Relatedly, session impact ratings have been found to be associated with the therapeutic alliance, an important aspect of the therapeutic process (Horvath & Luborsky, 1993; Lambert & Barley, 2001; Martin, Garske, & Davis, 2000; Shirk & Karver, 2003; Shirk, Karver, & Brown, 2011). For example, a recent study examining the relationship between session impact ratings and the therapeutic relationship in adults found that clients rated “relationship building” sessions higher in task impacts than other sessions; they were also more likely to report high task impacts when they were securely attached to their therapist (Janzen, Fitzpatrick, & Drapeau, 2008). Similarly, an unpublished doctoral study examining the relationship between session impact scores on the SIS and the formation of the working alliance in youths undergoing counseling found that task impacts and working alliance ratings have a reciprocal relationship (Ji, 2002). Youths who rated a session as having high task impacts were more likely to endorse a higher working alliance in the following sessions. Conversely, a high rating of working alliance in a session predicted a higher task impacts rating in the following session. These findings suggest that examining session impact ratings over time may help alert clinicians to a lack of treatment progress that can potentially affect the therapeutic relationship. Collectively, these studies suggest that examining ratings of session impact over time may be a clinically useful way of monitoring treatment progress.

Given the potentially important longitudinal information provided by measures of counseling impact, the YCIS was designed to be used regularly throughout the course of therapy. However, it is important to establish if and how YCIS scores change over time, particularly when used with youths receiving mental health treatment. Knowing the typical trajectory of the overall scale and subscales over time can provide clinicians with a benchmark by which to evaluate the progress of their clients in relation to other youths receiving similar services. Examining the trajectory of this process may also be helpful in exploring when progress is made in treatment or which points in therapy youths perceive as most important (Laurenceau et al. 2007) by noting where impact ratings are highest. Information about the trajectory of these constructs may also suggest whether it makes more sense to combine the subscales or examine them separately when measuring treatment progress. Therefore, the second aim of this paper is to examine the YCIS (v.2) longitudinally. Based on previous research with adults (Elliott and Wexler 1994; Reynolds et al. 1996) we expect to find that ratings of counseling impact and its subscales increase linearly over time.

Predictors of Counseling Impact

Presently, little is known about differences in the counseling impacts reported by different types of youths. Identifying the client characteristics that may impact these perceptions, however, could help clinicians recognize the groups of clients who may perceive therapy as more or less helpful and allow them to adjust their treatment plan accordingly. Therefore, as the third aim of the current paper, client characteristics were inspected for inclusion in the longitudinal model. However, research thus far has not identified client characteristics that could potentially influence clients' perceptions of therapy helpfulness or counseling impact. Thus, the current approach for selecting potential moderators of counseling impact is largely exploratory. As a starting point, however, two commonly proposed moderators of treatment outcome, sex and age (see Kazdin, 2008), will be used.

Although client gender has not been found to predict session impact ratings in adults (Elliott & Wexler, 1994), it is not clear whether the same holds true for youths. Furthermore, in a review of the literature, Clarkin and Levy concluded that despite the fact that gender has largely not been found to relate to treatment outcomes, a continued search for potential differences is still worthwhile (2004), because there have been a few exceptions to these findings (e.g. Bockting et al., 2006; Handwerk et al., 2006; Ogrodniczuk, Piper, Joyce, & McCallum, 2001; Spek, Nyklícek, Cuijpers, & Pop, 2008). Additionally, given that gender differences are commonly observed for both internalizing and externalizing disorders in youths (Zahn-Waxler et al. 2008), there may also be gender differences in how clients respond to interventions.

A second potentially important variable to examine is age. Although little is known about age-related differences in counseling impact specifically, the moderating effects of age have been examined extensively in treatment outcome studies. Age has been found to be a moderator of treatment outcome across several studies (e.g., Curry et al. 2006; Durlak et al. 1991; Le Grange et al. 2012; Southam-Gerow et al. 2001; Weisz et al. 1995) although there are several important exceptions (e.g., Alfano et al. 2009; Garcia et al. 2010; Owens et al. 2003). Consequently, it may be an important variable to examine in the search for possible moderators of counseling impact.

Aims of the Current Paper

In summary, the current paper has three main purposes. First, it will describe and evaluate the psychometric properties of a shortened and revised version of the YCIS in a sample of clinically referred youths receiving home-based mental health services. Second, it will

examine how total and subscale scores on the YCIS (v.2) function over the course of treatment. Third, it will examine whether gender or age predict counseling impact ratings. Unless specifically noted, the remainder of this paper will refer to the YCIS (v.2) simply as the YCIS.

Methods

Participants

We drew the sample from a larger evaluation study that examined the effects of providing clinicians with regular feedback via a measurement feedback system (CFS™) on youths' outcomes. Participants were from a large national provider of home-based mental health services and represented 28 regional offices in 10 different states. This service provider is a highly decentralized organization without a prescribed treatment model. Consequently there is a great deal of variation in treatments provided, which may include in-home counseling for individuals or families, intensive in-home services, treatment for substance abuse, crisis intervention, life-skills training, and case management. Clinicians employed therapeutic techniques that included those from cognitive-behavioral, integrative-eclectic, behavioral, family systems, and play therapy approaches. In total, CFS™ data were collected from 809 youths, 695 caregivers, and 285 clinicians. The majority of clinicians were female (78%) and had a Master's degree (80%). All study procedures were approved by the Institutional Review Board of Vanderbilt University.

Upon implementation of CFS™ for the larger evaluation study, clinicians began using the system and entering data for all their 'current clients' (i.e. clients who were already receiving services when CFS™ was implemented) and all 'new clients' who began treatment during the two and a half year data collection period. Thus, only clients who began treatment after the implementation of CFS™ contributed data at the beginning of treatment. This difference between current and new clients distinguishes the two different, but overlapping, samples used in the current paper. More detailed information about the differences between the longitudinal and psychometric samples can be found in the introductory chapter in this issue (Riemer et al., 2012).

The psychometric sample consisted of all youths entered into CFS™ at any time during their treatment and who completed at least one YCIS measure. Completion of the YCIS measure (total or subscale scores) required at least 85% non-missing item responses. If 15% or more of item responses were missing, scores were not computed. Our total sample consisted of N = 462 youths.

The sample for the longitudinal study contained all youths included in the larger study's evaluation analysis (Bickman, Kelley, Breda, Vides de Andrade, & Riemer, 2011). This includes only youths who began treatment (i.e., new clients) after the implementation of CFS™. Similar to the psychometric inclusion criteria, inclusion for the longitudinal study required youths to have at least one completed (85% non-missing item responses) YCIS measure.

As part of the CFS™ evaluation study, youths were scheduled to complete the YCIS every two weeks or at least once per month during treatment. In order to reduce measurement burden (given that the YCIS was only one of ten measures used within the system), the measurement schedule was changed so that youths were no longer required to complete the YCIS after January 1, 2008. Thus, all youths in the longitudinal sample began treatment prior to January 1, 2008 (N=203). There were no differences between youths who began treatment prior to January 1, 2008 and those who started after according to age, race, or baseline symptom severity as reported by youths, caregivers, or clinicians when controlling

for multiple tests. Of these 203, 53 did not have a completed YCIS measure. Therefore, the longitudinal sample contains a total of 150 youths. Tests comparing youths entering treatment before January 1, 2008 with and without a completed YCIS, found no differences based on youth gender, race, or baseline symptom severity as reported by youths, caregivers, or clinicians. The only difference found was for youth age ($F(1, 201) = 5.93, p = 0.02$), where youths who had a completed YCIS were younger ($M = 14.69$ years $SD = 1.89$) than those without ($M = 15.58$ years $SD = 3.23$).

Measures

The YCIS is one of eleven measures used in CFS™. Although the YCIS (v.1) had good psychometric properties, it was inspected for potential shortening in order to ease the overall respondent burden when used frequently with other PTPB measures. Therefore, the items of the YCIS (v.1) were inspected with regard to several criteria: 1) general psychometric quality and item properties, 2) item redundancy, 3) ability of items to discriminate between persons, 4) relationship to assumed factor structure, and 5) contribution to overall reliability. Based on careful inspection of these criteria, four items were identified for potential elimination from the scale in order to shorten the YCIS, while maintaining adequate psychometric properties. The current paper presents the results of a comprehensive psychometric analysis of the resulting 6-item YCIS. In the revised scale, the following items are expected to assess “Insight:” “I now understand my feelings better,” “I now have a better idea about how I can deal with my problems,” and “I now understand better what my strengths are.” The following items are expected to assess “Change:” “I tried things my counselor suggested,” “I used things that I learned in counseling,” and “I improved my behavior in my home.” We expect that the revised measure will also retain the original factor structure.

Procedures

Youths completed different measures (e.g., symptoms and functioning, therapeutic alliance, treatment motivation) that are part of the Peabody Treatment Progress Battery (PTPB; Bickman et al. 2007) once a week at the end of a clinical session. Each measure had a slightly different schedule. The YCIS was completed every other session, which would thus be every other week for those clients who had weekly sessions and once every 4 weeks for those who received treatment every other week. They were asked to complete paper versions of the measures and place them in a sealed envelope. This strategy was used in an attempt to encourage candid responses. The completed measures were entered into CFS™ by a data administrator at the clinical site and feedback reports were automatically generated for the clinician. Data used in the current study were received after a rigorous data processing protocol that included removing all personal identifiers (see Bickman et al. 2010). The Institutional Review Board of Vanderbilt University granted approval for all data collection procedures. For youths in the psychometric sample, we utilized the data from the first time the measure was completed. The longitudinal analyses utilized all YCIS data collected from the longitudinal sample.

Statistical Analysis

Psychometric Study—Evaluation of the psychometric properties of the YCIS included methods from classical test theory (CTT), confirmatory factor analysis (CFA) and item response theory (IRT)—specifically a rating scale model (RSM). Utilizing methods from CTT and IRT allowed for the examination of both the psychometric qualities of individual items as well as the overall scale, while capitalizing on the strengths of each approach. The strength of CTT includes its ease of use and wide familiarity. However, the resulting statistics of CTT analyses are sample dependent and include arithmetic operations that

require variables measured at an interval scale level, something that has not been empirically proven for rating scale items. On the other hand, IRT can provide detailed item-level information that is less sample dependent while also creating linear interval-level scales (Embretson, 1996). This is accomplished by utilizing a model that estimates both item-level and person-level parameters that are able to be ordered along the same latent trait continuum. SAS® version 9.2 software was used to conduct the CTT and CFA analyses, and WINSTEPS 3.36.0 to conduct the IRT analyses (Linacre, 2007). More detailed information about the specific procedures can be found in the introductory chapter in this issue (Riemer, et al., 2012).

In the CTT approach, the analysis of each YCIS item's distributional characteristics and relationship to the total score are examined. Summary statistics and Cronbach's alpha, an indicator of the internal reliability, are then used to describe the total scale score. An examination of the correlation between each item and the total scale score can be then used to identify items that are unrelated to the measure, as indicated by low correlations.

The YCIS was developed as a two-factor scale measuring the construct of counseling impact with two related factors: Insight and Change. In order to confirm this theoretical conceptualization of the construct, CFA was applied to the data using a G-Factor model with two secondary factors (Insight and Change). Here we forced items to load on to their respective latent variables as inferred from theory.

IRT analyses used the RSM with polytomously scored items (Andrich, 1978). The RSM, a 1-parameter logistic model, provides item difficulty ratings, which show where an item is most precise in estimating the level of counseling impact (on a logit scale). It also provides item fit statistics (infit and outfit), which quantify how well an item fits with the proposed model. WINSTEPS 3.63.0 (Linacre, 2007) also provides an estimate of each item's discrimination, or its ability to differentiate persons with high and low counseling impact.

Longitudinal Study—Although the measurement schedule of CFS™ arranged for the YCIS to be completed every other week or once a month, the actual frequency and number of completed measures varied due to the real-world nature of the study. Not only did data collection depend on the initiation and frequency of treatment for each youth, the length of time the youths received treatment was not under the control of researchers and the total number of sessions varied widely (mean = 10.6, $SD = 9.16$, range = 1 – 48). Additionally, clinicians used their own discretion in the administration of measures. At times, clinicians reported (within CFS™) skipping the administration of measures due to specific circumstances such as having a “crisis” session. As a result, youths had varying numbers of YCIS measurements: approximately 17% of youths completed only one YCIS, 24% completed two, 17% completed three, 22% completed four or five, and 13% completed six or more. Comparison of youths with different numbers of YCIS measurements found no significant differences based on age, gender, race, or initial symptom severity as rated by youths, caregivers, or clinicians. Additionally, youths with differing numbers of YCIS measurements did not differ based on their first YCIS score. The only difference found was in the total number of sessions ($F(2, 148) = 16.56$, $p < .001$) where youths with only three or more measurements had more sessions ($M = 16.66$, $SD = 8.83$) compared with youths with two measurements ($M = 10.89$, $SD = 7.24$) who had more sessions than those with only one measurement ($M = 7.04$, $SD = 7.44$).

Longitudinal analyses employed hierarchical linear modeling (HLM) using SAS®. This technique is appropriate given the nesting of data (time points within youths) as well as the unequal number and spacing of (YCIS) observations per youths. An advantage of HLM is the ability to utilize all the YCIS data, even for individuals who only contributed a single

data point. For example, individuals with only one data point contributed to the intercept parameter but did not influence estimation of parameters of change over time.

Separate models were run for the YCIS total scale and YCIS subscales (Change and Insight). An example of the within-youths (level 1) model is:

$$YCIS_{ti} = \pi_{0i} + \pi_{1i}(\text{Time}_{ti}) + e_{ti} \quad (1)$$

Where $YCIS_{ti}$ represents the youth's perception of counseling impact i at time t , Time_{ti} represents the time in weeks the youth had been in treatment. An example of the level-2 final model used is specified as follows:

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{YouthAGE}) + \beta_{02}(\text{YouthGender}) + r_{0i} \quad (2a)$$

$$\pi_{1i} = \beta_{10} \quad (2b)$$

which captures mean initial YCIS (β_{00}) and weekly rate of change in YCIS (β_{10}). Equation 2a also captures the relationship between youths' starting YCIS and age (β_{01}) and starting YCIS and gender (β_{02}). Youth age was grand mean centered to facilitate discussion concerning the average aged youth and youths who are older or younger.

The r_{0i} is a level-2 residual, also known as random effect. It indicates the deviation from the mean of an initial YCIS score. This residual is assumed to be normally distributed with variance τ_{00} .

Results

Psychometric Study

The psychometric properties of the YCIS in this population of clinically-referred youths have been investigated and replicated by Bickman and colleagues (2007, 2010) and are printed in a web-based PTPB manual. Most recently, the second edition of the PTPB was released, including a chapter on the current, shortened version of the YCIS (Bickman, et al., 2010). A total of 462 youths completed the YCIS; however, only 445 contributed YCIS total scores. The excluded youths failed to complete 85% of the YCIS items and thus total scores were not calculated. Total score and comprehensive item analysis for YCIS are found in Table 1. Scores can range from 1 to 5. The mean total score for youths was 3.65 with mean Insight and Change subscale scores of 3.68 and 3.58 respectively. Neither the total scores and subscale scores nor the item responses demonstrated significant skewness or kurtosis in their distributions (see Table 1). The total YCIS score demonstrated a satisfactory degree of internal consistency (Cronbach's $\alpha = .90$) with item-total correlations ranging from 0.68 to 0.86. The Insight and Change subscales also demonstrated adequate internal consistency ($\alpha = 0.87, 0.82$ respectively).

In order to aid clinicians in the interpretation of scores relative to the psychometric sample, total scores can be classified as high, medium, and low according to the 25th and 75th percentiles. This is in keeping with the conventions used within the larger feedback study to anchor the score in a larger sample of clinically referred youths. In this sample, total YCIS scores less than 3.0 were considered low and scores greater than 4.5 were considered high. Scores at or between 3.0 and 4.5 were medium in reference to this sample. Additionally, Insight and Change subscale scores less than 3.0 were considered low and scores greater than 4.67 were considered high. Based on the internal reliability of the scale and the standard error of the measurement, an index of minimum detectable change (MDC) was

calculated. The MDC indicates the smallest amount of change in scores from one time point to the next that is likely to reflect true change and not measurement error (Schmitt and Di Fabio 2004). Based on this index, there is 75% confidence that a change of 0.68 in the YCIS total score is not due to chance and measurement error in this population. MDC's for the Insight and Change subscales were 0.40 and 0.48 respectively.

Results from application of the RSM to the data are also found in Table 1. Item difficulties ranged from -0.27 to 0.19 on a logit scale. These scores indicate that the YCIS was most precise in measuring youths who are at the center of the latent trait, something not uncommon to clinical measurement within the IRT framework (Reise and Waller 2009). Fit indices (infit and outfit) were within the desirable range (0.6 – 1.4) indicating all items fit adequately to the RSM model (Wright and Linacre 1994). In other words, youths responded to the items in ways that were reliable and consistent with the predicted model. The items also displayed adequate discrimination as indicated by values close to one. This means that items of the YCIS were able to discriminate between youths who rate their sessions as having low vs. high counseling impact.

The expected factor structure inferred from the theory was a hierarchical G-Factor model with the two secondary factors (Insight and Change) loading on a higher-order counseling impact factor. This structure was well-supported by the data. We fit the data using a one-factor model as well as the G-Factor model to evaluate how well each model could explain the observed data. In order for the model to be identified, we set the loadings of the lower-order factors on the general counseling impact factor to be equal for the G-Factor model. The fit indices of the G-Factor provided strong support for the respective factor structure. The G-model provided an excellent fit for these data that was more satisfactory than the fit indices for the one-factor model. Further evidence for the superiority of the G-Factor model compared to a one factor model, comes from a χ^2 difference test ($\chi^2_{diff} = 127$, $df = 2$, $p < .001$; see Table 2).

Longitudinal Study

The gender breakdown of the final sample ($N=150$) was 45% female, 55% male. These youths were an average of approximately 15 years old (mean = 14.8 , $SD = 1.85$, Range = 11 – 18) and were in the study for an average of 4 months (mean = 3.95 , $SD = 3.4$, Range = 0 – 17).

Baseline models without age and gender predictors were conducted prior to final models. Table 3 summarizes the results of fitting the data to the baseline and final model for YCIS total scores. According to AIC and BIC indices, the final model showed better fit to the data. The final model indicates that, holding all else constant, the average initial YCIS total score was 3.39 with significant variation between youths ($\tau_{00} = 0.54$, $p < .001$). The average YCIS total score did not change linearly over time ($\beta_{10} = .01$, $p = .11$). There were no significant differences in initial YCIS total scores based on youth age ($\beta_{01} = -0.05$, $p = 0.21$) or youth gender ($\beta_{02} = 0.26$, $p = 0.08$). The final model accounted for 54% of the total variation between clients, the same amount as in the baseline model.

The parameter estimates for the baseline and final models on the YCIS subscales are found in Table 4. For both subscales, the AIC and BIC indices indicated the final models had superior fit over the baseline models. Results were similar compared to the YCIS total score models with one exception: the YCIS Insight subscale, on average, increased linearly over the course of treatment ($\beta_{10} = 0.01$, $p < .05$). This means that for every week in treatment, YCIS Insight scores increased an average of 0.01 points. On average, the Change subscale did not change linearly over time ($\beta_{10} = .01$, $p = 0.23$). There were no significant differences in initial YCIS subscale scores based on youth age or youth gender; however, there was a

1% increase in between-client variance explained in the Insight subscale final model over the baseline.

Discussion

The first portion of the current paper evaluated the psychometric properties of the revised version of the Youth Counseling Impact Scale (v.2) in a sample of clinically referred youths (aged 11–18) receiving home-based services. Application of methods from CTT, IRT, and CFA yielded results suggesting that the psychometric properties of the YCIS are satisfactory for this population. Scale scores and individual items were approximately normally distributed in the intended population. When fitted using a structural equation modeling framework, the YCIS had high internal consistency, adequate item-total correlations, and confirmed of the proposed G-factor model with two secondary factors. IRT analysis (specifically the RSM) indicated that all items fit reasonably to the model and, thus, demonstrated good scale characteristics. These findings suggest the utility of using the revised version in obtaining comparable information to that of the original measure, while reducing measurement burden through a shorter instrument. Further analyses are needed to examine the predictive validity, sensitivity to change, and test-retest reliability of the YCIS in this population.

The second portion of this paper examined whether the YCIS and its subscales change over time using a longitudinal sample. For this, we employed hierarchical linear modeling (HLM) to examine YCIS scores longitudinally. We found no significant differences in initial YCIS total or subscale scores based on youth gender or age. Our results also found that ratings on the Insight and Change subscales did not follow the same trajectory over time. While the Insight subscale showed significant linear increase over time, the YCIS total score and Change subscale did not change linearly over the course of treatment after controlling for age and gender. This finding means that, on average, youths' perceptions of experiencing insight during sessions increased as therapy progressed, but their perceptions of making between-session changes did not. In interpreting these scores, clinicians providing services at the study agency could thus expect to see insight scores progressively increase over the course of treatment. Decreases in insight scores or a lack of increase over time may alert clinicians that their client does not perceive him or herself to be making progress, relative to the average youth in the sample. In contrast, no clear pattern emerged at this point for the Change subscale that can be used to alert clinicians about their client's progress.

While the linear increase in the Insight subscale is consistent with research assessing session impact in adults (Elliott & Wexler, 1994; Reynolds, et al., 1996), the reasons for lack of a linear increase for the total YCIS score and the Change subscale score, however, are not clear. It could be that the pattern of change in these constructs is not linear and was thus not adequately captured by our model. Some researchers have suggested that change in psychotherapy is not linear (e.g., Hayes et al. 2007). In fact, some changes in therapy can occur through sudden therapeutic gains (Gaynor et al., 2003; Tang & DeRubeis, 1999; Tang, DeRubeis, Beberman, & Pham, 2005; Tang, Luborsky, & Andrusyna, 2002). Unfortunately, we did not have enough youths with three or more data points to test for non-linear patterns of change.

On the other hand, the Change subscale may be tapping into a fixed construct and/or one that can only be changed by direct intervention. Youths may have begun therapy with an established level of willingness or ability to make changes suggested in therapy. Indeed, some researchers have suggested that people may have a set range within which they are able to change (Harkness & Lillienfeld, 1997). Alternatively, clinicians in the sample may have been unable to address this specific concern sufficiently to facilitate increases in youth

reported Change because of other pressing concerns or a lack of knowledge on how to do so. Further research is thus needed to explore these implications.

The fact that the Change subscale did not show linear change similar to the Insight subscale suggests that, although these subscales load onto a common factor cross-sectionally, they may function separately over time. In keeping with this idea, Prochaska et al. have cautioned that “Insight alone does not necessarily bring about behavior change” (Prochaska et al. 1992, p. 1110). If the Insight and Change subscales are indeed functioning separately over time, it may be more useful to examine them individually to assess treatment progress. Combining them to create a total score in a longitudinal study may be masking important information.

In addition, given the varying time frames asked about in the Insight and Change subscales (i.e., the Insight subscale asks about the immediate effects of therapy whereas the Change subscale asks about between-session changes), if a total YCIS score is meant to capture the overall short term effects of therapy, it may be more useful to combine Insight and Change subscale scores in a cross-lagged fashion (i.e., using subscale scores that refer to the same session) as opposed to combining the subscale scores from the same administration. We were unfortunately unable to examine this possibility using our current dataset.

Strengths and Limitations

The current study has several strengths worth mentioning. First, it was conducted with a diverse sample receiving treatment from a national treatment provider with offices across 28 sites in 10 states. Consequently, it helps provide information on youths’ ratings of counseling impact in a real-world setting. Second, as one of the few studies that have been done examining youths’ perceptions of the impact of therapy, it provides new knowledge on a largely understudied but potentially useful area of research. Third, its rigorous data analysis plan used multiple techniques to help triangulate psychometric findings.

Despite its strengths, the study also had several limitations. First, the study included the uneven timing of measures and unevenness in the administration and completion of measures (as a result of the real-world setting for data collection). This resulted in missing data and prevented a more thorough examination of temporal relationships between Insight and Change and prevented us from testing for non-linear patterns of change over time.

Second, the current study used a sample of youths receiving home-based services and the items for the “Change” subscale came from interviews with home-based service providers. Therefore, the scale and the results may not be generalizable to youths receiving standard outpatient mental health treatment. Furthermore, we did not have information on the specific treatment models used by the clinicians, which can influence therapeutic impact ratings (Bussell 2001; Elliott and Wexler 1994; King et al. 2006; Reynolds et al. 1996). Thus we were unable to investigate differences between therapy modalities. As a result, the average trajectory found in this paper may not be characteristic of individual clients whose trajectories are influenced by the type of treatment they received, or any other unmeasured variable. In addition, an implicit assumption in using the YCIS as a measure of treatment progress—particularly for the Insight subscale—is that change in treatment takes place within the individual. In such cases, where the primary therapeutic interaction is with the caregivers or in which there is minimal emphasis on client-driven change, the YCIS may not be an appropriate measure of treatment progress. The extent to which these types of treatments may have been used in this sample, however, is not known. Despite this possibility, all of the items chosen for inclusion in the measure were approved (and, in the case of the Change subscale, generated) by teams of clinicians and supervisors employed by the study’s treatment provider. These employees’ intimate involvement with the selection

and generation of items suggests that they felt the measure would fit with the services typically provided at their agency.

Third, this study assessed youths' perceptions of treatment progress. Although some researchers have argued that the client's perspective is the most important (e.g. Ogles, Lambert, & Fields, 2002) it may be limited if youths have trouble accurately reporting their perceptions of progress. Relatedly, the fact that youths' knew that their responses would later be seen by clinicians may have led to social desirability effects in their reporting of counseling impacts. Some researchers have suggested that while children may be better reporters of internal experiences, parents may be better reporters of behaviors (Thomas & Guskin, 2001). Unfortunately, a caregiver version of counseling impact did not exist at the time of the study to test this potential explanation, but one has been recently been developed (Riemer, Williams, & Kearns, unpublished manuscript) that may allow for examining this possibility in future research. Additionally, future work is needed to investigate the potential role of social desirability in youth responses.

Conclusions

In sum, the YCIS (v.2) is a psychometrically sound instrument (available in both English and Spanish) that can provide a free and potentially useful tool for clinicians to assess youths' perceptions of counseling impact. In fact, in 2006 the American Psychological Association (APA) Presidential Task Force on Evidence-Based Practice recommended the use of outcome monitoring for all practices (American Psychological Association, 2006). The YCIS can aid clinicians in following this recommendation and in providing their clients with more individualized services by alerting them to when an adjustment to the treatment approach may be necessary. Information that could allow clinicians to change the treatment approach when needed could potentially help them avoid a treatment failure or problems in the therapeutic relationship associated with a lack of treatment progress.

The YCIS may also provide researchers with a way to measure mechanisms of change. Specifically, the YCIS could be utilized to link the specific interventions used within a session to clients' ratings of counseling impact; ratings of counseling impact could in turn be linked to symptom change. In this sense, this type of measure could be used as bridge between psychotherapy process and outcome (Stiles, 1980). As the field moves toward widespread dissemination and implementation of evidence-based treatments, understanding treatment mechanisms will be important for deciding what aspects of treatment to transport, how best to match clients to specific treatment, and how to help clinicians effectively influence change in their clients. More research is needed, however, to explore trajectories of YCIS ratings across different treatment models and populations as well as the relationship of these ratings to treatment outcome and other therapy process variables.

Acknowledgments

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

References

- Addis ME, Jacobson NS. Reasons for depression and the process and outcome of cognitive-behavioral psychotherapies. *Journal of Consulting and Clinical Psychology*. 1996; 64:1417–1424. [PubMed: 8991328]
- Alfano CA, Pina AA, Villalta IK, Beidel DC, Ammerman RT, Crosby LE. Mediators and moderators of outcome in the behavioral treatment of childhood social phobia. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2009; 48:945–953. [PubMed: 19625981]

- American Psychological Association. Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist*. 2006; 61:271–285. [PubMed: 16719673]
- Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978; 43:561–573.
- Athay MM, Riemer M, Bickman L. The Symptoms and Functioning Severity Scale (SFSS): Psychometric evaluation and differences among youth, caregiver, and clinician ratings over time. *Administration and Policy in Mental Health and Mental Health Services Research*. 2012;10.1007/s10488-012-0403-2
- Bickman L, Kelley S, Breda C, Vides de Andrade AR, Riemer M. Effects of routine feedback to clinicians on youth mental health outcomes: A randomized cluster design. *Psychiatric Services*. 2011; 62:1423–1429. [PubMed: 22193788]
- Bickman, L.; Riemer, M.; Kelley, S.; Tempesti, T.; Brannan, A.; Athay, M., et al. Manual of the Peabody Treatment Progress Battery. Nashville, TN: Vanderbilt University; 2010.
- Bickman, L.; Riemer, M.; Lambert, E.; Kelley, S.; Breda, C.; Dew, S., et al. Manual of the Peabody treatment progress battery [Electronic version]. Nashville, TN: Vanderbilt University; 2007.
- Bockting CLH, Spinhoven P, Koeter MWJ, Wouters LF, Visser I, Schene AH. Differential predictors of response to preventive cognitive therapy in recurrent depression: A 2-year prospective study. *Psychotherapy and Psychosomatics*. 2006; 75:229–236. [PubMed: 16785772]
- Browne, MWC.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Newbury Park: Sage; 1993. p. 136-162.
- Bussell JR. Exploring the role of therapy process and outcome in interventions that target adolescent identity and intimacy. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. 2001; 61(10-B):5553.
- Clarkin, J.; Levy, K. The influence of client variables on psychotherapy. In: Lambert, MJ., editor. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 5. New York: Wiley; 2004. p. 194-226.
- Curry J, Rohde P, Simons A, Silva S, Vitiello B, Kratochvil C, et al. Predictors and moderators of acute outcome in the Treatment for Adolescents with Depression Study (TADS). *Journal of the American Academy of Child and Adolescent Psychiatry*. 2006; 45:1427–1439. [PubMed: 17135988]
- Dunne A, Thompson W, Leitch R. Adolescent males' experience of the counselling process. *Journal of Adolescence*. 2000; 23:79–93. [PubMed: 10700374]
- Durlak JA, Fuhrman T, Lampman C. Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin*. 1991; 110:204–214. [PubMed: 1835106]
- Elliott R. Helpful and nonhelpful events in brief counseling interviews: An empirical taxonomy. *Journal of Counseling Psychology*. 1985; 32:307–322.
- Elliott R, Wexler M. Measuring the impact of sessions in process-experiential therapy of depression: The Session Impacts Scale. *Journal of Counseling Psychology*. 1994; 41:166–174.
- Embretson SE. The new rules of measurement. *Psychological Assessment*. 1996; 8:341–349.
- Eugster SL, Wampold BE. Systematic effects of participant role on evaluation of the psychotherapy session. *Journal of Consulting and Clinical Psychology*. 1996; 64:1020–1028. [PubMed: 8916631]
- Forgatch, MS.; Patterson, GR. Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents. In: Weisz, JR.; Kazdin, AE., editors. *Evidence-based psychotherapies for children and adolescents*. 2. New York: Guilford; 2010. p. 159-178.
- Garcia AM, Sapyta JJ, Moore PS, Freeman JB, Franklin ME, March JS, et al. Predictors and moderators of treatment outcome in the Pediatric Obsessive Compulsive Treatment Study (POTS I). *Journal of the American Academy of Child & Adolescent Psychiatry*. 2010; 49:1024–1033. [PubMed: 20855047]
- Garland AF, Hawley KM, Brookman-Frazee L, Hurlburt MS. Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2008; 47:505–514. [PubMed: 18356768]
- Gaynor ST, Weersing VR, Kolko DJ, Birmaher B, Heo J, Brent DA. The prevalence and impact of large sudden improvements during adolescent therapy for depression: A comparison across

- cognitive-behavioral, family, and supportive therapy. *Journal of Consulting and Clinical Psychology*. 2003; 71:386–393. [PubMed: 12699032]
- Handwerk ML, Huefner JC, Smith GL, Clopton K, Hoff KE, Lucas CP. Gender differences in adolescents in residential treatment. *American Journal of Orthopsychiatry*. 2006; 76:312–324. [PubMed: 16981810]
- Hansen NB, Lambert MJ. An evaluation of the dose–response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research*. 2003; 5:1–12. [PubMed: 12602642]
- Harkness A, Lillienfeld SO. Individual differences science for treatment planning: Personality traits. *Psychological Assessment*. 1997; 9:349–360.
- Harmon SC, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, et al. Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research*. 2007; 17:379–392.
- Hayes AM, Laurenceau JP, Feldman G, Strauss JL, Cardaciotto LA. Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*. 2007; 27:715–723. [PubMed: 17316941]
- Henggeler SW. Multisystemic therapy: An overview of clinical procedures, outcomes, and policy implications. *Child and Adolescent Mental Health*. 1999; 4:2–10.
- Horvath AO, Luborsky L. The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology*. 1993; 61(4):561–573. [PubMed: 8370852]
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Janzen J, Fitzpatrick M, Drapeau M. Processes involved in client-nominated relationship building incidents: Client attachment, attachment to therapist, and session impact. *Psychotherapy: Theory, Research, Practice, Training*. 2008; 45:377–390.
- Ji PY. The role of attachment status in predicting longitudinal relationships between session-impact events and the working alliance within an adolescent client population. *Dissertation-Abstracts-International:-Section-B:-The-Sciences-and-Engineering*. 2002; 62(12-B)
- Kazdin AE. Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*. 2008; 63:146–159. [PubMed: 18377105]
- Kellett S, Clarke S, Matthews L. Session impact and outcome in group psychoeducative cognitive behavioural therapy. *Behavioural and Cognitive Psychotherapy*. 2006; 35:335–342.
- King R, Bambling M, Reid W, Thomas I. Telephone and online counselling for young people: A naturalistic comparison of session outcome, session impact and therapeutic alliance. *Counselling & Psychotherapy Research*. 2006; 6:175–181.
- Kolden GG. Change in early sessions of dynamic therapy: Universal processes and the generic model of psychotherapy. *Journal of Consulting and Clinical Psychology*. 1996a; 64:489–496. [PubMed: 8698941]
- Kolden GG, Chisholm-Stockard SM, Strauman TJ, Tierney SC, Mullen EA, Schneider KL. Universal session-level change processes in an early session of psychotherapy: Path models. *Journal of Consulting and Clinical Psychology*. 2006; 74:327–336. [PubMed: 16649877]
- Kolden GG, Howard KI. An empirical test of the generic model of psychotherapy. *Journal of Psychotherapy Practice & Research*. 1992; 1:225–236. [PubMed: 22700099]
- Kolden GG, Strauman TJ, Gittleman M, Halverson JL, Heerey E, Schneider KL. The Therapeutic Realizations Scale—Revised (TRS-R): Psychometric characteristics and relationship to treatment process and outcome. *Journal of Clinical Psychology*. 2000; 56(9):1207–1220. [PubMed: 10987692]
- Lambert MJ, Barley DE. Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*. 2001; 38:357–361.
- Lambert MJ, Whipple JL, Vermeersch DA, Smart DW, Hawkins EJ, Nielsen SL, et al. Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy*. 2002; 9:91–103.

- Laurenceau JP, Hayes AM, Feldman GC. Some methodological and statistical issues in the study of change processes in psychotherapy. *Clinical psychology review*. 2007; 27:682–695. [PubMed: 17328996]
- Le Grange D, Lock J, Agras WS, Moye A, Bryson SW, Jo B, et al. Moderators and mediators of remission in family-based treatment and adolescent focused therapy for Anorexia Nervosa. *Behaviour Research and Therapy*. 2012; 50:85–92. [PubMed: 22172564]
- Linacre, J. *Winsteps Rasch Measurement (Version 3.64. 2)*. Chicago, IL: MESA Press; 2007.
- Mallinckrodt B. Session impact in counseling process research: Comment on Elliott and Wexler (1994) and Stiles et al. (1994). *Journal of Counseling Psychology*. 1994; 41:186–190.
- Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 2000; 68:438–450. [PubMed: 10883561]
- Maton, KI.; Schellenbach, CJ.; Leadbeater, BJ.; Solarz, AL., editors. *Investing in children, youth, families, and communities: Strengths-based research and policy*. Washington, D.C: American Psychological Association; 2004.
- Ogles, BM.; Lambert, MJ.; Fields, SA. *Essentials of outcome assessment*. John Wiley & Sons Inc; 2002.
- Ogrodniczuk JS, Piper WE, Joyce AS, McCallum M. Effect of patient gender on outcome in two forms of short-term individual psychotherapy. *Journal of Psychotherapy Practice Research*. 2001; 10:69–78.
- Orlinsky DE, Howard KI. The good therapy hour: Experiential correlates of patients' and therapists' evaluations of therapy sessions. *Archives of General Psychiatry*. 1967; 16:621–632. [PubMed: 6023951]
- Orlinsky, DE.; Roonestad, M.; Willutzki, U. Fifty years of psychotherapy process-outcome research: Continuity and change. In: Lambert, M., editor. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 5. New York: Wiley; 2004. p. 307-389.
- Owens EB, Hinshaw SP, Kraemer HC, Arnold LE, Abikoff HB, Cantwell DP, et al. Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *Journal of Consulting and Clinical Psychology*. 2003; 71:540–552. [PubMed: 12795577]
- Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: Applications to addictive behaviors. *American psychologist*. 1992; 47:1102–1114. [PubMed: 1329589]
- Reise SP, Waller NG. Item response theory and clinical measurement. *Annual Review of Clinical Psychology*. 2009; 5:27–48.
- Reynolds S, Stiles WB, Barkham M, Shapiro DA, Hardy GE, Rees A. Acceleration of changes in session impact during contrasting time-limited psychotherapies. *Journal of Consulting and Clinical Psychology*. 1996; 64:577–586. [PubMed: 8698952]
- Riemer, M.; Athay, MM.; Bickman, L.; Breda, C.; Kelley, SD.; Vides de Andrade, A. *Administration and Policy in Mental Health and Mental Health Services Research*. 2012. The Peabody Treatment Progress Battery: History and methods for developing a comprehensive measurement battery for youth mental health.
- 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; 1007/s10488-012-0404-1
- Riemer, M.; Williams, NJ.; Kearns, MA. Caregiver reports on youth counseling impact: Description and psychometric evaluation of the Youth Counseling Impact Scale – Caregiver Version. (unpublished manuscript)
- Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *Journal of clinical epidemiology*. 2004; 57:1008–1018. [PubMed: 15528051]
- Shirk SR, Karver M. Prediction of treatment outcome from relationship variables in child and adolescent therapy: A meta-analytic review. *Journal of Consulting and Clinical Psychology*. 2003; 71:452–464. [PubMed: 12795570]

- Shirk SR, Karver MS, Brown R. The alliance in child and adolescent psychotherapy. *Psychotherapy*. 2011; 48:17–24. [PubMed: 21401270]
- Southam-Gerow MA, Kendall PC, Weersing VR. Examining outcome variability: Correlates of treatment response in a child and adolescent anxiety clinic. *Journal of Clinical Child Psychology*. 2001; 30:422–436. [PubMed: 11501258]
- Spek V, Nyklíček I, Cuijpers P, Pop V. Predictors of outcome of group and internet-based cognitive behavior therapy. *Journal of Affective Disorders*. 2008; 105:137–145. [PubMed: 17543392]
- Stiles WB. Measurement of the impact of psychotherapy sessions. *Journal of Consulting and Clinical Psychology*. 1980; 48:176–185. [PubMed: 7365056]
- Stiles WB, Reynolds S, Hardy GE, Rees A, Barkham M, Shapiro DA. Evaluation and description of psychotherapy sessions by clients using the Session Evaluation Questionnaire and the Session Impacts Scale. *Journal of Counseling Psychology*. 1994; 41:175–185.
- Tang TZ, DeRubeis RJ. Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*. 1999; 67:894–904. [PubMed: 10596511]
- Tang TZ, DeRubeis RJ, Beberman R, Pham T. Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*. 2005; 73:168–172. [PubMed: 15709844]
- Tang TZ, Luborsky L, Andrusyna T. Sudden gains in recovering from depression: Are they also found in psychotherapies other than cognitive-behavioral therapy? *Journal of Consulting and Clinical Psychology*. 2002; 70:444–447. [PubMed: 11952204]
- Thomas JM, Guskin KA. Disruptive behavior in young children: What does it mean? *Journal of the American Academy of Child & Adolescent Psychiatry*. 2001; 40:44–51. [PubMed: 11195561]
- Timulak L. Significant events in psychotherapy: An update of research findings. *Psychology and Psychotherapy: Theory, Research and Practice*. 2010; 83:421–447.
- Weisz JR, Weiss B, Han SS, Granger DA, Morton T. Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*. 1995; 117:450–468. [PubMed: 7777649]
- Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch measurement transactions*. 1994; 8:370.
- Zahn-Waxler C, Shirtcliff EA, Marceau K. Disorders of Childhood and Adolescence: Gender and Psychopathology. *Annual Review of Clinical Psychology*. 2008; 4:275–303.

Table 1

Item and Total Score Analysis of YCIS for Sample of Clinically Referred Youths (N = 445)

Item/Scale	Subscale	Mean	SD	Skew	Kurtosis	CFA	Corr	Measure	Infit	Outfit	Disc.
1	I	3.54	1.37	-0.59	-0.84	0.73	0.79	0.19	1.02	1.01	1.03
2	I	3.72	1.28	-0.82	-0.30	0.78	0.86	-0.12	0.79	0.8	1.24
3	I	3.80	1.26	-0.82	-0.31	0.77	0.85	-0.27	0.85	0.86	1.17
4	C	3.65	1.29	-0.65	-0.54	0.66	0.68	0.00	1.19	1.27	0.77
5	C	3.54	1.32	-0.46	-0.90	0.76	0.76	0.19	0.88	0.88	1.13
6	C	3.65	1.29	-0.61	-0.66	0.66	0.68	0.00	1.22	1.28	0.70
Total Scale Score		3.65	1.06	-0.73	-0.13						
Insight (I) Subscale		3.68	1.16	-0.82	-0.11						
Change (C) Subscale		3.58	1.12	-0.53	-0.54						

SD = Standard Deviation; CFA = Confirmatory Factor Analysis standardized factor loadings; Corr = Correlation with total; Measure = item difficulty; Disc = Discrimination

Table 2

Evaluation of the YCIS Factor Structure

Scale	χ^2	df	χ^2 / df	χ^2 diff	Bentler CFI	Joreskog GFI	SRMR
YCIS One-Factor Model	157.55	10	15.76		0.91	0.88	0.18
YCIS G-Factor Model	30.56	8	3.82	126.99	0.99	0.98	0.08

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 at or below 0.08 shows a good fit (Hu & Bentler, 1999).

Table 3

Parameter Estimates of Baseline and Final Two-Level Growth Curve Models of YCIS Total Score

	Baseline Model		Final Model	
	Parameter Estimate	95% CI	Parameter Estimate	95% CI
Fixed Effects				
Intercept (β_{00})	3.53 **	3.38, 3.68	3.39 **	3.18, 3.61
AGE(β_{01})			-0.05	-0.12, 0.03
GENDER(β_{02})			0.26	-0.05, 0.54
Time				
Intercept (β_{10})	0.01 *	0.00, 0.00	0.01	-0.03, 0.55
Variance Estimates				
Intercept (τ_{00})	0.54 **	0.38, 0.70	0.54 **	0.37, 0.71
Fit Statistics REML				
AIC	1376		1232	
BIC	1389		1251	
Intraclass Correlation Coefficients				
Between client	54%		54%	
Residual	46%		46%	

** indicates significance at $p < .001$;

* indicates significance at $p < .05$

Note: Time scaled in weeks and zero corresponds to intake. CI's were constructed using $1.96 \times SE$; AGE is grand-mean centered; Gender (0) = males

Table 4
Parameter Estimates for Baseline and Final Two-Level Growth Curve Models of YCIS Subscales

	YCIS Change Subscale				YCIS Insight Subscale			
	Baseline Model		Final Model		Baseline Model		Final Model	
	Parameter Estimate	95% CI	Parameter Estimate	95% CI	Parameter Estimate	95% CI	Parameter Estimate	95% CI
Fixed Effects								
Intercept (β_{00})	3.53**	3.37, 3.68	3.38**	3.16, 3.59	3.52**	3.35, 3.69	3.39**	3.17, 3.62
AGE(β_{01})			-0.04	-0.11, 0.04			-0.06	-0.14, 0.02
GENDER(β_{02})			0.25	-0.05, 0.54			0.28	-0.03, 0.58
Time								
Intercept (β_{10})	0.01	-0.01, 0.02	0.01	-0.00, 0.02	0.02*	0.00, 0.03	0.01*	0.00, 0.026
Variance Estimates								
Intercept (τ_{00})	0.54**	0.37, 0.70	0.53**	0.36, 0.70	0.57**	0.38, 0.75	0.57**	0.38, 0.76
Fit Statistics REML								
AIC	1480		1332		1572		1394	
BIC	1493		1351		1585		1413	
Intraclass Correlation Coefficients								
Between client	50%		50%		46%		47%	
Residual	50%		50%		54%		53%	

Note: Time scaled in weeks and zero corresponds to intake. CI's were constructed using 1.96*SE; AGE is grand-mean centered; Gender (0) = males

** indicates significance at $p < .001$;

* indicates significance at $p < .05$