Published in final edited form as:

Adm Policy Ment Health. 2012 March ; 39(1-2): 78–89. doi:10.1007/s10488-011-0398-0.

The relationship between change in therapeutic alliance ratings and improvement in youth symptom severity: Whose ratings matter the most?

Leonard Bickman, Ana Regina Vides de Andrade, M. Michele Athay, J.I Chen, A.S De Nadai, B Jordan-Arthur, and M.S. Karver

Abstract

This paper presents the psychometric evaluation of brief measures of therapeutic alliance (TA) for youths, clinicians and caregivers and a longitudinal analysis of relationships between changes in TA and changes in youth symptom and functioning severity. Psychometric analyses using methods from Classical Test Theory, Item Response Theory, and Factor Analysis indicate that the measures of TA used in this study offer something new for both practice and research. The measures have variability, sensitivity to change over time, brevity and can be used with multiple parties through parallel forms. The longitudinal analyses, employing hierarchical linear modeling with time-varying covariates, found the TA ratings of the clinician correlated with symptom improvement as rated by the clinician, caregiver and youth. Additional analysis showed that decreases in clinician-rated youth TA was most important in predicting a lower rate of youth improvement. Implications for future research and clinical practice are discussed.

Keywords

therapeutic alliance; youth mental health; improvement; ratings of therapeutic alliance

Longitudinal Perspectives on the Therapeutic Alliance in Youth

Therapeutic alliance (TA) is a well-established therapeutic process factor that has been shown to be associated with youth psychotherapy outcomes (Shirk & Karver, 2003; Shirk, Karver, & Brown, 2011). Therapeutic alliance is part of a larger framework of what are known as common process factors, which are trans theoretical constructs that influence treatment outcomes across a variety of psychotherapies (see Karver, Handelsman, Fields, & Bickman (2005) for a review). Numerous studies have found that TA is associated with better treatment outcomes across several youth and family mental health treatment settings (e.g. Anker, Owen, Duncan, & Sparks, 2010; Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2009; Friedlander, Lambert, & de la Pe a, 2008; Hawley & Garland, 2008; Karver, Handelsman, Fields, & Bickman, 2006; Kazdin, Marciano, & Whitley, (2005); Kazdin, Whitley, & Marciano (2006); Kazdin, & Whitley, (2006); Rector, Zuroff, & Segal, 1999; Shelef, Diamond, Diamond, & Liddle, 2005; Shirk & Karver, 2003; Shirk, et al., 2011). Therapeutic alliance has also been found to be associated with maintaining youth involvement in treatment (Chu & Kendall, 2004; Karver et al., 2008; Shirk, et al., 2011), which is particularly relevant given that approximately 25 to 60 percent of youth do not complete treatment (Armbruster & Kazdin, 1994; Fernandez & Eyberg, 2009; Gonzalez, Weersing, Warnick, Scahill, & Woolston, 2011; Hawley & Weisz, 2005; Miller, Southam-

Gerow, & Allin, 2008; Topham, Wampler, Titus, & Rolling, 2011). This high drop-out rate has broad individual and public health implications; influencing treatment outcome, the dissemination and implementation of evidence-based practices, and healthcare costs (Kazdin, 1996; National Institute of Mental Health, 2001). Developing a strong relationship with youth clients through TA may improve treatment outcomes and, consequently, may enhance the efficiency of and decrease burden on mental health treatment systems.

However, some researchers have reported null or weak findings when studying the relationship between TA and treatment outcome (DeRubeis, Brotman, & Gibbons, 2005; DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gelfand, 1999; McLeod, 2011). This discrepancy in the literature may be the result of the methodologies employed to measure TA and the way the field conceptualizes the very nature of the alliance, especially in the case of youth mental health research. Reviewing current measures of youth TA, Shirk, Caporino, & Karver (2010) found several problematic issues with current TA measurement methods. In particular, for self-report measures of TA several issues may be responsible for ceiling effects that have been found and may subsequently influence the validity of TA measures. For example, many TA measures have socially desirable, easy to endorse items. The measures also tend to be written in a general rather than a session-specific manner. After a session, clients are not asked to endorse an aspect of the alliance from the specific session, but are instead asked very general questions about the relationship with the therapist. Additionally, contemporary alliance measures frequently ask for amorphous impressions of the overall therapeutic relationship, which are more susceptible to measurement error than questions that evaluate specific attributes (e.g., alliance at a specific session; Crocker & Algina, 2005).

Furthermore, many youth TA measures have only undergone limited, if any, psychometric development studies. Many youth TA measures are solely downloads of adult measures with some slight modifications, while other measures have only been examined with small youth samples and very specific populations such as substance abusers. The lack of adequate psychometric development studies with sufficiently sized youth samples may explain why prior youth studies have only found youth alliance to have one factor (Faw, Hogue, Johnson, Diamond & Liddle, 2005; Hogue, Dauber, Stambaugh, Cecero & Liddle, 2006; Shirk, Caporino, & Karver, 2010). Underdeveloped research in the youth TA field may underlie the paucity of good therapeutic alliance items for youth and even the proper identification of youth therapeutic alliance subconstructs. This is especially salient for measures adapted from adult TA research, as the factor structure and measurement properties of adult measures and their items may not generalize, or even be applicable, to youth populations. Another shortcoming of some of these measures is that they are often lengthy, which precludes use in routine clinical practice. Given the great utility of regular monitoring of treatment progress (Bickman, 2008; Bickman et al., 2000; Lambert, Hansen, & Finch, 2001; Lambert et al., 2001), a brief but accurate measure of TA in youth psychotherapy is needed.

Considering the issues of length, social desirability-associated ceiling effects, and limited psychometrics in prior measures, a new measure addressing these critical flaws in TA measurement is necessary. Consequently, the Therapeutic Alliance Quality Scale (TAQS) and Therapeutic Alliance Quality Rating (TAQR) have been developed to address such criticisms in TA measurement. The TAQS is composed of five items and is rated by the caregiver and youth while the TAQR is a single item rated by the clinician. Development of these measures was based upon Bordin's conceptualization of TA, which focuses on the client-therapist relationship, agreement on the tasks to be performed in therapy, and agreement on treatment goals (Bordin, 1979). As TA may be expressed differently in youth populations, a large set of items was administered during initial development of these measures. These items were specifically developed with the intention of having some

difficult to endorse items to try to counteract the pervasive problem of ceiling effects in the alliance literature (Hughes & Kendall, 2007; Kendall et al., 1997; Shelef et al., 2005). The initial TAQS and TAQR item pool then underwent an extensive, iterative process, using multiple youth samples in order to remove, add, and modify items to develop a scale with optimum reliability and validity (Bickman et al., 2010). Another advantage of these new measures is their focus on session-based TA as opposed to general TA. Prior measures have evaluated TA in a more general manner without focusing on a specific session. Using a session-based measure may lead to more accurate responding and help to limit the influence of social desirability, as asking the client to judge the TA of a specific recent session may lessen the likelihood of a client focusing on judgment of the therapist or therapy or other halo effects.

While these measures represent an important potential advance in the measurement of youth therapeutic alliance, utility may be diminished if they are utilized in the manner of most therapeutic alliance measures. The majority of alliance research has focused on single time point measures of alliance and little research has examined the relationship between changes in alliance over time and treatment outcomes (Shirk et al., 2011). This methodology is problematic because the therapeutic alliance is probably not a static state for which a single sampling is an adequate measure of its association with treatment outcome. Therapeutic alliance is most likely an active, modifiable process in which the changes in TA are the truly predictive aspect of the alliance. Consistent with this, prior research has found that the dependability of TA measurement may be contingent upon the number of measurement time points, where reliability of measurement increases with the number of time points assessed (Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011). Consequently, multiple measurements of TA over time may provide for a more accurate representation of the trajectory of TA and of how TA is related to treatment outcome.

Within the adult literature, several TA trajectories have been explored as candidates for optimum treatment outcome in adults. Stiles and colleagues (2004) found that a rupture-repair or V-shaped trajectory of TA was associated with better outcomes, while Kramer, de Roten, Beretta, Michel & Despland (2009) found that a linear trajectory predicted the best outcomes (also see Kivlighan & Shaughnessy, 1995; Kivlighan & Shaughnessy, 2000). Within youth psychotherapy, the evidence base remains extremely limited regarding TA trajectories, especially in the case of youth mental health treatment, and has focused predominantly on measurement at discrete time points as opposed to longitudinal measurements of treatment outcome over time (Chiu, McLeod, Har, & Wood, 2009; Eltz, Shirk, & Sarlin, 1995; Hawley & Weisz, 2005; Hawley & Garland, 2008; Kendall et al., 2009). Overall, current findings from the literature regarding the exact relationship between TA trajectories and treatment outcome remain inconclusive.

Another limitation of the current research literature is scant attention paid to whose rating of TA is being compared with whose rating of severity. While there is debate in the adult literature concerning if the client's rating of TA is more predictive than the clinicians (De Bolle, Johnson, & De Fruyt, 2010; Horvath, 2001; Horvath, Del Re, Flückiger, & Symonds, 2011), in treatment with youth there is the added complexity of caregiver involvement in treatment (Green, 2009; Hawley & Garland, 2008). In the present study these three respondents' (clinician, youth and caregiver) ratings of severity of mental health status (symptoms and functioning) are compared with three ratings of TA provided by them.

The current study, as part of a larger community mental health treatment study of the effects of feedback to clinicians on mental health outcomes (Bickman, Kelley, Breda, De Andrade, & Riemer, 2011), examines a novel TA assessment system that evaluates TA at multiple points throughout treatment from caregiver, child, and clinician perspectives as part of the

Peabody Treatment Progress Battery (PTPB) 2nd edition (Bickman et al., 2010). The purpose of this paper is twofold: to evaluate the psychometric properties of the TAQR and the TAQS and to assess the influence of TA trajectories on treatment outcome, with both TA and symptom severity being measured at multiple sessions over time. Specifically, we hypothesized that positive trajectories of TA reports from caregivers, youths, and clinicians would each be associated with greater decrease in youth symptom severity. Examining this relationship may help to capture patterns of TA change on symptom severity change over time.

Method

Participants

Participants were drawn from a larger study evaluating the effects of a measure-driven feedback system (Contextualized Feedback Systems™) on youth mental health outcomes. These participants were in services managed by a large, national provider of home- and community-based mental health services and were from 28 regional offices in 10 states (Bickman et al., 2011). Vanderbilt University's Institutional Review Board granted study approval. To be included in the current study, participants must have at least one completed (defined as 85% non-missing items) therapeutic alliance measure and one measure of severity.

This paper uses two different but overlapping samples, one for the psychometric analyses and another for the longitudinal study. The longitudinal sample consists of those who were included in the larger evaluation study mentioned above. These were youth, and their respective caregivers and clinicians, who began treatment during the two and a half year data collection period. This resulted in a total of 288 youth, 225 caregivers, and 300 Clinicians. In addition to these participants, data from additional youth, caregivers, and clinicians were collected for the psychometric analyses. The psychometric analyses include a total of 679 youth, 561 caregivers, and 713 individual clinician ratings (one per youth). For those with multiple completed measures, the first completed measure was used for psychometric purposes. For a more detailed description of the samples used to conduct the psychometric and longitudinal analyses, please see Riemer, Athay, Bickman, Breda, Kelley & Vides de Andrade (2012) in this issue.

Measures

Therapeutic Alliance Quality Scale (TAQS)—Given the multiple participants in youth therapy the TAQS includes both youth and caregiver rated scales. Composed of five questions each, the TAQS is a session-based measure that assesses the bond with the clinician as well as agreement on goals and tasks. The TAQS has undergone extensive development and testing including expert review of items, pilot studies, and several comprehensive psychometric analyses (see Bickman et al., 2007). At each phase of development, items were removed that were unclear or redundant. The current paper describes the most recent revised TAQS measures that were shortened to reduce participant burden and eliminate item redundancy. The TAQS produces a total score.

Therapeutic Alliance Quality Rating (TAQR)—During initial development of the TAQS a parallel measure for clinicians was tested. A pilot study tested two measures of clinicians' therapeutic alliance, a single item vs. a 52- item clinician scale used by the clinician to rate youth and caregivers TA. The correlation between the long scale and a single TA item that the clinician used to rate the TA for the youth and caregiver was .75 and .81 respectively. On the TAQR, clinicians rate how they thought the TA was for that

session with the caregiver and youth. Clinicians are requested to complete a single item for the youth and a similar item for the caregiver.

Symptoms and Functioning Severity Scale (SFSS)—The SFSS is completed by the youth, caregiver and clinician and rates youth emotional and behavioral functioning. The SFSS contains 26 items (27 for clinician version) in which respondents rate the frequency of various emotional or behavioral symptoms/functioning on a five-point Likert-type scale ranging from one (*never*) to five (*very often*). The SFSS Total Score is calculated based on a linear transformation of the item average if at least 85% of the items are completed. There is a different linear transformation for each respondent. The SFSS yields a total score as well as internalizing and externalizing subscale scores. The SFSS has demonstrated sound psychometric properties for all three respondent forms (see Bickman et al., 2010). See the Athay, Riemer, and Bickman (2012) in this issue for more information on the SFSS.

Analyses

Psychometric analysis—The psychometric analysis of the TAQS evaluated each item in the scale for its reliability and validity. Multiple models were utilized in the psychometric analysis including classical test theory (CTT), confirmatory factor analysis (CFA), and Rasch modeling. All items were evaluated according to their distribution, relationship to the underlying construct of therapeutic alliance, fit with the scale as a whole and their ability to discriminate between individuals with high and low TA. Reliability was assessed with standardized Cronbach's alpha as well as item-total correlations. Assessment of the factorial validity of the TAQS was done with CFA. Additional analyses were conducted to facilitate score interpretation including total score quartiles, standard error of measurement (SEM), and a reliable change index (RCI). For more detailed description of the psychometric procedures utilized, please see Riemer et al., (2012) in the current issue. The clinician TAQR does not constitute a scale since it is only one item rating of the youth and one rating of the caregiver. Thus, only basic descriptives for each item's responses will be provided.

Results

Psychometric Study

Total scores and comprehensive item analysis for youth and caregiver TAQS are found in Table 1. With total scores able to range from 1 – 5, high TA is generally reported by the youth (Mean = 4.12, SD = 0.89) and caregiver (Mean = 4.24, SD = 0.72). All items and total scores demonstrate significant negative skewness, although kurtosis values were within acceptable range for most items. Rasch analysis results indicate that item locations (i.e. item difficulty) ranged from –0.66 to 0.47 for youth items and –0.80 to 0.88 for caregiver items. All infit and outfit values were within the acceptable range of 0.6 to 1.4 (Wright & Linacre, 1994) indicating items demonstrated good fit to the Rasch model. Additionally discrimination values were close to 1, indicating the ability of items to discriminate between those with high and low TA. The TAQS demonstrated adequate internal consistency for the youth version (standardized Cronbach's alpha = 0.85) and the caregiver version (standardized Cronbach's alpha = 0.86).

Scores were classified as high, medium, and low according to the 25th and 75th quartiles. Youth TAQS scores less than 3.8 are considered low and scores above 4.8 are considered high. Scores falling in between 3.8 and 4.8 are medium. The high Caregiver TAQS means resulted in an inability to distinguish between medium and high scores. Therefore, caregiver scores were only classified as low or medium/high. Caregiver TAQS scores less than 3.8 are considered low and scores at or above 3.8 are considered medium/high.

The SEM for youth TAQS total score is 0.34 points and for caregiver TAQS total score is 0.27. Based on these values and the reliability of the measure (Cronbach's alpha), a reliable change index was calculated. This value indicates with 75% confidence the amount a TAQS changes that are not due to chance. The RCI for the youth TAQS is 0.62 and for the caregiver TAQS is 0.49. This means that if a caregiver's TAQS score changes by 0.49 points or more, we are 75% confident that change is not due to chance.

Finally, CFA results confirmed the proposed one-factor model fit the data for the youth measure (Bentler's CFI = 0.99; Joreskog GFI = 0.97; SRMR = 0.02) and the caregiver measure (Bentler's CFI = 0.94; Joreskog GFI = 0.94; SRMR = 0.04). Standardized factor loadings ranged from 0.64 to 0.77 for the youth TAQS and 0.69 to 0.79 for the caregiver TAQS. This means a single construct is being measured by the TAQS.

The single item response for clinicians TAQR rating the session TA with caregiver and youth are normally distributed without significant skewness or kurtosis. Clinicians TA ratings with the youth were generally high ($N = 641$, Mean = 3.80, SD = 0.82, range 1 – 5, kurtosis = .51, skewness = -0.53) and TA with the caregiver even higher ($N = 505$, Mean = 4.22, SD = 0.71, range 1 – 5, kurtosis = 1.04, skewness = -0.74).

Longitudinal Study

The second purpose of this study is to assess the relationship between TA and youth's symptom and functioning severity as indicated by the three respondents. Specifically, we hypothesized that positive trajectories of TA reports from parents, youths, and clinicians would each be associated with greater decrease in youth severity. This analysis uses four pairs of TA measures: youth- rating clinician, caregivers- rating clinician, clinician- rating youth, and clinician-rating caregiver. The TA-SFSS analysis is done in three parts. First, we examine the relationship between the initial TA and youth outcomes. Second, we examine whether TA changes over time. Thirdly, we examine whether the TA trajectory (change over time) is related to severity.

Sample

There are 288 youth, 225 caregivers, and 300 clinicians reporting at least one TA and SFSS measure. On average, these youths are 14.8 years of age (SD = 1.82 years), 60% are Caucasian, 50% female, 10% are Hispanic. Initial severity on average was 51.18 (SD = 10, range 32–80), 51.43 (SD = 10, range 30–81) and 50.40 (SD = 9, range 31–76) as reported by youths, caregivers and clinicians respectively. Youths had an average of 13 sessions (SD = 9 sessions) within 18.9 weeks of participation (SD = 13 weeks). The number of youth and informants varies per measure. Table 2 shows the number of informants reporting TA and severity measures and number of youth being rated. For instance, 288 youths reported at least one self-report TA data point, each reporting an average of 4.23 TA data points. One hundred twenty one clinicians reported at least one SFSS data point on 294 youths, with an average of 5.04 SFSS data points per youth. The HLM models use all available records per measure. Ratings of TA and measures of severity were scheduled and collected every other week. However, to reduce the burden of data collection these measures were scheduled on different weeks (i.e. TA in even weeks and severity in odd weeks). Sixty-seven percent of clinicians, 37% youths, and 26 % of caregivers reported their TA and SFSS ratings within one week apart. Rest of informants reported these two measures 2 to 3 weeks apart.

Respondent Initial Therapeutic Alliance

Initial TA measures were reported at baseline or within the first two weeks of participation in the study. Table 3 shows the bivariate correlations of initial TA measures among informants; statistically significant correlations are bolded. Initial TA measures correlated

very poorly among informants, the paired correlations ranged from 0.01 to .27. Only the ratings provided by the clinicians of youth and -caregivers TA ratings correlated 0.57 ($p < 0.0001$), indicating that clinicians perceived there is a substantial relationship between how the caregiver and youth rated the TA during that session; when in fact, the actual correlation ($r = .13$), as reported by the caregivers and youths, was not statistically significant.

Does initial TA correlate with rate of improvement in severity?

We estimated four separate models to examine whether initial TA predicted youth's change in severity. To be clear we are using the term predict in a statistical sense and stress that this does not imply these data indicate a causal relationship. To account for the nesting of repeated measures within youth, and youth within clinicians, and clinicians within sites, we used hierarchical longitudinal linear models (HLM). These models offer important advantages over older models for better handling missing values and unequal time intervals between sessions and clients (Singer & Willett, 2003; Hedeker & Gibbons, 1997; Nich & Carroll, 1997). They also offer increased statistical power (by using all available data), describe the shape of change over time, and avoid the psychometric problems with pre-post change scores (Lambert et al., 2001). We used SAS (Proc Mixed v9.2) to estimate all HLM models using maximum likelihood estimation (ML). Hierarchical linear models include fixed effects and random intercepts for youth, clinician, and site levels, allowing an exchangeable correlation structure at each of the levels.

Before estimating the HLM models, we used multiple group bootstrap testing (based on 100K re-samples with replacement) to examine whether initial severity was similar across youth's demographics groups like age, gender, and race. Based on clinicians and caregivers ratings of severity, Whites were more severe at baseline ($p = .001$ and $p = .02$, respectively) than non-Whites. Each severity HLM model then includes an intercept, youth's race (White = 1, other race = 0), initial TA, slope, and the interaction between the slope and the initial TA. To facilitate the model interpretation, initial TA measures are centered at the informant grand mean. Therefore, the intercept indicates the average SFSS for non-White youths (the reference group) at the average initial TA at time zero. The other estimated parameters are deviations from this intercept. The time variable measures the SFSS slope or rate of change per week. Negative and significant SFSS slope shows the average rate of youth's severity improvement per week. A negative and significant interaction coefficient of time and centered initial TA indicates that youth with higher initial TA, higher than the grand mean, improve faster. Table 4 shows the HLM results on severity. Regardless of informant, the estimated coefficient of the slope was negative and significant indicating that youths rates of severity improved over time. When starting treatment, ratings of severity were higher for Whites according to caregivers and clinicians. All estimated interaction coefficients of initial TA and slope were not significant concluding that measures of initial TA do not predict youth's rates of improvement in symptoms and functioning.

Does TA change?

On average at the beginning of treatment, all informants reported positive and favorable TA total scores. The average youth TA, i.e. youth rating of the TA, was 4.07 (SD = 0.86, range 1–5), caregivers 4.23 (SD = 0.69, range 2–5), clinician rating of youth TA was 3.77 (SD = 0.71, range 2–5), and clinician-caregivers TA was 3.84 (SD = 0.71, range 1–5). To determine whether TA changes over time we estimated four TA HLM models, each model including an intercept, initial TA, and the slope. Initial TA was also centered at the informant grand mean. The initial TA main effect controls for possible TA rating differences informants might have when starting treatment (i.e. soft vs. tough raters). The slope or rate of change is measured in months of participation in the study, if positive and significant the

slope will be indicative of TA improvement per month. Table 5 shows the HLM TA results. Regardless of informant, TA had small positive change indicating that TA improved over time. At four months in CFS (the average length of participation), the TA slope effect sizes were 0.17 and 0.11 for youth and caregivers and 0.11 and 0.13 for clinicians TA with youth and caregivers respectively.

Do changes in TA correlate with changes in severity?

To examine whether changes in TA correlates with improvement in severity, we estimated SFSS growth models using HLM controlling for time (rate of SFSS improvement in weeks), youth's race, predicted TA change (in months), and the interaction of total TA change by time. A negative and significant interaction coefficient between the rate of improvement and TA change will indicate that youth with improving TA also have a higher rate of psychopathology improvement. We estimated four HLM models per each SFSS measure, twelve models total; the results are shown in table 6. According to these findings, increasing TA rated by the clinicians has a positive relationship with the rate of youth improvement as rated by all three respondents; the youth, caregiver and the clinician. According to all three psychopathology improved faster when clinicians reported TA improvements with their clients. At four months, the effect sizes of an improvement of one clinician TA point on psychopathology are 0.10, 0.21, and 0.15 as reported by youth, caregiver and clinicians respectively.

What is the effect of more Extreme TA change on SFSS?

From the TA HLM models, of the 288 youths with TA measure, 25 percent (N = 73 youths) had a decreasing TA of about .8 points (equivalent to 1 SD of initial TA). Another 35 percent of youth (N = 102 youths) reported an average TA increase of about 0.85 TA points. To distinguish between the overall effects of these more extreme TA changes on severity, we classified youths into three groups: youths with increasing TA, youths with decreasing TA, and youths with no TA change. We then estimated four HLM models per informant of severity. Outcome was a function of time, youth's race (White = 1, other race = 0), negative TA (equal to 1 for youth with decreasing TA and zero otherwise), positive TA (= 1 for youth with increasing overall TA and zero otherwise), and two interactions of positive TA by time, and negative TA by time. The reference groups were non-White youths with stable or no TA change. The time interactions terms let us examine whether youths with overall positive or negative TA change also have higher or smaller rates of improvement compared to youths with no TA change. Table 7 shows the HLM results. Youths with increasing and no TA change improved at the same rate. However, regardless of informant, youths whose clinicians reported a decreasing TA did not improve as fast as other two groups. Decreasing TA reported by clinicians had effect sizes of .24, .25 and .21 for youth, caregiver and clinicians SFSS respectively. Also youth reporting negative TA also reported less (self) improvement (ES = 0.37). All effect sizes are estimated at the average length of stay.

Discussion

The first purpose of this study was to evaluate the psychometric properties of three very brief measures of TA that can be used with youth, caregivers and clinicians and second to assess the relationship between change in TA and change in symptom severity across multiple sessions of therapy. The results of the Rasch analysis as well as the classic psychometric analyses showed that while the measures had excellent psychometric properties, they still had characteristically high TA ratings as has been found in previous research. It is unclear if this is a problem of the measures used or if it is, in fact, a reflection of the true ratings of TA. Alliance measures may be prone to social desirability artifacts that distort the accuracy of respondents' evaluations. Despite years of attempts to reduce this

effect from TA self-report measurement by multiple research groups (including our own), little success has been achieved, even when making the assessment specific to a single session. On the other hand the high TA scores may well reflect the real client perceptions of positive skills and attributes of clinicians. It is plausible that people with excellent social skills self-select to become clinicians, even if they do not receive training in those skills during graduate school. However, it is also possible that clients, especially youth, are not able to differentiate between levels of quality of therapeutic alliance due to lack of experience with quality relationships with family, friends, and even other clinicians.

Some evidence examining the role of social desirability is available in previous research directed by some of the authors of the present paper (Bickman, Andrade, Lambert, et al., 2004). They conducted two studies to examine the social desirability of a previous version of the TA measure used in the current research. Youth were asked to complete the TA rating form as usual. However, after they turned in their TA rating form, they were asked for their permission to share this information with the counselor. Sixty-three percent of youth participants gave their permission to release and share the TA scores with their counselors. There was no difference between the TA ratings for youth willing to share their ratings and those youth choosing not to share this information. A second, more direct study was conducted in which youth were told *prior* to completing the TA rating that their TA scores might be shown to their counselors. The main hypothesis was whether TA increased after youths knew that counselors might know their scores. A linear longitudinal regression showed that TA scores did *not* change as a result of being informed that their ratings might be shared with their counselors. These studies have some evident limitations, but the results do suggest that social desirability might not play as strong a role as many believe. Regardless of possible ceiling effects, it does appear that the measures of TA used in this study offer something new for both practice and research. The measures have variability, sensitivity to change over time, brevity and can be used with multiple parties through parallel forms.

The present study also examined TA over time. The majority of previous youth TA research has focused on single time points and little research has examined the relationship between changes in TA over time and treatment outcomes (Shirk et al., 2011). Therapeutic alliance may instead be, as the current results suggest, a process in which it is the changes in TA are the predictive aspect of the alliance and not the level of TA. We did not find that TA at the start of treatment was predictive of improvement but found that the clinicians' change in ratings of TA did predict change in change in severity as reported by all three respondents. This appears to be a robust effect that suggests clinicians need to be made aware of changes in their own ratings of TA as factor related to client improvement. However, it is improbable that without consistent and timely feedback clinicians can even be aware of the subtle linear changes of their own ratings found in this study.

One of the limitations of the current analysis is that it tested a linear model of change. Future research should not only consider more complex models but also examine, as did Stiles and colleagues (2004), whether a rupture-repair or V-shaped trajectory of TA is associated with better outcomes in youth treatment. The data also show a significant amount of within and between respondent variance over time that suggests other patterns of alliance might exist that would explain some of this variance. Using a data collection system that supports session-by-session data collection is needed to provide the wealth of fine-grained data that will allow a more in-depth and comprehensive understanding of TA.

Finally, one of the major limitations of our understanding of TA is the lack of tested interventions designed to improve TA. We need to move the alliance literature out of the correlational domain and develop interventions to directly improve therapeutic alliance. The

lack of any significant literature is extraordinary considering the thousands of studies that have been conducted and the significance TA has for clinicians and researchers alike.

References

- Anker MG, Owen J, Duncan BL, Sparks JA. The alliance in couple therapy: Partner influence, early change, and alliance patterns in a naturalistic sample. *J Consult Clin Psychol*. 2010; 78(5):635–645. [PubMed: 20873899]
- Armbruster P, Kazdin AE. Attrition in child-psychotherapy. *Advances in Clinical Child Psychology*. 1994; Vol 16(16):81–108.
- Athay MM, Riemer M, Bickman L. The Symptoms and Functioning Severity Scale (SFSS): Psychometric evaluation and differences of youth, caregiver, and clinician ratings over time. Administration and Policy in Mental Health and Mental Health Services Research. 2012
- Barber JP, Connolly MB, Crits-Christoph P, Gladis L, Siqueland L. Alliance Predicts Patients' Outcome Beyond In-Treatment Change in Symptoms. *Personality Disorders: Theory, Research, and Treatment*, S. 2009; (1):80–89.
- Bickman L. A measurement feedback system (MFS) is necessary to improve mental health outcomes. *J Am Acad Child Adolesc Psychiatry*. 2008; 47(10):1114–1119. [PubMed: 20566188]
- Bickman, L.; Athay, MM.; Riemer, M.; Lambert, EW.; Kelley, SD.; Breda, C.; Vides de Andrade, AR. Manual of the Peabody Treatment Progress Battery. 2nd ed.. 2010. Retrieved from <http://peabody.vanderbilt.edu/ptpb/>
- Bickman, L.; Riemer, 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/>
- Bickman L, Rosof-Williams J, Salzer MS, Summerfelt WT, Noser K, Wilson SJ, Karver MS. What Information Do Clinicians Value for Monitoring Adolescent Client Progress and Outcomes? *Professional Psychology: Research and Practice*. 2000; 31(1):70–74.
- Bickman L, Andrade AR, Lambert EW, Doucette A, Sapyta J, Boyd AS, Rumberger DT, Moore-Kurnot J, McDonough LC, Rauktis MB. Youth therapeutic alliance in intensive treatment settings. *Journal of Behavioral Health Services & Research*. 2004; 31(2):134–148. [PubMed: 15255222]
- 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(12):1423–1429. [PubMed: 22193788]
- Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*. 1979; 16(3):252–260.
- Chiu AW, McLeod BD, Har K, Wood JJ. Childtherapist alliance and clinical outcomes in cognitive behavioral therapy for child anxiety disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2009; 50(6):751–758.
- Chu BC, Kendall PC. Positive association of child involvement and treatment outcome within a manual-based cognitive-behavioral treatment for children with anxiety. *Journal of consulting and clinical psychology*. 2004; 72(5):821–829. [PubMed: 15482040]
- Crits-Christoph P, Gibbons MBC, Hamilton J, Ring-Kurtz S, Gallop R. The Dependability of Alliance Assessments: The Alliance-Outcome Correlation Is Larger Than You Might Think. *Journal of consulting and clinical psychology*. 2011; 79(3):267–278. [PubMed: 21639607]
- Crocker, L.; Algina, J. Introduction to classical and modern test theory. Mason, OH: Wadsworth; 2006.
- De Bolle M, Johnson JG, De Fruyt F. Patient and clinician predictors of therapeutic alliance as predictors of improvement in depression. *Psychotherapy Psychosomatics*. 2010; 79(6):378–385.
- DeRubeis RJ, Brotman MA, Gibbons CJ. A Conceptual and Methodological Analysis of the Nonspecifics Argument. *Clinical Psychology: Science and Practice*. 2005; 12(2):174–183.
- DeRubeis RJ, Feeley M. Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research*. 1990; 14(5):469–482.
- Eltz MJ, Shirk SR, Sarlin N. Alliance formation and treatment outcome among maltreated adolescents. *Child Abuse & Neglect*. 1995; 19(4):419–431. [PubMed: 7606521]

- Faw L, Hogue A, Johnson S, Diamond GM, Liddle HA. The Adolescent Therapeutic Alliance Scale (ATAS): Initial psychometrics and prediction of outcome in family-based substance abuse prevention counseling. *Psychotherapy Research*. 2005; 15(1–2):141–154.
- Feeley M, DeRubeis RJ, Gelfand LA. The Temporal Relation of Adherence and Alliance to Symptom Change in Cognitive Therapy for Depression. *Journal of consulting and clinical psychology*. 1999; 67(4):578–582. [PubMed: 10450629]
- Fernandez M, Eyberg S. Predicting treatment and follow-up attrition in Parent–Child Interaction Therapy. *Journal of Abnormal Child Psychology*. 2009; 37(3):431–441. [PubMed: 19096926]
- Friedlander ML, Lambert JE, de la Pe a CM. A step toward disentangling the alliance/improvement cycle in family therapy. *Journal of Counseling Psychology*. 2008; 55(1):118–124.
- Gonzalez A, Weersing V, Warnick E, Scahill L, Woolston J. Predictors of treatment attrition among an outpatient clinic sample of youths with clinically significant anxiety. *Administration and Policy in Mental Health and Mental Health Services Research*. 2011; 38(5):356–367. [PubMed: 20976618]
- Green J. The therapeutic alliance. *Child: Care, Health, and Development*. 2009; 35(3):298–301.
- Hawley KM, Garland AF. Working alliance in adolescent outpatient therapy: Youth, parent and therapist reports and associations with therapy outcomes. *Child and Youth Care Forum*. 2008; 37:59–74.
- Hawley KM, Weisz JR. Youth versus parent working alliance in usual clinical care: distinctive associations with retention, satisfaction, and treatment outcome. *J Clin Child Adolesc Psychol*. 2005; 34(1):117–128. [PubMed: 15677286]
- Hedeker D, Gibbons RD. Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*. 1997; 2(1):64–78.
- Hogue A, Dauber S, Stambaugh LF, Cecero JJ, Liddle HA. Early therapeutic alliance and treatment outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology*. 2006; 74(1):121–129. [PubMed: 16551149]
- Horvath AO. The alliance. *Psychotherapy: Theory, Research & Practice*. 2001; 38(4):365–372.
- Horvath AO, Del Re AC, Flückiger C, Symonds D. Alliance in individual psychotherapy. *Psychotherapy: Theory/Research/Practice/Training*. 2011; 48(1):9–16. [PubMed: 21401269]
- Hughes AA, Kendall PC. Prediction of cognitive behavior treatment outcome for children with anxiety disorders: Therapeutic relationship and homework compliance. *Behavioural and Cognitive Psychotherapy*. 2007; 35(4):487–494.
- Karver MS, Handelsman JB, Fields S, Bickman L. Meta-analysis of therapeutic relationship variables in youth and family therapy: the evidence for different relationship variables in the child and adolescent treatment outcome literature. *Clin Psychol Rev*. 2006; 26(1):50–65. [PubMed: 16271815]
- Karver MS, Shirk S, Handelsman JB, Fields S, Crisp H, Gudmundsen G, McMakin D. Relationship Processes in Youth Psychotherapy. *Journal of Emotional and Behavioral Disorders*. 2008; 16(1):15–28.
- Kazdin AE. Dropping Out of Child Psychotherapy: Issues for Research and Implications for Practice. *Clinical Child Psychology and Psychiatry*. 1996; 1(1):133–156.
- Kazdin AE, Marciano PL, Whitley M. The therapeutic alliance in cognitive-behavioral treatment of children referred for oppositional, aggressive, and antisocial behavior. *Journal of Consulting and Clinical Psychology*. 2005; 73:726–730. [PubMed: 16173860]
- Kazdin AE, Whitley M, Marciano PL. Child-therapist and parent-therapist alliance and therapeutic change in the treatment of children referred for oppositional, aggressive, and antisocial behavior. *Journal of Child Psychology and Psychiatry*. 2006; 47:436–445. [PubMed: 16671927]
- Kazdin AE, Whitley MK. Pretreatment social relations, therapeutic alliance, and improvements in parenting practices in parent management training. *Journal of Consulting and Clinical Psychology*. 2006; 74:346–355. [PubMed: 16649879]
- Kendall PC, Comer JS, Marker CD, Creed TA, Puliafico AC, Hughes AA, Hudson J. In-Session Exposure Tasks and Therapeutic Alliance Across the Treatment of Childhood Anxiety Disorders. *Journal of consulting and clinical psychology*. 2009; 77(3):517–525. [PubMed: 19485592]

- Kendall PC, Flannery-Schroeder E, Panichelli-Mindel SM, Southam-Gerow M, Henin A, Warman M. Therapy for Youths With Anxiety Disorders: A Second Randomized Clinical Trial. *Journal of consulting and clinical psychology*. 1997; 65(3):366–380. [PubMed: 9170760]
- Kivlighan DM, Shaughnessy P. Analysis of the Development of the Working Alliance Using Hierarchical Linear Modeling. *Journal of Counseling Psychology*. 1995; 42(3):338–349.
- Kivlighan DM, Shaughnessy P. Patterns of working alliance development: A typology of client's working alliance ratings. *Journal of Counseling Psychology*. 2000; 47(3):362–371.
- Kramer U, de Roten Y, Beretta V, Michel L, Despland JN. Alliance patterns over the course of short-term dynamic psychotherapy: The shape of productive relationships. *Psychotherapy Research*. 2009; 19(6):699–706. [PubMed: 19606388]
- Lambert MJ, Hansen NB, Finch AE. Patient-Focused Research: Using Patient Outcome Data to Enhance Treatment Effects. *Journal of Consulting and Clinical Psychology*. 2001; 69(2):159–172. [PubMed: 11393594]
- Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research*. 2001; 1(1):49–68.
- McLeod BD. Relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical psychology review*. 2011; 31(4):603–616. [PubMed: 21482319]
- Miller L, Southam-Gerow M, Allin R. Who stays in treatment? Child and family predictors of youth client retention in a public mental health agency. *Child and Youth Care Forum*. 2008; 37(4):153–170. [PubMed: 19774098]
- National Institute of Mental Health. Blueprint for change: Research on Child and Adolescent Mental Health's Council's Workgroup on Child and Adolescent Mental Health Intervention Development and Deployment. Washington, DC: US Government Printing Office; 2001. (DHHS Publication No. CG 031-591)
- Nich C, Carroll K. Now you see, it now you don't: A comparison of traditional versus random-effects regression models in the analysis of longitudinal follow-up data from a clinical trial. *Journal of Consulting and Clinical Psychology*. 1997; 65(2):252–261. [PubMed: 9086688]
- Rector NA, Zuroff DC, Segal ZV. Cognitive change and the therapeutic alliance: the role of technical and nontechnical factors in cognitive therapy. *Psychotherapy*. 1999; 36(4):320–328.
- 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
- Shelef K, Diamond GM, Diamond GS, Liddle HA. Adolescent and Parent Alliance and Treatment Outcome in Multidimensional Family Therapy. *Journal of consulting and clinical psychology*. 2005; 73(4):689–698. [PubMed: 16173856]
- Shirk, SR.; Caporino, N.; Karver, M. The alliance in adolescent therapy: conceptual, operational, and predictive issues. In: Castro-Blanco, D.; Karver, M., editors. *Elusive alliance: treatment engagement strategies with high-risk adolescents*. Washington, DC: American Psychological Association Press; 2010.
- Shirk SR, Karver M. Prediction of treatment outcome from relationship variables in child and adolescent therapy: a meta-analytic review. *J Consult Clin Psychol*. 2003; 71(3):452–464. [PubMed: 12795570]
- Shirk SR, Karver MS, Brown R. The alliance in child and adolescent psychotherapy. *Psychotherapy: Theory/Research/Practice/Training*. 2011; 48(1):17–24. [PubMed: 21401270]
- Singer, JD.; Willett, JB. *Applied Longitudinal Data Analysis*. New York: Oxford University Press.; 2003.
- Stiles WB, Glick MJ, Osatuke K, Hardy GE, Shapiro DA, Agnew-Davies R, Barkham M. Patterns of alliance development and the rupture-repair hypothesis: Are productive relationships U-shaped or V-shaped? *Journal of Counseling Psychology*. 2004; 1(1):81–92.
- Topham GL, Wampler KS, Titus G, Rolling E. Predicting parent and child outcomes of a filial therapy program. *International Journal of Play Therapy*. 2011; 20(2):79–93.

Table 1

Total Score and Item Analysis for Youth and Caregivers rated TAQS

Item*	Mean	SD	Kurtosis	Skew	Corr	Std CFA loadings	Measure	Infit	Outfit	Discr
Youth Rated TAQS (N = 679)										
1	3.91	1.17	0.21	-0.95	0.66	0.73	0.47	0.95	0.96	1.04
2	4.40	0.94	3.20	-1.83	0.59	0.64	-0.66	1.19	1.13	0.86
3	4.15	1.14	0.85	-1.30	0.70	0.77	-0.03	0.96	0.95	1.06
4	4.22	1.10	1.32	-1.44	0.68	0.75	-0.20	1.02	0.97	1.05
5	3.94	1.23	0.13	-1.04	0.69	0.77	0.42	0.94	0.94	1.06
Total	4.12	0.89	1.56	-1.32						
Caregiver rated TAQS (N = 561)										
1	4.13	0.98	1.38	-1.23	0.64	0.71	0.44	1.14	1.14	0.92
2	4.27	0.89	1.68	-1.27	0.66	0.74	-0.03	1.09	1.07	0.96
3	4.39	0.81	2.07	-1.37	0.70	0.77	-0.49	0.92	0.95	1.07
4	3.98	1.01	0.57	-0.92	0.73	0.79	0.88	0.80	0.81	1.19
5	4.46	0.81	3.97	-1.83	0.63	0.69	-0.80	1.21	1.10	0.89
Total	4.24	0.72	1.76	-1.10						

SD = Standard Deviation; Measure = Item Difficulty; Corr = Correlation with Total; Discr = Discrimination

* All items range 1–5.

Table 2

Measures of Therapeutic Alliance by Informant

Measure/Informant	# informants/ # Youths	Total Records	Measures per Youth	
			Mean	Range
Therapeutic Alliance (TA)				
Youth rating – Clinician	288/288	1217	4.23	1 – 15
Caregiver rating – Clinician	280/225	753	3.45	1 – 13
Clinician rating – Youth	130/300	1447	4.82	1 – 15
Clinician rating - Caregiver	125/274	1171	4.27	1 – 15
Symptoms and Functioning Severity Scale (SFSS)				
Youth self-report	340/340	1497	4.40	1 – 16
Caregivers report of youth	381/307	1051	3.42	1 – 12
Clinicians report of youth	121/294	1483	5.04	1 – 18

Table 3

Spearman Correlations of First TA Measures by Informant

	Youth - Clinician TA	Caregiver - Clinician TA	Clinician - Youth TA	Clinician- Caregiver TA
Youth-Clinician TA	1			
Caregiver-Clinician TA	0.13	1		
Clinician-Youth TA	0.27 ***	0.09	1	
Clinician-Caregiver TA	0.04	0.14 **	0.57 ***	1

 $p < 0.001$;**
 $p < 0.01$;*
 $p < 0.05$

Table 4

HLM results of initial TA on Youth Severity by Informant

Severity per Informant	Youth reported TA ^a		Caregiver reported TA ^a		Clinician Youth reported TA ^b		Clinician-Caregiver Reported TA ^b	
	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE
Youth SFSS Intercept	49.48***	0.85	49.73***	0.94	50.16***	0.84	50.10***	0.86
Youth Race ^c	0.12	1.10	0.92	1.21	-10.13	1.08	0.02	1.13
Initial TA ^d	-0.38	0.65	1.36	0.86	0.97	0.79	0.51	0.81
Slope (weeks)	-0.21***	0.02	-0.22***	0.02	-0.21***	0.02	-0.21***	0.02
Slope*(Initial TA) ^d	0.02	0.02	-0.05	0.03	0.02	0.03	0.00	0.03
Caregiver SFSS Intercept	49.17***	0.89	50.04***	0.95	49.30***	0.87	49.16***	0.89
Youth Race ^c	3.41***	1.12	2.57*	1.21	3.33***	1.10	3.23***	1.13
Initial TA ^d	-0.26	0.64	0.81	0.88	0.31	0.79	1.18	0.84
Slope (weeks)	-0.11***	0.02	-0.12***	0.02	-0.13***	0.02	-0.12***	0.02
Slope*(Initial TA) ^d	-0.03	0.03	-0.04	0.03	0.06	0.03	0.00	0.03
Clinician SFSS Intercept	49.26***	0.87	50.19***	0.99	49.40***	0.86	49.50***	0.90
Youth Race ^c	2.85***	0.99	2.20	1.16	2.77***	0.99	2.51**	1.06
Initial TA ^d	-1.29*	0.56	-0.22	0.75	-0.73	0.71	0.83	0.73
Slope (weeks)	-0.15***	0.02	-0.15***	0.02	-0.14***	0.02	-0.15***	0.02
Slope*(Initial TA) ^d	0.02	0.02	-0.04	0.03	-0.02	0.03	-0.04	0.03

Note:

^aTherapeutic Alliance quality Scale (TAQS) is a 5-point Likert scale (1 = not at all to 5 = totally), where higher scores indicate better alliance.

^bClinicians therapeutic Alliance Quality Rating is a 5-point Likert scale (1 = very poor to 5 = excellent), where higher scores indicate higher alliance.

^cCaucasian = 1, otherwise = 0.

^dInitial TA centered at the grand mean.

500 > $\bar{d}$ *
; 100 > $\bar{d}$ **
; 1000 > $\bar{d}$ ***

Table 5

Therapeutic Alliance HLM Results

Parameters	Youth-Clinician TA ^a		Caregiver-Clinician TA ^a		Clinician-Youth TA ^b		Clinician-Caregiver TA ^b	
	Est. Coef	SE	Est. Coef	SE	Est. Coef	SE	Est. Coef	SE
Fixed Effects								
Intercept	4.10 ***	0.03	4.25 ***	0.03	3.89 ***	0.03	3.90 ***	0.03
Initial TA ^c	0.63 ***	0.03	0.73 ***	0.04	0.53 ***	0.04	0.48 ***	0.03
Slope (months)	0.03 ***	0.01	0.02 **	0.01	0.02 *	0.01	0.02 ***	0.01
Variance Component								
Between-youths	0.08 ***		0.07 ***		0.05 ***		0.02 ***	
Between-clinicians	0.00		0.01		0.04 ***		0.04 ***	
Between-sites	0.00		0.00		0.00		0.00	
Within-youth	0.39 ***		0.18 ***		0.34 ***		0.32 ***	
Fit Statistics REML								
-2 Res Log Likelihood	2494		1040		2776		2110	
Intra-Class Correlation Coefficients								
Between youths	17%		27%		12%		5%	
Between clinicians	1%		2%		9%		11%	
Between sites	0%		0%		0%		1%	
Residual (within youths)	82%		71%		79%		84%	

Note:

^aTherapeutic Alliance quality Scale (TAQS) is a 5-point Likert scale (1 = not at all to 5 = totally), where higher scores indicate better alliance.^bClinicians therapeutic Alliance Quality Rating is a 5-point Likert scale (1 = very poor to 5 = excellent), where higher scores indicate higher alliance.^cInitial TA centered at the grand mean.***
 $p < 0.001$;**
 $p < 0.01$;*
 $p < 0.05$

Table 6

HLM Results of TA Change on Youth Severity by Informant

SFSS Measures	Youth TA ^a		Caregiver TA ^a		Clinician Youth TA ^b		Clinician Caregiver TA ^b	
	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE
Youths SFSS								
Intercept	46.15 ***	3.42	41.77 ***	4.64	44.46 ***	3.83	43.06 ***	4.38
Youth Race ^c	-0.01	1.19	0.36	1.37	-0.26	1.20	0.18	1.27
TA Change	0.79	0.72	-0.13	1.02	0.62	0.81	2.03	0.92
Slope (weeks)	-0.19 ***	0.02	-0.23 ***	0.02	-0.21 ***	0.02	-0.22 ***	0.02
Slope *TA Change	-0.04	0.02	0.03	0.02	-0.05 **	0.02	-0.01	0.02
Caregivers SFSS								
Intercept	49.27 ***	0.97	49.93 ***	1.12	49.38 ***	0.95	49.56 ***	0.99
Youth Race ^c	2.74 *	1.20	3.20 *	1.39	2.85 **	1.18	3.04 **	1.23
TA Change	-0.25	0.62	-0.61	0.90	-0.40	0.69	0.06	0.73
Slope (weeks)	-0.11 ***	0.02	-0.12 ***	0.03	-0.13 ***	0.02	-0.12 ***	0.02
Slope *TA Change	-0.01	0.03	0.01	0.03	-0.11 ***	0.03	-0.07 *	0.03
Clinicians SFSS								
Intercept	49.49 ***	0.94	50.18 ***	1.15	49.34 ***	0.92	49.58 ***	1.03
Youth Race ^c	2.27 *	1.03	2.93 *	1.35	2.39 *	1.02	2.45 *	1.18
TA Change	0.26	0.52	-1.06	0.79	0.49	0.58	0.25	0.63
Slope (weeks)	-0.15 ***	0.02	-0.17 ***	0.02	-0.15 ***	0.02	-0.16 ***	0.02
Slope *TA Change	-0.01	0.02	0.01	0.02	-0.07 ***	0.02	-0.03	0.02

Note:

^aTherapeutic Alliance quality Scale (TAQS) is a 5-point Likert scale (1 = not at all to 5 = totally), where higher scores indicate better alliance.

^bClinician therapeutic Alliance Quality Rating is a 5-point Likert scale (1 = very poor to 5 = excellent), where higher scores indicate higher alliance.

^cCaucasian = 1, otherwise = 0.

500 > $\bar{d}$ *
; 100 > $\bar{d}$ **
; 1000 > $\bar{d}$ ***

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 7

HLM Youth SFSS Results Controlling for Negative, Positive vs. No TA Change

SFSS Measures	Youth TA ^a		Caregiver TA ^a		Clinician Youth TA ^b		Clinician Caregiver TA ^b	
	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE
Youths SFSS								
Intercept	50.34 ***	2.96	45.13 ***	3.69	45.28 ***	3.10	46.75 ***	3.32
Youth Race ^c	0.19	1.10	0.96	1.21	-0.16	1.08	-0.02	1.12
Positive TA	0.56	1.31	1.39	1.45	0.54	1.23	0.58	1.28
Negative TA	-1.19	1.43	2.19	1.54	-2.55	1.57	-1.86	1.71
Slope (weeks)	-0.29 ***	0.04	-0.18 ***	0.03	-0.24 ***	0.04	-0.24 ***	0.03
Slope *Positive TA	0.06	0.05	-0.05	0.05	-0.01	0.04	0.05	0.04
Slope *Negative TA	0.19 ***	0.06	-0.05	0.05	0.13 ***	0.05	0.02	0.06
Caregivers SFSS								
Intercept	50.44 ***	1.00	49.11 ***	0.95	49.80 ***	0.95	49.95 ***	0.95
Youth Race ^c	2.56 *	1.08	2.70 *	1.07	2.55 *	1.06	2.49 *	1.08
Positive TA	-1.52	1.24	-0.05	1.31	-1.39	1.27	-1.02	1.27
Negative TA	-1.52	1.41	2.57	1.45	1.98	1.58	0.03	1.63
Slope (months)	-0.12 **	0.05	-0.09 **	0.04	-0.18 ***	0.05	-0.13 ***	0.04
Slope *Positive TA	0.01	0.06	-0.07	0.05	0.03	0.06	0.00	0.05
Slope *Negative TA	0.02	0.07	-0.01	0.06	0.13 *	0.06	0.06	0.06
Clinicians SFSS								
Intercept	49.11 ***	1.03	48.37 ***	0.94	49.19 ***	0.97	49.23 ***	0.95
Youth Race ^c	2.51 *	0.98	2.60 **	0.96	2.48 *	0.97	2.46 *	0.97
Positive TA	0.95	1.03	0.82	1.01	0.73	1.01	1.04	0.99
Negative TA	-0.29	1.21	4.06 **	1.16	0.12	1.27	-0.40	1.33
Slope (months)	-0.12 **	0.04	-0.10 **	0.03	-0.14 ***	0.04	-0.15 ***	0.03
Slope *Positive TA	-0.05	0.05	-0.05	0.04	-0.08	0.05	-0.01	0.04

SFSS Measures	Youth TA ^a		Caregiver TA ^a		Clinician Youth TA ^b		Clinician Caregiver TA ^b	
	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE	Est. Coeff	SE
Slope ^a Negative TA	0.02	0.06	-0.09	0.05	0.11 [*]	0.05	0.07	0.05

Note:

^aTherapeutic Alliance quality Scale (TAQS) is a 5-point Likert scale (1 = not at all to 5 = totally), where higher scores indicate better alliance.

^bClinicians therapeutic Alliance Quality Rating is a 5-point Likert scale (1 = very poor to 5 = excellent), where higher scores indicate higher alliance.

^cCaucasian = 1, otherwise = 0.

p< 0.001;

**
p< 0.01;

*
p< 0.05