

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Special Education and Communication Disorders  
Faculty Publications

Department of Special Education and  
Communication Disorders

---

2013

# Therapeutic Alliance Between Youth and Staff in Residential Group Care: Psychometrics of the Therapeutic Alliance Quality Scale

Kristin Duppong-Hurley  
*University of Nebraska-Lincoln*, [kristin.hurley@unl.edu](mailto:kristin.hurley@unl.edu)

Matthew C. Lambert  
*University of Nebraska-Lincoln*, [mlambert2@unl.edu](mailto:mlambert2@unl.edu)

Mark Van Ryzin  
*Oregon Social Learning Center*

Justin Sullivan  
*University of Nebraska-Lincoln*

Amy Stevens  
*Father Flanagan's Boys Town*

Follow this and additional works at: <https://digitalcommons.unl.edu/specedfacpub>

 Part of the [Special Education and Teaching Commons](#)

---

Duppong-Hurley, Kristin; Lambert, Matthew C.; Ryzin, Mark Van; Sullivan, Justin; and Stevens, Amy, "Therapeutic Alliance Between Youth and Staff in Residential Group Care: Psychometrics of the Therapeutic Alliance Quality Scale" (2013). *Special Education and Communication Disorders Faculty Publications*. 112.  
<https://digitalcommons.unl.edu/specedfacpub/112>

This Article is brought to you for free and open access by the Department of Special Education and Communication Disorders at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Special Education and Communication Disorders Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in final edited form as:

*Child Youth Serv Rev.* 2013 January 1; 35(1): 56–64. doi:10.1016/j.chilyouth.2012.10.009.

## Therapeutic Alliance Between Youth and Staff in Residential Group Care: Psychometrics of the Therapeutic Alliance Quality Scale

**Kristin Duppong Hurley,**  
University of Nebraska-Lincoln

**Matthew C. Lambert,**  
University of Nebraska-Lincoln

**Mark Van Ryzin,**  
Oregon Social Learning Center

**Justin Sullivan,** and  
University of Nebraska-Lincoln

**Amy Stevens**  
Father Flanagan's Boys Town

### Abstract

Therapeutic alliance has been frequently studied in individual counseling sessions; however, research on therapeutic alliance in residential settings for youth with mental health diagnoses has been limited. This may be due, in part, to the presence of multiple service providers often in caregiving roles. The purpose of this study was to examine the psychometric quality of a widely utilized measure of therapeutic alliance used in psychotherapy with youth in residential care where the treatment is provided by a trained married couple. We also compared the relationship between youth ratings of their male and female service provider, as well as examined correlations in ratings between youth and staff on therapeutic alliance. Finally, we investigated the direction, magnitude, and trajectory of change in therapeutic alliance over a 12-month period following admission into residential care. The method was a longitudinal assessment of 135 youth and 124 staff regarding therapeutic alliance over the course of 12 months or discharge from services. Results indicated strong psychometric properties and high correlations for youth ratings of both their male and female service providers. However, the correlation was low between youth and service provider ratings of alliance. Longitudinal analyses indicated that rates of therapeutic alliance changed over time.

### Keywords

Residential group homes; therapeutic alliance; psychometrics; adolescents; out of home care

---

© 2012 Elsevier Ltd. All rights reserved.

Corresponding Author: Kristin Duppong Hurley, Department of Special Education and Communication Disorders, 247E Barkley Memorial Center, University of Nebraska-Lincoln, Lincoln, NE 68583-0738, kristin.hurley@unl.edu, phone: 402-472-5501, fax: 402-472-7697.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## 1 Introduction

Therapeutic alliance, also known as the working relationship between a therapist and client, has been extensively researched in the context of adult psychotherapy with over 7,000 articles on the topic (Horvath, Del Re, Fluckiger, & Symonds, 2011). Meta-analyses that have been predominately focused on outpatient adult counseling have demonstrated that therapeutic alliance is a strong predictive indicator of treatment engagement and outcomes (e.g., Horvath & Bedi, 2002; Horvath et al., 2011; Horvath & Symonds, 1991; Martin, Grasko, & Davis, 2000; Sharf, Primavera, & Diener, 2010). Therapeutic alliance has been conceptualized as consisting of three components: the bond or affective components of the relationship, agreement on the tasks or activities of the therapy, and shared agreement on the goals of the therapy (Bordin, 1979). This perspective also often incorporates a perception of openness, honesty, and trustfulness within the working relationship (Doucette & Bickman, 2001).

Recently, therapeutic alliance research has been extended to child and adolescent-centered therapy. As with the adult research, a number of review studies have found evidence that therapeutic alliance between the youth and therapist is predictive of treatment outcomes (Karver, Handelsman, Fields, & Bickman, 2006; Shirk & Karver, 2003; Shirk, Karver, & Brown, 2011), although a few studies have found only weak associations among therapeutic alliance and youth outcomes (Green, 2006; McLeod, 2011). The vast majority of these studies with youth are focused on traditional, out-patient settings (e.g., McLeod 2011; Shirk et al., 2011). However, many youth receive mental health treatment in alternative settings, such as in-home family based services or out-of-home services such as residential group homes or treatment foster. Such settings bring with them a variety of challenges to the assessment of therapeutic alliance.

While 15% of youth in out-of-home placements are served in residential care (Kid's Count Data Center, 2012), there has been minimal research into the role of therapeutic alliance in this intensive treatment setting that serves youth 24/7. Most of the work on therapeutic alliance in out-of-home care has been descriptive in nature. For example, Manso, Rautkis and Boyd's (2008) study of therapeutic alliance in a youth wilderness camp setting conducted focus groups with 11 youth regarding their relationships with their counselors. This study found that youth identified three essential aspects of their counselors: the counselor's personal qualities (e.g., caring, trustworthiness), the counselor's behaviors (e.g., listening to youth, providing accurate feedback), and the relationship between the youth and counselor (e.g., respect, caring, trust). Notably, youth were also looking for counselors to be professional, yet provide a personal, more parental role. Another study based in a similar wilderness camp program found a minimal correlation between youth and counselor ratings of therapeutic alliance (Bickman et al., 2004), a discrepancy that has also been found in outpatient settings (e.g. Hawley & Garland, 2008). Findings from a descriptive study of therapeutic alliance in a residential setting in the Netherlands suggested that attachment representations to counselors could predict relationship quality (Zegers, Schuengel, Van IJzendoorn, & Janssens, 2006). Likewise, a descriptive study by Moses (2000) indicated that it was beneficial if counselors focused on rules and expectations when dealing with interpersonal conflicts among the youth, but used personalized, youth-focused approaches when addressing individual youth behavior.

Few studies of youth and therapeutic alliance examine how it predicts youth outcomes. A recent meta-analysis requiring temporal order of therapeutic alliance in predicting outcomes in more traditional outpatient settings found only 16 studies that meet the inclusion criteria (Shirk, Karver, & Brown, 2011). A couple of studies in residential settings have attempted to examine the role of therapeutic alliance in relation to the prediction of treatment

outcomes. For example, a large residential service provider examined the role of therapeutic alliance between youth and their professional mental health counselors (whom they visit with one to two hours per week) and found no relationship between therapeutic alliance and youth outcomes (Handwerk et al., 2008). Another study found that gains in therapeutic alliance with residential staff from intake to three months predicted improvement in youth outcomes; whereas youth with high alliance scores at intake seemed to be related to negative outcomes (Florsheim, Shortorbani, Gues-Warnick, Barrat, & Hwang, 2000). This finding is intriguing, as minimal research has been conducted on the developmental trajectories for youth in psychotherapy treatment. However, it is consistent with a recent study that found evidence suggesting that increasing youth therapeutic alliance scores over time was related to improved outcomes in an outpatient setting (Bickman et al., 2012). This is also supported by study in foster care that examined how therapeutic alliance scores for youth changed over time (Rautktis, Vides De Andrade, Doucette, McDonough, and Reinhart, 2005). They found that youth therapeutic alliance scores typically started high, declined, and then gradually began to increase. This study was limited by a small sample of 25 youth but is a promising approach for future research. Research seeking to understand the relationship between therapeutic alliance trajectories and outcomes would help to better inform practice and has led to a call for more research examining therapeutic alliance longitudinally (Bickman et al., 2012; Hawley & Garland, 2008; Shirk, Karver, & Brown, 2011).

Conducting research on therapeutic alliance for youth in residential settings is inherently complex. In some treatment settings, youth live in homes served by rotating shift staff, and thus may have different types of working relationships with the different providers. In other situations, such as Teaching Family Model programs (Wolf et al., 1976), youth reside with a married couple that lives with them daily. In these situations, the youth alliance would be shared between the male and female spouses. In both the married couple or shift-staff arrangements, the providers are the primary source of discipline and decision-making; far more like a parent role than a traditional clinician role. This brings to the forefront issues regarding the measurement of therapeutic alliance in residential settings.

There are few measures of therapeutic alliance with strong psychometric properties that have been used with youth in traditional psychotherapy settings (Bickman, Athay, Riemer et al., 2010; Bickman et al., 2012; Fjermestad et al., 2012; Hawley & Weisz, 2005; Horvath & Greenberg, 1989; Kazdin, Marciano, & Whitley, 2005; McLeod & Weisz, 2005; Shirk & Saiz, 1992; Tracey & Kokotovic, 1989). Early versions of youth therapeutic alliance measures were used in two studies of residential care (Bickman et al., 2004; Florsheim et al., 2000); however, the psychometric properties of the alliance measures were only briefly addressed. Moreover, the studies did not examine if there were differential ratings of alliance among the youth's residential providers. Thus, it is important to discover if the existing therapeutic alliance assessment approaches based upon the outpatient model can be modified to be used in out-of-home settings such as group residential care while retaining the strong psychometric properties of the assessments. If not, perhaps new therapeutic alliance instruments will need to be developed for use in residential care.

Consequently, the purpose of this study was three-fold: 1) to examine the psychometric quality of the TAQS (Bickman et al., 2010) a widely utilized measure of alliance used in outpatient settings; 2) to investigate the relationship among the male and female service provider TAQS ratings as provided by the youth, as well as compare the youth-provided TAQS ratings to the perspectives of the service providers; and 3) to examine the direction, magnitude, and trajectory of change in therapeutic alliance over a 12-month period following admission into care for a sample of youth in residential care. This was accomplished with a psychometric study to address goals one and two, and a longitudinal study to assess goal three.

## 2. Methods

### 2.1 Setting

The study was conducted at a residential facility that serves over 500 girls and boys in 70 family-style group homes in a large Midwestern city. The agency has implemented an adaptation of the Teaching Family Model (TFM; Davis & Daly, 2003; Wolf et al., 1976) since the 1980's and employs a married couple as the primary service delivery agents. This treatment-providing couple live in a family-style home with up to eight adolescent girls or boys. The focus of the TFM intervention is to teach youth skills to manage their behavior using coaching, practice, and feedback along with a token economy (Wolf et al., 1976).

### 2.2 Participants

Youth eligible to participate in the study were identified with a disruptive behavior diagnosis (via a professional diagnosis, Diagnostic Interview Schedule for Children (DISC; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000), or the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001), were at least 10 years old, were experiencing their first admission to the agency, and were assigned to service providers participating in the study (124 service providers participated in this study representing 62 group homes, almost 79% of the group homes at this agency). Based on these criteria, over a two year period, 170 youth were eligible for participation and 145 (85%) had guardian consent and youth assent to participate. All recruitment and consent procedures for youth and staff were approved by the University of Nebraska-Lincoln IRB and the agency IRB.

For the psychometric study, we used a subset of 135 youth that had responded to at least 80% of therapeutic alliance items in the first wave of data collection (about two months after admission). Youth demographics include 59 girls and 76 boys, 30 youth indicated that they were Hispanic, 71 Caucasian, 50 African American, and 29 other. Age at enrollment ranged from 10 to 17 years, with a mean age of 15.2 (SD = 1.39).

For the longitudinal study, we used a subset of the total sample consisting of youth who had at least two data points and did not change group home during the course of data collection. Thus, the sample for the longitudinal study consisted of 101 youth. We excluded youth that changed group-homes during data collection because their ratings of therapeutic alliance where with different service providers at different points in time.

### 2.3 Measures

The *Therapeutic Alliance Quality Scale* (TAQS; Bickman et al., 2010) is often used to measure the working relationship between a clinician and a youth receiving out-patient counseling. The items on the TAQS capture the bond between the client and clinician as well as the agreement on the tasks and goals of therapy. The TAQS was developed through an extensive iterative process, and comprehensive psychometric evaluations have been undertaken within the context of out-patient counseling settings (see Bickman et al., 2010 and Riemer et al., 2012). The current version of the TAQS consists of 5 items rated by youth on a 5-point scale (1 = *Not at all*, 4 = *Totally*). Higher total scores are associated with higher quality of therapeutic alliance. Mean total scores tend to be relatively high (4.12 on a 5-point scale) and negatively skewed (Bickman et al., 2010); most youth report good relationships with their clinician. The TAQS, by design, has only a single underlying factor (Bickman et al., 2012; Riemer et al., 2012) and includes two items on the bond between youth and the therapist and three questions regarding agreement on goals and tasks.

As the TAQS was designed to assess the relationship between a single clinician and a client in outpatient counseling settings, the assessment was slightly modified for use in residential

care where each youth has a therapeutic bond with two service providers. We reworded items (working closely with the developers of the TAQS) to fit the setting (substituting words referring to outpatient therapy or to a single session to be more reflective of residential treatment). We also asked youth to rate each service provider separately (i.e., provide a rating of alliance with the male service provider and a rating of alliance with the female service provider). We also had service providers rate their therapeutic alliance with youth using a single item used by the developers of the TAQS (Bickman et al., 2010; Bickman et al., 2012).

## 2.4 Procedures

Youth were given the TAQS as a part of an assessment battery for a larger, longitudinal study (Duppong Hurley, Lambert, & Sullivan, 2011) two months following admission to the residential group home, and then every 2 months thereafter. The final assessment was at discharge or after the 12<sup>th</sup> month of residential care, whichever occurred first for the youth. The assessments were given individually via an online survey. Research staff were present during the administration of the assessment to help the youth log-in to the computer and address any difficulties the youth might experience while completing the survey. Service providers completed their ratings of alliance on the same schedule as the youth. They also recorded their ratings via a web-based form, which they completed at their own convenience.

## 2.5 Data Analysis

**2.5.1 Psychometric study**—To investigate the psychometrics of the TAQS we used only cases from the first wave of data (collected two months after admission to the residential setting). We used classical tests theory (CTT), Rasch modeling, and confirmatory factor analysis (CFA) approaches to evaluate the psychometric quality of the TAQS. Each approach produces different indicators of the psychometric quality of the individual items and of the overall scale.

Rasch modeling provides detailed item-level indicators of psychometric quality such as item difficulty, item discrimination and item fit in addition to indicators of scale functioning (i.e., category characteristic curve plots). These indicators are sample independent and allow for the evaluation of person and item facets as different dimensions of psychometric information (Crocker & Algina, 1986). CFA modeling yields sample dependent indicators of psychometric quality, but provides a comprehensive evaluation of the internal structure of the assessment. The classical test theory approach provides summary statistics and familiar indicators of reliability such as Cronbach's alpha.

SPSS v19 was used to calculate item and scale means, standard deviations, skewness, kurtosis, internal consistency and correlations between the male and female service provider TAQS ratings provided by youth, as well as between ratings on the TAQS and the single-item service provider rating of alliance. Winsteps v3.73 (Linacre, 2012a) was used to fit a rating-scale Rasch model (Andrich, 1978) to the youth-provided TAQS data to estimate item difficulty and discrimination parameters, item fit indices and category characteristic curve plots. Item difficulty refers to the degree to which individuals endorse the item using high ratings; the more difficult the item the less likely an individual is to endorse the item with a high rating. While the Rasch model is a 1-parameter logistic model, item discrimination can be estimated in a subsequent model in Winsteps 3.73. Item discrimination refers to the degree to which items can identify individuals with high alliance from individuals with low alliance. Item fit refers to the degree to which observed responses to an item correspond to the expected responses given the difficulty of the items and the theta level of the examinee (i.e., the quality of alliance between the youth and the service provider). Infit statistics

represent fit for items for examinees in the middle of the distribution of the sample and outfit statistics represent fit for items for examinees in the tails of the distribution of theta (i.e., extreme measure scores) (Bickman et al., 2010).

The purpose of fitting data to a Rasch model was to explore the item-level psychometric quality of the assessment and overall rating scale functioning. Evidence of the item-level psychometric quality of the TAQS has been provided for clinical populations by Bickman et al. (2012) and Bickman et al. (2010). For evidence of item-level psychometric quality, we focused on item difficulty, discrimination and fit. For evidence of rating scale functioning, we focused on the category characteristic curves plots and the corresponding statistics.

Mplus v6 (Muthén & Muthén, 1998–2009) was used to fit confirmatory factor analysis (CFA) models to the TAQS data. We fit the youth-rated male service provider and female service provider data to separate CFA models. CFA models were specified such that all items loaded on a single latent factor without correlated residuals. As usual, the factor loading of the first item was fixed and all other parameters were freely estimated. Missing data for the CFA models were treated as *missing at random* (MAR) and included in the analysis by using the full information maximum likelihood (FIML) estimator. We chose to use direct maximum likelihood estimation rather than listwise deletion or imputation because FIML is a robust method for handling missing data, does not reduce statistical power and produces consistent, efficient and asymptotic normal parameters (Allison, 2009).

The purpose of fitting data to a CFA model was to establish that the internal structure of the TAQS was the same for youth in residential settings as for youth in outpatient counseling settings. Evidence of the internal structure of the TAQS has been provided for clinical populations by Bickman et al. (2012) and Bickman et al. (2010). For evidence of internal structure, we focused on goodness-of-fit indicators and factor loadings.

**2.5.2 Longitudinal study**—Mplus v6 was used to estimate a series of latent curve models (LCM) to address two main questions: 1) what was the initial level of therapeutic alliance, and 2) did alliance change over participants' stay in care? Four separate models were fit: 1) youth TAQS ratings of male service providers, 2) youth TAQS ratings of female service providers, 3) single-item staff rating provided by the male service provider, and 4) single item staff rating provided by the female service provider. Figure 1 is the path diagram for the latent curve models (specifically the youth-rated TAQS models). TAQS models were estimated as curvilinear (i.e., with a quadratic term ( $\beta_2$ )) and the staff-rating models were estimated as linear. Staff-rating models were estimated as linear because initial analyses indicated that the means and variances of the quadratic terms for the staff-rating models were non-significant (note that only results for the linear models are presented in the results section). All models were estimated using individually-varying times of observations with random *growth factors* (e.g., intercept and slopes). FIML was used to estimate all models and missing data was included in the analysis.

### 3. Results

#### 3.1 Psychometric Study

Table 1 contains item-level descriptive statistics including the mean, standard deviation, skewness and kurtosis for TAQS items. The TAQS item about the youth understanding their treatment was rated the highest by the youth. The item about the service provider understanding what it feels like to be the youth was rated the lowest (most difficult to rate high). There were no differences in these rankings for TAQS ratings of the female or the male service provider (see Table 1). Most items were negatively skewed indicating that the majority of youth reported good relationships and only a few youth reporting low alliance

with service providers. As the item difficulty measures suggest, most items were easy for youth to endorse high on the rating scale, leading to mostly high scores with only a few low scores. Several items also have high kurtosis values indicating less spread around the mean than would be found in a normal distribution.

Total TAQS scores were computed for the youth ratings of therapeutic alliance with the male or female service provider by summing the ratings for the items and dividing the sum by the number of items. The descriptive statistics and internal consistency for the total scores are reported in Table 2. Total scores for youth ratings of male and female service providers were highly correlated ( $r = .788$ ,  $n = 133$ ). Using the cut-scores from the TAQS manual, 39.6% of the sample had low alliance, 47% had medium alliance, and 13.4% had high alliance with the male service providers. Proportions were similar for youth alliance scores with female service providers (37.8%, 48.9%, and 13.3%).

**3.1.1 Rasch**—The results presented for the Rasch model include only non-extreme cases ( $n = 117$ ). Extreme cases represent examinees that had maximum or minimum total scores; these cases are excluded because theta ( $\theta$ )<sup>1</sup> cannot be estimated reliably for these examinees as theta can take on an infinite amount of different values for these individuals (Bond & Fox, 2002). In this case, theta (also referred to as *measure* in Rasch output) represents the level of therapeutic alliance quality. For person parameters, higher measure values indicate higher alliance quality as rated by the youth. For item parameters, higher measure values indicate more difficult items for youth to endorse high on the rating scale (i.e., a youth would need to have a high level of alliance with the service provider to endorse a difficult item with a high rating); effective tests include a mix of easy to difficult items (Bickman et al., 2010).

Item level indicators of psychometric quality are presented in Table 1. All of the items on the TAQS except for the fourth item in the table (service provider would help) demonstrate adequate or desirable item fit and discrimination properties (Bond & Fox, 2002; Linacre, 2006). Acceptable unstandardized infit and outfit values range from 0.5 to 1.5. The discrimination parameter was centered at 1 for the analyses, so items with values below 1 indicate less discriminating efficacy (i.e., the slope of the logistic curve is flatter when discrimination values are less than 1). The fourth item has a relatively low discrimination parameter (.67 and .66 for ratings with male and female service providers, respectively) indicating that the item is not as useful for distinguishing between youth with high alliance and youth with low alliance (Bond & Fox, 2002). This finding is consistent for youth-ratings of both male and female service providers.

In addition to adequate item fit indicators, both the item and person reliability statistics were acceptable (.98 and .77, respectively) indicating that the model can accurately place items and people on the theta continuum. It should be noted that, ideally, the person reliability would be above .80 (Chapman et al., 2008) corresponding to at least a 2.00 separation index. The separation index indicates the relative number of statistically different groups of examinees. Therefore, a separation index less than 2 indicates less than two statistically distinguishable groups of examinees.

The items, as a set, were quite easy for examinees to endorse high as indicated by the item means. Further evidence is provided by the results of the Rasch model. The item facet was centered to aid with the interpretation of the mean difficulty of the assessment (Linacre, 2012b) (i.e., the mean measure score for the set of items was centered at zero). In such a model, a test with a fairly even mix of difficult to easy items will yield a mean measure statistic around zero. The mean measure statistics for the TAQS with male service providers and with female service providers were 1.06 ( $sd = 1.42$ ) and .93 ( $sd = 1.35$ ), respectively.



The larger than expected means and standard deviations indicate that youth, on average, rate the therapeutic alliance with service providers high with more spread around the mean than would be normally expected.

The category characteristic curve plot (see Figure 2) is presented for only the youth ratings of the male service provider due to space considerations. The plot for the youth ratings of the female service provider is nearly identical to the plot for the male service provider. The curves in the plot represent the mean (across all items) likelihood that an individual with a given theta level (i.e., level of therapeutic alliance) would endorse items using a certain rating category. Each curve should peak at a unique place on the theta continuum representing that each category is distinguishable from one another (Linacre, 2002). The curves should also be ordered monotonically from *Not at all* to *Totally*. This is achieved for all categories except category two (*Only a little*). This suggests that youth completing the TAQS cannot differentiate between categories one (*Not at all*) and two (*Only a little*).

**3.1.2. Association between TAQS and staff-ratings**—The single-item staff-rating of alliance indicated, on average, high alliance with youth;  $m = 3.90$  ( $sd = 0.62$ ) for male staff-ratings and  $m = 3.87$  ( $sd = 0.67$ ) for female staff-ratings. The relationship between male and female staff ratings and between staff and youth ratings were evaluated using Spearman Rho correlations ( $\rho$ ) since the staff ratings are categorical in nature. Male and female staff-ratings were significantly correlated ( $\rho = .418$ ) indicating that they view therapeutic alliance with youth in a similar way. We also tested male single-item staff ratings with the youth TAQS scores for the male service provider, and female single-item staff ratings with the youth TAQS scores for the female service provider. With regard to the male service providers, the magnitude of the association between the youth rated TAQS total scores and the single-item staff ratings was moderate and statistically significant ( $\rho = .181$ ). There was no statistically significant association between the youth TAQS total score and the single-item ratings for the female service providers ( $\rho = .064$ ).

**3.1.3. CFA**—The goodness-of-fit indicators listed in Table 3 include chi square ( $\chi^2$ ; Pearson, 1900), comparative fit index (CFI; Bentler, 1990), Tucker-Lewis index (TLI; Tucker & Lewis, 1973), root mean square error of approximation (RMSEA; Steiger & Lind, 1980), and the mean square residual (SRMR). Each indicator gauges fit or misfit in unique ways and therefore multiple indices are reported.

Exact tests of model fit, such as  $\chi^2$ , test the hypothesized model against a perfect null model (Chen, Curran, Bollen, Kirby, & Paxton, 2008). A significant  $\chi^2$  statistic indicates misfit to the data; however, this test is often considered too conservative of a measure of fit (Browne & Cudeck, 1993). Comparative tests of fit such as the TLI or the CFI test the hypothesized model against a null model derived from the noncentrality parameter (Boomsma, 2000). It is generally considered that a good fitting model has a TLI or CFI greater than .95 (Brown, 2006); adequate fit is considered to be between .90 and .95 (Browne & Cudeck, 1993). Absolute tests of fit such as the RMSEA or SRMR test the magnitude of error between the observed covariance matrix and the predicted covariance matrix (Boomsma, 2000). Both of these absolute tests are standardized and reported on a scale of 0 to 1; values closer to zero indicate better fit. It is widely held that RMSEA and SRMR point estimates less than .05 indicate good fit while values between .05 and .08 indicate adequate fit (Hu & Bentler, 1999).

The RMSEA approach also yields confidence intervals (CI) around the point estimate that can be used to test model fit. There are three tests of significance for the RMSEA confidence intervals (see Chen et al, 2008); we used only the test of exact fit ( $H_0: \epsilon = 0$ ; where  $\epsilon$  is the

lower CI). Since the lower bound of the confidence interval is 0 (see Table 3), we fail to reject the null for both models (i.e., the models fit the data).

All of the indicators of goodness-of-fit support the conclusion that the CFA model fits the TAQS data quite well and unstandardized and standardized factor loadings were all positive and large in magnitude ranging from .60 to .84 (see Table 4). We can be confident that, for this sample, the TAQS has only a single underlying factor.

### 3.2 Longitudinal Study

The latent curve models were fit with random growth factors: every individual was allowed to take on unique values for intercept ( $\alpha$ ), slope ( $\beta_1$ ) and quadratic terms ( $\beta_2$ ) (if quadratic terms were included in the particular model). The intercept represents the model-implied initial level of therapeutic alliance upon admission into the residential setting. The slope factor takes on different interpretations for linear and quadratic models. For linear models, the slope factor describes the change over time (i.e., the change that occurs between any two adjacent time points). For quadratic models, the slope describes the instantaneous change (Bollen & Curran, 2006); the change that occurs between the first and second data points, or in more technical terms, the slope factor in quadratic models represents the tangent to the curve. The quadratic factor describes how the slope (i.e., the rate of change) changes over time (i.e., acceleration or deceleration of change).

Since individuals were allowed to take on unique values for growth factors, we can represent the distribution of factor scores with means and variances. A statistically significant mean for the intercept growth factor indicates that, for the sample as a whole, the model-implied intercept is different than zero. A statistically significant variance for the intercept growth factor indicates significant inter-individual differences in model-implied initial therapeutic alliance. A significant mean for the slope growth factor indicates that, for the sample as a whole, therapeutic alliance is changing significantly over time. A significant variance for the slope factor indicates significant inter-individual differences in magnitude and/or direction of change over time. Lastly, a significant mean quadratic factor indicates that, for the sample as a whole, the rate of change (i.e., the slope) significantly accelerates or decelerates over time. A significant variance for the quadratic factors indicates significant inter-individual difference in the acceleration or deceleration of change over time. Table 5 contains the fit indicators and the growth factor means and variances for all four therapeutic alliance models.

**3.2.1. Model fit**—The log-likelihood ratio test (LRT) is presented as the model fit indicator. The LRT is interpreted as a  $X^2$  statistic; a significant statistic represents misfit to the data. Although LRT is a conservative indicator of fit similar to  $X^2$  and therefore the Type I error rate is inflated (Browne & Cudeck, 1993), there are no other interpretable model fit indicators available when estimating a latent curve model with individually-varying times of observations (Muthén & Muthén, 1998–2009). The log-likelihood ratio tests suggested that only models for the youth ratings of the male service provider (LRT = 7.4) and the female service provider ratings of the youth (LRT = 18.7) fit the data well. The other two models demonstrated less than adequate fit to the data as indicated by the significant LRT values.

**3.2.2. Initial level of therapeutic alliance**—The mean model-implied initial level of alliance was 3.88 for the youth-rated TAQS for the male service provider model and 3.79 for the TAQS model for the female service provider. According to cut-off values used by the TAQS developers to categorize youth into high, medium and low alliance groups, the initial levels of alliance for this sample were low to medium. The variances of the intercepts were

both significant indicating that there were substantial inter-individual differences in initial therapeutic alliance. Model-implied initial levels of therapeutic alliance for youth TAQS ratings of their male and female service providers' ratings were significantly correlated ( $r = .83$ ).

The mean model-implied initial levels of therapeutic alliance as rated by staff using a single item were 3.95 for the male-rated model and 3.96 for the female-rated model. Again, the variances of both intercepts were significant indicating inter-individual differences in initial therapeutic alliance ratings. Model-implied initial levels of therapeutic alliance for male and female staff single-item alliance ratings were significantly correlated ( $r = .78$ ).

**3.2.3. Change in therapeutic alliance**—Therapeutic alliance as rated by the youth on the TAQS changed significantly over the 12-month period for ratings for the male service provider, but not the female service provider; however, youth TAQS ratings for both the male and female service providers followed the same general trajectory. Youth TAQS ratings, on average, decreased initially, leveled off and increased toward the end of care. TAQS ratings by the youth for both their male and female service providers indicated that the intercept, slope and quadratic factors had significant variances indicating substantial inter-individual differences in trajectories.

Therapeutic alliance as rated by staff using the single-item did not change significantly. The male staff ratings, on average, indicated slight decreases in alliance over time. The female staff ratings indicated a slight increase in alliance with youth over time. The variances of the male staff ratings growth factors were significant indicating substantial inter-individual differences in trajectories, but the female-staff score growth factor variances were non-significant except for the intercept factor, indicating that change was rather consistent across individuals in the sample. It should be noted that, while not presented here, the staff-rating models were also fit with quadratic terms, but for both male and female ratings the means and variances of the quadratic terms were non-significant indicating that the trajectories were in fact linear.

**3.2.4. Association between youth-rated TAQS and single-item staff-ratings**—Model-implied correlations between youth TAQS total scores and staff ratings over time ranged from  $r = .07 - .33$  for ratings of the male service provider and  $r = .09 - .52$  for ratings of the female service provider. The correlations between the staff-rated and the TAQS model-implied initial levels of therapeutic alliance were significant for the male service provider ( $r = .50$ ) and female service provider ( $r = .33$ ). We only compared the TAQS and staff-rated trajectories at the intercept because TAQS models were quadratic, staff-rated models were linear, and the interpretations of the slope parameters differed between the models.

## 4. Discussion

The primary goal of this study was to examine if the psychometric properties of a measure of youth therapeutic alliance (TAQS) were similar for youth in a residential care setting in comparison to outpatient clinical settings. The findings suggest that the psychometrics of the TAQS for the sample of youth and service providers in a 24/7, family-style, residential setting were very similar to the original outpatient clinical samples (Bickman et al., 2010). Thus, the short five-item TAQS seems a promising measure for assessing therapeutic alliance for youth in residential care.

Most indicators of psychometric quality suggest that the TAQS performs well with youth in residential care. Both the Cronbach's alphas for the scale and the unidimensional factor

structure are acceptable and similar to the clinical sample. The Rasch analyses also indicate the TAQS is performing adequately; although, the findings of the Rasch models did suggest that the tests were relatively 'easy'. Youth rated alliance high on four out of the five items on the TAQS. This finding is consistent with recent studies in outpatient settings (Bickman et al., 2012); however, this may hinder the interpretation of youth standing on the underlying factor due to potential ceiling effects. This is further evidenced by the acceptable, but relatively low person reliability (.77) and the consequently low person separation index indicating just less than two distinct levels of alliance (i.e., the test is not quite as sensitive as it perhaps could be for this population; Linacre, 2002). Substantive interpretation of therapeutic alliance may be aided by increasing the person separation index of the test (i.e., Rasch person reliability) by adding a few more difficult items representing higher levels of therapeutic alliance.

In addition, the Rasch measurement model indicated that there was some difficulty differentiating between the two lowest response options (see Figure 2), likely due to the tendency to rate alliance high. Linacre (2002) suggests that researchers should collect at least 10 observations of each rating category on each item to estimate category characteristic curves. Our data do not always meet this rule of thumb. We do not always observe adequate frequencies of low categories (i.e., *Not at all* and *Only a little*) and therefore the estimated category characteristic curves for those categories may not be completely reliable. Alternatively, there is evidence that other, similar common therapeutic process factor assessments display this same pattern of category characteristic curves (Riemer & Kearns, 2010). Researchers may consider using a three point scale as most respondents rate alliance in the top three categories anyway. Such a change may help scale properties align more consistently with Rasch measurement properties. However, it still may be useful to have the four-point scale to differentiate those rare, but noteworthy, occasions when youth do rate therapeutic alliance as extremely low.

The TAQS item and total score means were lower than those reported in the outpatient settings (Bickman et al., 2010; Bickman et al., 2012). While there are several possible explanations for the lower therapeutic alliance ratings by youth in residential care than outpatient care, one likely reason is the difference in settings between outpatient and residential services. First, many youth may have low therapeutic alliance ratings as they have typically been involuntarily removed from their home environment (including family, friends, school) to live with their service providers and other teens facing similar issues, and thus may be resentful to and detached from the service provider. Likewise, one role of the service providers is that of disciplinarian, routinely administering positive and negative consequences to the youth, which likely influences the youth's ratings of the working relationship. Clinical therapists are not placed in similar types of 24/7 parenting and disciplinarian roles with their clients; thus, youth ratings of their interactions may be more positive.

The youth ratings of the male and female service providers were very similar, with a correlation of .79. This suggests that few youth are creating differential relationships with the male and female service providers. The psychometric findings are also essentially the same for youths' ratings of male and female service providers. These findings suggest that in future studies of therapeutic alliance in residential settings, researchers may not need to measure alliance between the youth and both the male and female service providers. However, it should be noted that while ratings of alliance with male and female service providers were highly correlated, growth process displayed small, but significant differences. While the direction and trajectory of change was similar for alliance with male and female service providers, the magnitude of change was greater for alliance with male service providers.

The single-item staff ratings indicated high levels of therapeutic alliance from the perspective of the adult service providers. The mean item values were similar between the male and female providers. The perspectives of the two staff-raters are also highly correlated ( $\rho = .418$ ). Interestingly, the therapeutic alliance ratings among the service providers and youth were not highly correlated ( $\rho = .181$  and  $.064$ ) in the psychometric study, similar to the findings of previous research that have not found strong agreement between youth and clinician ratings (Bickman et al., 2012; Hawley & Garland, 2008). This suggests that youth and the service providers do not rate therapeutic alliance in a similar fashion and is also supported by other research in out-of-home care settings (Bickman et al., 2004; Rauktis, Vides De Andrade, Doucette, McDonough, & Reinhart, 2005). However, the model-implied intercepts of the latent curve models are more highly correlated for both youth and staff ratings ( $r = .5$  for male service providers and  $.33$  for female service providers) for the longitudinal study sample. This apparent discrepancy could be a function of using different samples between the two studies or simply that the two sets of correlations are of different variables. The correlations presented for the psychometric study were between the manifest variables while the correlations presented for the longitudinal study were between the latent intercept factors. It is important to note that the manifest single-item staff ratings were categorical while the latent intercept factor scores were continuously distributed. In addition, the Spearman Rho correlations between the staff and youth ratings may be artificially low because of the consistently high ratings of alliance by staff (Field, 2009). Given these seemingly inconsistent findings, it is difficult to identify the true relationship between youth- and staff-ratings of therapeutic alliance. More research is needed in this area, with larger sample sizes, to better elucidate these findings.

The longitudinal growth model analyses indicated considerable differences in therapeutic alliance trajectories among youth, suggesting that there were significant individual differences in therapeutic alliance patterns over time. However, the overall model suggests that for many youth there was a tendency for alliance ratings to decline and then slowly rebound which seems in line with the experiences of many youth in residential care; it takes time to get used to the new treatment setting, but eventually the situation improves. Other research in out-of-home care has found a similar trajectory of declining ratings in alliance and begins to rebound near the end of services (Rauktis, Andrade, Doucette, McDonough, & Reinhart, 2005). It will be helpful for future research to examine these trajectories for therapeutic alliance overtime to see if there are specific sub-groups, or profiles, of youth that respond to their service providers in different ways. A descriptive study (Rauktis et al., 2005) began to examine therapeutic alliance trajectories for youth with different characteristics, which is an approach that future research could replicate. Likewise, it would be interesting to investigate if there were differences among youth that had discrepancies in their ratings of alliance between their male and female service provider, compared to those that had similar ratings between providers. It would also be helpful to examine if different trajectories of youth therapeutic alliance were predictive of youth outcomes.

### 3.3 Limitations

The most serious limitations surround the sample of youth in residential care. First, this study was limited to youth with a disruptive behavior diagnoses on their first admission to this residential service provider. It is possible that therapeutic alliance ratings would vary for youth without significant behavior needs or that were readmitted to group care. Future research should examine therapeutic alliance with more diverse samples of youth in residential care. Second, data were collected from a single treatment facility in the Midwest. Future research would be strengthened with the inclusion of data from a variety of facilities. This study was also unable to collect ratings of the parental therapeutic alliance with the service provider, which is another important aspect of the working relationship surrounding

a child's care (Hawley & Garland, 2008). Finally, while the study was able to collect multiple instances of therapeutic alliance, it only occurred every other month and the data was not shared with agency staff. It would be helpful to examine if the routine and frequent collection of therapeutic alliance data could be used as feedback to the agency to improve the working relationship between youth and service providers.

An analysis of missing data revealed that the assumptions under which the models used for the longitudinal study were estimated might not hold as true. We found evidence that missing data was not missing at random starting at the 10-month observations for youth TAQS ratings of the male-service provider and at the 8-month observation for youth TAQS ratings of the female service provider. Missingness was related to low initial youth provided TAQS ratings for both the male and female service providers. We estimated the models under the assumptions of missing at random because there is no consensus on how to handle data missing not at random in latent curve models (Bollen & Curran, 2006) and we believe that the association between low alliance and missingness was likely an artifact caused by some youth having shorter than 12-month stays in care which is actually related to a number of external factors such as limited funding sources or court-orders mandating other placements.

### 3.4 Implications

This study provides support that the psychometric properties of a version of the TAQS modified for residential care can be replicated in a residential care setting, thus offering residential providers an efficient method for assessing therapeutic alliance. Even though the settings between outpatient and residential care are markedly different, it is encouraging that a similar approach to assess treatment alliance can be used. Most importantly, the ability to assess therapeutic alliance in residential care with a brief, five-item measure opens up opportunities for care providers to routinely monitor levels of therapeutic alliance between youth and staff. Both researchers and practitioners in residential care settings can begin to study and monitor therapeutic alliance over time to examine if specific profiles of youth have differing trajectories of therapeutic alliance during care. Moreover, continued advances in the assessment of therapeutic alliance may lead to improvements in matching the optimal therapeutic relationship fit between youth and providers, as well as methods to examine the role of therapeutic alliance on youth outcomes during residential care and following discharge.

### Acknowledgments

We appreciate the work of Heidi Menard in editing and formatting the manuscript and the participation of the youth and staff at the residential groups homes that made this study possible.

The research reported herein was supported, in part, by the National Institute for Mental Health, through Grant R34 MH080941 and Institute of Education Sciences, U.S. Department of Education, through Grant R324B110001 to the University of Nebraska-Lincoln. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Dr. Duppong Hurley is an investigator with the Implementation Research Institute (IRI), at the George Warren Brown School of Social Work, Washington University in St. Louis; through an award from the National Institute of Mental Health (R25 MH080916-01A2) and the Department of Veterans Affairs, Health Services Research & Development Service, Quality Enhancement Research Initiative (QUERI).

### References

Achenbach, T.M.; Rescorla, L.A. *Manual for ASEBA School-Age Forms & Profiles*. Burlington: University of Vermont, Research Center for Children, Youth, & Families; 2001.

- Allison, PD. Missing data. In: Millsap, R.; Maydeu-Olivares, A., editors. *The Sage handbook of quantitative methods in psychology*. London, UK: Sage; 2009. p. 72-89.
- Andrich D. A rating formulation for ordered response categories. *Psychometrika*. 1978; 43:561–573.
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107(2):238–246. [PubMed: 2320703]
- Bickman, L.; Athay, MM.; Riemer, M.; Lambert, EW.; Kelley, SD.; Breda, C.; Andrade, AR., editors. *Manual of the Peabody Treatment Progress Battery*. 2nd ed.. Nashville, TN: Vanderbilt University; 2010. [Electronic version].
- Bickman L, Vides de 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. *The Journal of Behavioral Health Services and Research*. 2004; 31:134–148.
- Bickman L, Vides de Andrade AR, Athay MM, Chen JI, De Nadai AS, Jordan-Arthur BL, Karver MS. The relationship between change in therapeutic alliance ratings and improvement in youth symptom severity: whose ratings matter the most? *Administration and Policy in Mental Health*. 2012; 39:78–89. [PubMed: 22407555]
- Bollen, KA.; Curran, PJ. *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley; 2006.
- Bond, TG.; Fox, CM. *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum; 2002.
- Boomsma A. Reporting analyses of covariance structures. *Structural Equation Modeling*. 2000; 7:461–483.
- Bordin E. The generalizability for the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice*. 1979; 16:252–260.
- Brown, T. *Confirmatory factor analysis for applied research*. New York, NY: Gilford Press; 2006.
- Browne, MW.; Cudeck, R. Alternative ways of accessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Newbury Park, CA: Sage; 1993. p. 136-162.
- Chapman JE, Sheidow AJ, Henggeler SW, Halliday-Boykins CA, Cunningham PB. Developing a measure of therapist adherence to contingency management: An application of the many-facets Rasch model. *Journal of Child and Adolescent Substance Abuse*. 2008; 17(3):47–68. [PubMed: 20871744]
- Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research*. 2008; 36:462–494. [PubMed: 19756246]
- Crocker, L.; Algina, J. *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt; 1986.
- Davis, J.; Daly, DL. *Long-Term Residential Program training manual*. 4th ed.. Boys Town, NE: Father Flanagan's Boy's Home; 2003.
- Doucette, A.; Bickman, L. *Child Adolescent Measurement System: User manual*. TN: The Center for Mental Health Policy, Vanderbilt University; 2001. Unpublished manuscript
- Duppong Hurley, K.; Lambert, M.; Sullivan, J. Examining implementation from many perspectives: how different are the views of implementation quality from observers, supervisors, staff, and clients?. Paper presented at the American Evaluation Association Annual Conference; Anaheim, CA. 2011 Nov.
- Field, A. *Discovering statistics using SPSS*. 3rd ed.. Los Angeles, CA: Sage; 2009.
- Fjermestad KW, McLeod BD, Heiervang ER, Havik OE, Öst L, Haugland BSM. Factor structure and validity of the therapy process observational coding system for child psychotherapy-alliance scale. *Journal of Clinical Child & Adolescent Psychology*. 2012; 41:246–254. [PubMed: 22417197]
- Florsheim P, Shotorbani S, Guest-Warnick G, Barratt T, Hwang WC. Role of the working alliance in the treatment of delinquent boys in community-based programs. *Journal of Clinical Child Psychology*. 2000; 29:94–107. [PubMed: 10693036]
- Green J. Annotation: The therapeutic alliance – a significant but neglected variable in child mental health treatment studies. *Journal of Child Psychology and Psychiatry*. 2006; 47:425–435. [PubMed: 16671926]

- Handwerk ML, Huefner JC, Ringle JL, Howard BK, Soper SH, Almquist JK, Chmelka MB. The role of therapeutic alliance in therapy outcomes for youth in residential care. *Residential Treatment for Children & Youth*. 2008; 25:145–165.
- 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: Journal of Research and Practice in Children's Services*. 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. *Journal of Clinical Child and Adolescent Psychology*. 2005; 34:117–128. [PubMed: 15677286]
- Horvath, AO.; Bedi, RP. The alliance. In: Norcross, JC., editor. *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. New York, NY: Oxford University Press; 2002. p. 37-69.
- Horvath AO, Del Re AC, Fluckiger C, Symonds D. Alliance in individual psychotherapy. *Psychotherapy*. 2011; 48(1):9–16. [PubMed: 21401269]
- Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*. 1989; 36:223–233.
- Horvath A, Symonds B. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*. 1991; 38:139–149.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6(1):1–55.
- Karver MS, Handelsman J, 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. *Clinical Psychology Review*. 2006; 26:50–65. [PubMed: 16271815]
- Kazdin AE, Marciano PL, Whitley MK. The therapeutic alliance in cognitive-behavioral treatment of children referred for opposition, aggressive, and antisocial behavior. *Journal of Consulting and Clinical Psychology*. 2005; 73:726–730. [PubMed: 16173860]
- Kids Count Data Center. 2012 <http://datacenter.kidscount.org/data/acrossstates/Rankings.aspx?loct=2&by=a&order=a&ind=6247&dtm=12994&ch=a&tf=133>.
- Linacre JM. Optimizing rating scale category effectiveness. *Journal of Applied Measurement*. 2002; 3:85–106. [PubMed: 11997586]
- Linacre JM. Misfit diagnosis: infit outfit mean-square standardized. 2006 from <http://www.winsteps.com/winman/diagnosingmisfit.htm>.
- Linacre, JM. Winsteps® Rasch measurement computer program. Beaverton, OR: 2012a. [Winsteps.com](http://www.winsteps.com).
- Linacre, JM. Winsteps® Rasch measurement computer program user's guide. Beaverton, OR: 2012b. [Winsteps.com](http://www.winsteps.com).
- Manso A, Rauktis ME, Boyd AS. Youth expectations about therapeutic alliance in a residential setting. *Residential Treatment for Children and Youth*. 2008; 25(1):55–72.
- Martin D, Graske J, Davis M. 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]
- McLeod BD. The relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical Psychology Review*. 2011; 31:603–616. [PubMed: 21482319]
- McLeod BD, Weisz JR. The Therapy Process Observational Coding System-Alliance Scale: Measure characteristics and prediction of outcome in usual clinical practice. *Journal of Consulting and Clinical Psychology*. 2005; 73:323–333. [PubMed: 15796640]
- Moses T. Attachment theory and residential treatment: A study of staff-client relationships. *American Journal of Orthopsychiatry*. 2000; 70(4):474–490. [PubMed: 11086526]
- Muthén, LK.; Muthén, BO. Mplus user's guide. 6th ed.. Los Angeles, CA: Muthén & Muthén; 1998–2009.
- Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*. 1900; 50(5):157–175.



- Rauktis ME, Andrade AR, Doucette A, McDonough L, Reinhart S. Treatment foster care and relationships: Understanding the role of therapeutic alliance between youth and treatment parent. *International Journal of Child & Family Welfare*. 2005; 4:146–163.
- Riemer M, Athay MM, Bickman L, Breda C, Douglas Kelley S, Vides de Andrade AR. The Peabody Treatment Progress Battery: History and methods for developing a comprehensive measurement battery for youth mental health. *Administration and Policy in Mental Health*. 2012; 39:3–12. [PubMed: 22421933]
- Riemer M, Kearns MA. Description and psychometric evaluation of the youth counseling impact scale. *Psychological Assessment*. 2010; 22(2):259–268. [PubMed: 20528053]
- Shaffer D, Fisher P, Lucas C, Dulcan M, Schwab-Stone M. NIMH Diagnostic Interview Schedule for Children version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2000; 39:28–38. [PubMed: 10638065]
- Sharf J, Primavera LH, Diener MJ. Dropout and therapeutic alliance: A meta-analysis of adult individual psychotherapy. *Psychotherapy Theory, Research, Practice, Training*. 2010; 41(4):637–645.
- Shirk SR, Karver M. Prediction of treatment outcome from relationship variables in child and adolescent therapy: A meta-analytic review. *Journal of Consulting & Clinical Psychology*. 2003; 71(3):452–464. [PubMed: 12795570]
- Shirk SR, Karver MS, Brown R. The alliance in child and adolescent psychotherapy. *Psychotherapy*. 2011; 48:17–24. [PubMed: 21401270]
- Shirk SR, Saiz C. The therapeutic alliance in child therapy: Clinical, empirical, and developmental perspectives. *Development and Psychopathology*. 1992; 4:713–728.
- Steiger, JH.; Lind, JC. Statistically based tests for the number of common factors; Paper presented at the annual meeting of the Psychometric Society; Iowa City, IA. 1980 May.
- Tracey TJ, Kokotovic AM. Factor structure of the Working Alliance Inventory. *Psychological Assessment*. 1989; 1:207–210.
- Tucker L, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38(1):1–10.
- Wolf MM, Phillips E, Fixsen D, Baukman CJ, Kirigin KA, Willner AG, Schumaker J. Achievement Place: The Teaching-Family Model. *Child Care Quarterly*. 1976; 5:92–103.
- Zegers MAM, Schuengel C, Van IJzendoorn MH, Janssens JMAM. Attachment representations of institutionalized adolescents and their professional caregivers: Predicting the development of therapeutic relationships. *American Journal of Orthopsychiatry*. 2006; 76:325–334. [PubMed: 16981811]

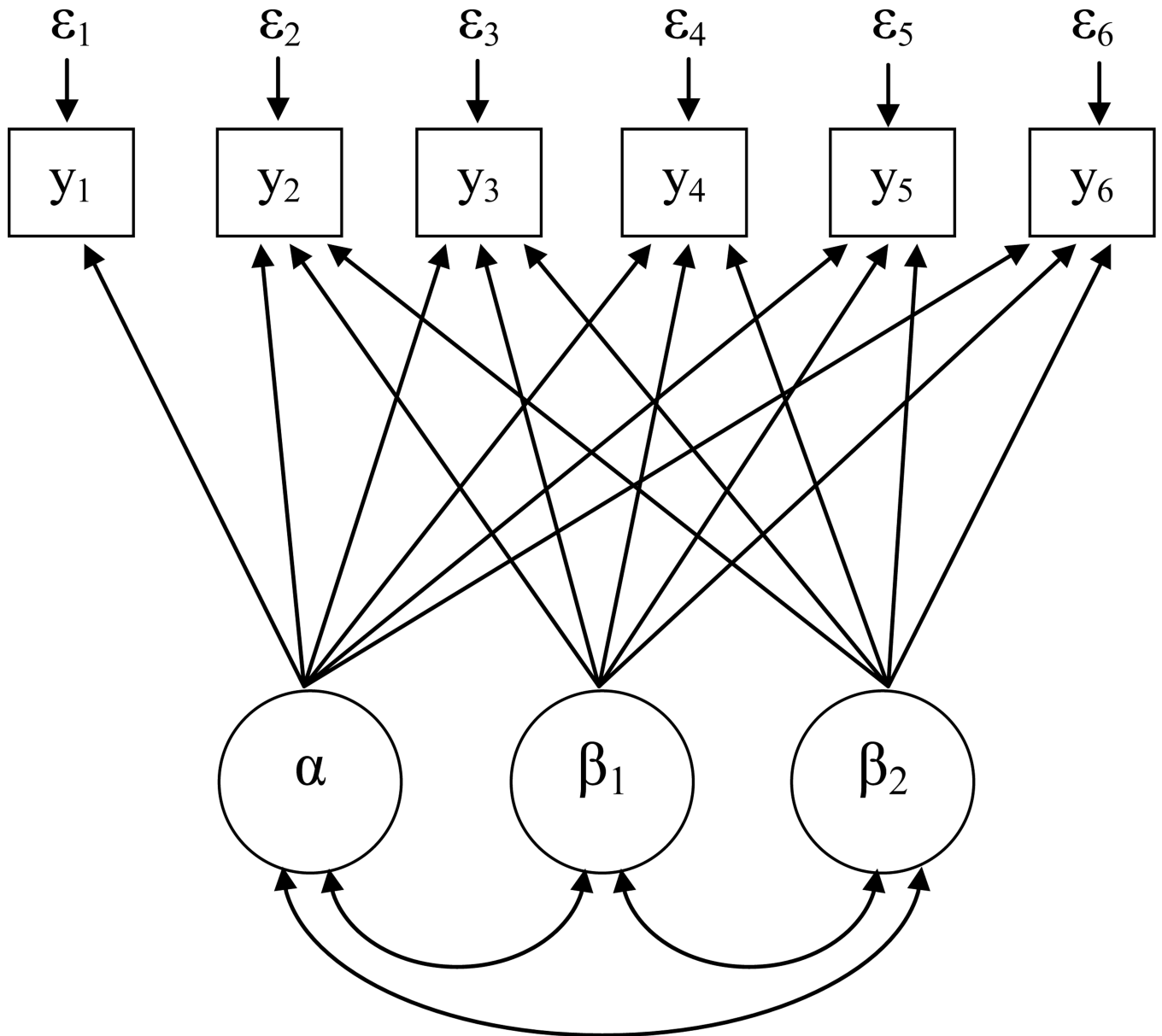
### Highlights

- The TAQS has strong psychometric properties with residential care populations
- Youth ratings of therapeutic alliance with male and female staff are correlated
- Youth and staff ratings of therapeutic alliance are only slightly correlated
- Alliance tended to start high, decline over time, and then slowly rebound
- Developmental trajectories of alliance differed significantly between youth

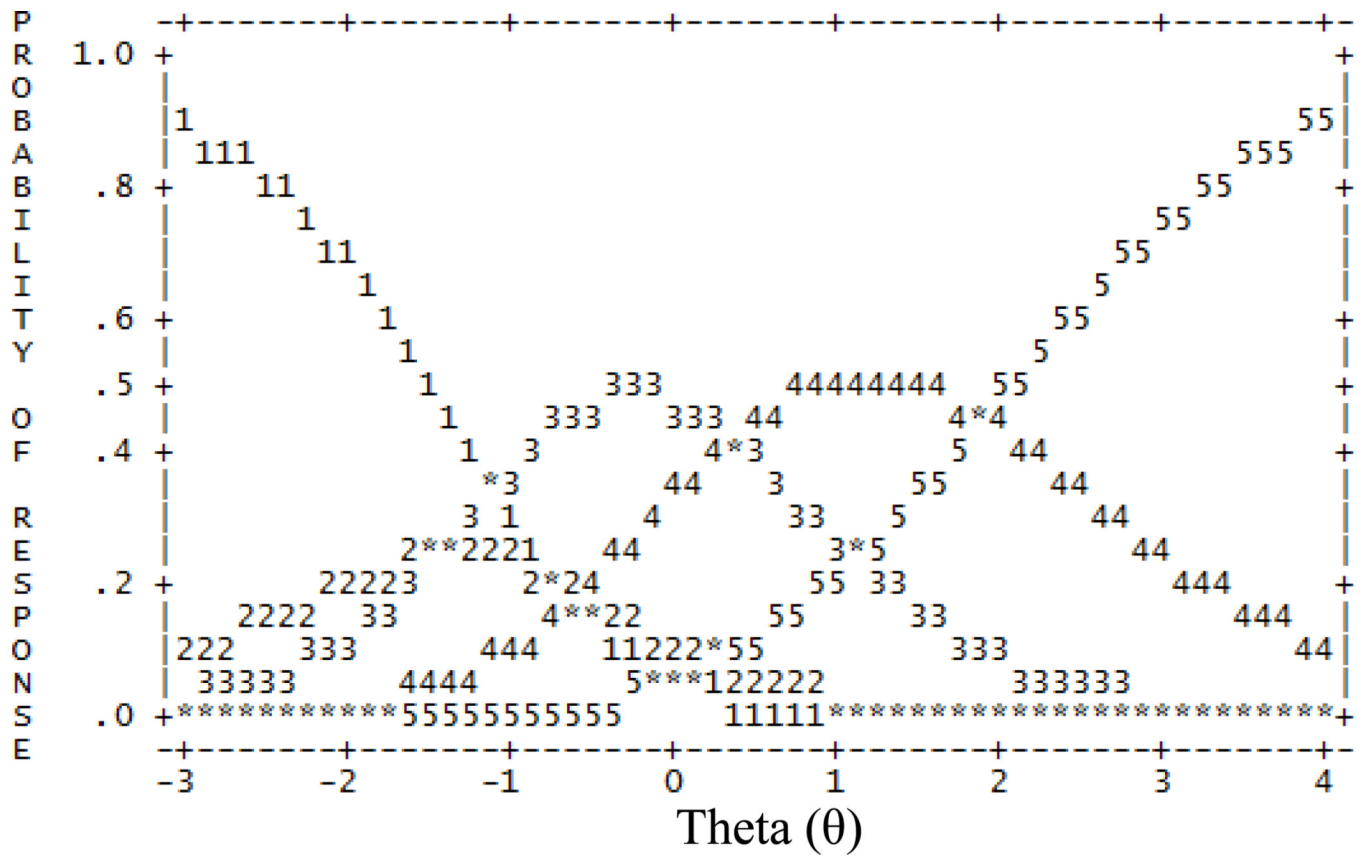
\$watermark-text

\$watermark-text

\$watermark-text



**Figure 1.** Path diagram of the latent growth model for TAQS models  
**Note.**  $\alpha$  = intercept growth factor;  $\beta_1$  = slope growth factor;  $\beta_2$  = quadratic growth factor;  $y_1$ – $y_6$  = therapeutic alliance measure over six time points.  
 The staff-ratings models include only an intercept and slope factor.



**Figure 2.** Male service provider TAQS category characteristic curves plot (generated by Winsteps v3.73)  
**Note.** Category Probabilities: Modes – Structure Measures at Intersections (1 = *Not at all*, 2 = *Only a little*, 3 = *Somewhat*, 4 = *Quite a bit*, 5 = *Totally*)

**Table 1**

TAQS Item Descriptive Statistics

	n	M	SD	Skewness	Kurtosis	Measure	Disc.	Infit	Outfit
<i>TAQS with Male Service Provider</i>									
Happy with treatment	134	4.09	1.03	-0.94	0.23	-0.35	1.26	0.70	0.73
Youth understands treatment	133	4.36	0.88	-1.65	3.20	-1.04	1.08	0.83	0.97
Work together on treatment	132	4.11	0.96	-1.16	1.26	-0.38	1.21	0.71	0.72
Service provider would help	134	3.99	1.13	-0.91	0.01	-0.13	0.67	1.26	1.36
Service provider understands youth	134	2.85	1.53	0.05	-1.47	1.90	0.84	1.28	1.18
<i>TAQS with Female Service Provider</i>									
Happy with treatment	133	4.08	1.05	-0.96	0.17	-0.50	1.05	0.90	0.91
Youth understands treatment	132	4.28	0.93	-1.46	2.23	-0.98	0.92	0.90	1.08
Work together on treatment	131	4.01	1.08	-1.21	1.16	-0.30	1.38	0.70	0.65
Service provider would help	133	3.91	1.21	-0.97	0.08	-0.13	0.66	1.33	1.36
Service provider understands youth	130	2.72	1.48	0.24	-1.32	1.91	1.05	1.02	0.92

**Table 2**

## TAQS Scale Descriptive Statistics and Internal Consistency

	<b>Male Service provider (n =135)</b>	<b>Female Service provider (n = 134)</b>
<i>M</i>	3.87	3.80
SD	0.85	0.89
Skewness	-0.72	-0.77
Kurtosis	0.23	0.44
Cronbach's alpha	0.82	0.83

**Table 3**

## Goodness-of-Fit Indicators

	Male Treatment provider (n = 135)	Female Service provider (n = 134)
$\chi^2(5)$	6.76	8.01
CFI	0.99	0.98
TLI	0.99	0.98
RMSEA	.051 [0 – .138]	.067 [0 – .149]
SRMR	0.02	0.02

**Note.** Both chi-square statistics were non-significant.

CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual.

**Table 4**

Unstandardized Loadings (Standard Error) and Standardized Loadings

	Male Service provider (n = 135)		Female Service provider (n = 134)	
	Un-standardized	Standardized	Un-standardized	Standardized
Happy with treatment	1.00 (-)	.82	1.00 (-)	.74
Youth understands treatment	0.76 (.088)	.73	0.84 (.109)	.71
Work together on treatment	0.87 (.100)	.77	1.16 (.141)	.84
Service provider would help	0.85 (.116)	.64	0.98 (.144)	.63
Service provider understands youth	1.09 (.156)	.60	1.26 (.178)	.66



**Table 5**

Model Fit and Unstandardized Latent Growth Factor Means (Variances)

	Male TAQS (n = 101)	Female TAQS (n = 99)	Male Service Provider (n = 100)	Female Service Provider (n = 101)
<i>LRT</i>	7.4	28.75 *	34.92 ***	18.70
$\alpha$	3.88 (0.42 ***)	3.79 (0.50 ***)	3.95 (0.25 ***)	3.96 (0.18 **)
$\beta_1$	-0.85 ** (5.22 *)	-0.40 (8.64 *)	-0.05 (0.29 *)	0.14 (0.17)
$\beta_2$	0.84 (8.36 *)	0.44 (13.40 *)	-	-

\*  
*p* .05

\*\*  
*p* .01

\*\*\*  
*p* .001

**Note.**  $LRT = -2 * [\log \text{likelihood } H_1 - \log \text{likelihood } H_0]$

All intercept means were significant at the .001 alpha level.