10-1-2017

Assessing Clinical Improvement in School-Based Treatment for Social Anxiety Disorder: Agreement Between Adolescents, Parents, and Independent Evaluators

Jeremy K Fox
Montclair State University, foxjer@mail.montclair.edu

Carrie Masia
Montclair State University, masiac@mail.montclair.edu

Child Psychiatry Hum Dev (2017) 48:721–727
DOI 10.1007/s10578-016-0697-5

ORIGINAL ARTICLE

Assessing Clinical Improvement in School-Based Treatment for Social Anxiety Disorder: Agreement Between Adolescents, Parents, and Independent Evaluators

Jeremy K. Fox[1] · **Carrie Masia Warner**[2,3,4]

Published online: 22 November 2016

Abstract The Clinical Global Impressions of Improvement (CGI-I) scale is widely used in clinical trials to monitor clinically meaningful change during treatment. Although it is standard practice in research to have independent evaluators (IEs) complete the CGI-I, this approach is not practical in school and community settings. Few studies have explored the potential utility of other informants, such as youth and parents. Therefore, this study aimed to investigate agreement between IEs and both adolescents and parents in CGI-I improvement ratings in the context of a randomized controlled trial of cognitive-behavioral therapy for social anxiety disorder, as delivered by psychologists and school counselors. Multilevel growth models indicated that IEs were generally more conservative in their ratings of positive treatment response across time and treatment conditions, though greater agreement was observed between parents and IEs by post-intervention and 5-month follow-up. Possible explanations for these findings and suggestions for alternative approaches are discussed.

Keywords Social anxiety · Cognitive-behavioral therapy · Schools · Independent evaluator · Rater agreement · Treatment response

✉ Carrie Masia Warner
masiac@wpunj.edu

[1] Department of Psychology, Montclair State University, Montclair, NJ, USA

[2] Department of Psychology, William Paterson University, Wayne 07470, NJ, USA

[3] Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NJ, USA

[4] Department of Child and Adolescent Psychiatry, NYU Langone Medical Center, New York, NJ, USA

Introduction

Quick and efficient in administration, global ratings of clinical improvement are designed to assess changes in an individual's overall well-being following treatment and are commonly used in randomized controlled trials (RCTs) to assess treatment response. One measure, the Clinical Global Impressions of Improvement scale (CGI-I) [1, 2], is widely considered the gold standard assessment of clinical improvement. As an assessment tool with numerical values denoting treatment response (i.e. clinically significant treatment change), the CGI-I has been utilized extensively as a primary outcome measure in RCTs of interventions for mood, anxiety, and eating disorders [3], including several prominent studies in youth mental health [4–6].

RCTs for youth mental health treatments typically rely on independent evaluators (IEs) to provide CGI-I ratings by integrating various sources of information (e.g., child and parent interviews, clinical judgment). Using IEs, who are trained in evidence-based assessment and blinded to treatment condition, may yield more objective ratings, compared to youth, parents, and clinicians, whose judgments may rely less on strict behavioral markers and instead be influenced by subjective factors, such as social comparisons (e.g., to peers), social desirability effects, demand characteristics, and personal motivations for improvement [7, 8]. However, this practice can be costly and impractical, particularly given increasing emphasis on longitudinal methods requiring frequent assessments to map patient change trajectories. Moreover, it is important to consider assessment options that are feasible and cost-effective for real-world community settings, in light of growing awareness of the need to enhance children's access to evidence-based services. For example, schools have been increasingly promoted as a prime setting in which to address the high rates

of youth with untreated anxiety disorders [9, 10]. Given their competing priorities and limited time and resources, it is unrealistic to expect schools to depend on outside evaluators to determine the usefulness of school-based treatments. Therefore, it may be beneficial to utilize youth and parents as reporters of improvement.

To our knowledge, only one published study has explored youth and parents as reporters of global improvement during treatment using the CGI-I. Lewin et al. [11] evaluated agreement between youth (ages 8–17), parents, and IEs in ratings of treatment response on the CGI-I, in the context of an efficacy trial for pediatric obsessive–compulsive disorder comparing family-based cognitive-behavioral therapy (CBT) to a control condition emphasizing psychoeducation and relaxation training. Findings indicated that IEs were initially more conservative than youth and parents in rating treatment response. However, agreement between IEs and both youth and parents increased over the course of treatment, with the authors suggesting that youth and parents "caught up" to IEs over time. Additionally, no differences in cross-rater agreement were observed between treatment conditions.

Although Lewin et al. provide some support for the potential utility of youth and parents as raters of treatment response, additional research is needed, particularly in light of recommendations from NIH for increased study of patient-reported outcomes [12]. Furthermore, consistent with the growing movement toward transporting evidence-based treatments for childhood disorders to real-world community settings, such as schools, it is important to study agreement between youth, parents, and IEs as raters of treatment response in the context of an effectiveness trial. In addition, as noted by Lewin et al. [11], it is uncertain whether their findings will generalize to other populations, such as children of different ages and with other psychological disorders (e.g., social anxiety). For example, given that cross-informant agreement in ratings of internalizing problems decreases as children age [13], greater disagreement in treatment response may be expected in adolescents.

The primary goal of the present study was to evaluate agreement between IEs and both adolescents and parents in ratings of CGI-I clinical improvement during an effectiveness trial of school-based CBT for social anxiety disorder (SAD), as delivered by psychologists and school counselors. In line with Lewin et al. [11], we examined associations between IE, adolescent, and parent CGI-I ratings across time (midpoint, post-intervention, and 5-month follow-up) and randomly assigned treatment conditions (CBT vs. attention control). Specifically, we examined (a) whether treatment response rates changed over time, (b) whether type of rater predicted treatment response rates and changes in response rates over time (i.e. do adolescents and parents agree with IEs, and does agreement change over time?), and (c) whether rater differences in treatment response varied by treatment condition. We expected, based on Lewin et al., that adolescents and parents would endorse higher treatment response rates than IEs across treatment conditions, though agreement would increase over time. However, given the aforementioned findings from van der Ende et al. [13], we anticipated that there would be greater disagreement between adolescents and IEs than between parents and IEs.

Methods

Participants and Procedures

Data for the present study was collected as part of a controlled trial of *Skills for Academic and Social Success* (SASS), a school-based group CBT intervention for adolescent SAD [14]. The trial was approved by the Institutional Board of Research Associates at the New York University School of Medicine. It was funded by the National Institute of Mental Health (R01MH081881) and registered on the U.S. National Institutes of Health Clinical Trials database (ref: NCT01320800). Recruitment involved a sequence of three steps: initial school-wide screenings, telephone screenings with parents, and diagnostic evaluations with adolescents and parents using the *Anxiety Disorders Interview Schedule for DSM-IV: Parent and Child Versions* (ADIS-P/C) [15]. Adolescents were eligible for the trial if assigned a primary diagnosis of SAD. For further description of trial procedures, see Masia Warner et al. [14].

Based on these recruitment procedures, 138 ninth through 11th graders (ages $M=15.42$; $SD=0.81$; 68% female, 72% white, median family income between $100,000–$150,000) were randomized to one of three 12-week treatment groups: SASS delivered by school counselors (C-SASS), SASS delivered by psychologists (P-SASS), and *Skills for Life* (SFL), a non-specific group counseling program included as an attention control condition (see Ryan and Masia Warner [16] for additional description of SASS). Because no differences in primary outcomes (e.g., CGI-I scores) were observed between C-SASS and P-SASS (see Masia Warner et al. [14] for primary outcome results), the two groups were combined (SASS; $N=95$) and compared to SFL ($N=43$) in subsequent analyses.

Raters (adolescents, parents, and IEs) completed the *Clinical Global Impression Scale-Improvement-Revised Version* (CGI-I) [2] during scheduled diagnostic evaluations at three time points: midpoint (week six of treatment), post-intervention (post; following week 12), and follow-up (five months after treatment). The CGI-I has demonstrated good reliability and validity in prior studies [3, 17], and

the revised CGI-I has been used in multiple clinical trials, including the Pediatric Psychopharmacology (RUPP) Anxiety Group [18]. Raters were asked to indicate how much their (or the adolescent's) social anxiety (i.e. social evaluative fears/worries and functioning in social situations) changed since the start of treatment on a scale from 1 (*Completely Recovered*) to 8 (*Much Worse*). Ratings of 1, 2 (*Much Improved*), and 3 (*Improved*) denote treatment response, whereas ratings of 4 (*Minimally Improved*) and higher denote treatment non-response. Adolescents and parents completed the CGI-I separately, and CGI-I ratings were not shared between raters. IEs were doctoral-level psychologists or psychology graduate students who completed extensive training in evidence-based assessment of SAD and other anxiety disorders, as well as in the CGI-I. IEs conducted the diagnostic evaluations (ADIS-P/C) to assess current functioning and improvement at midpoint, post, and follow-up but were blinded to treatment condition and did not have access to any treatment records or the adolescent and parent CGI-I forms. For most participants, the same IE was involved at all three time points; in the event that a different IE had to complete an evaluation at a particular time point, the replacement IE reviewed that participant's previous evaluations and listened to an audio recording of the pre-treatment evaluation in order to have knowledge of the participant's baseline functioning.

Data Analysis

To evaluate differences among raters in treatment response rates on the CGI-I over time, we performed multilevel growth modeling in Mplus 7.4 [19], using maximum likelihood estimation with robust standard errors. Analyses followed an intent-to-treat approach, including all randomized subjects to create treatment response probability trajectories. In line with Mplus procedures described by Heck and Thomas [20], as well as the model building framework used by Lewin et al. [11], we evaluated a series of five models in sequential fashion. This approach allows for examining change in model fit as predictors of treatment response (e.g., time, rater type, treatment condition) and interactions between predictors are added to each successive model. Consistent with Lewin et al. and our aims, we entered treatment response as a binary variable, with CGI-I ≤ 3 representing responder status and CGI-I ≥ 4 representing non-responder status. All variables had <5% missing values. Multiple imputation was conducted to handle missing data, with subsequent analyses conducted across 15 complete datasets. Results represent an average of the 15 separate analyses with Rubin's correction of standard error [21].

First, we tested an unconditional means (null) model to evaluate the mean response rate. Initial testing of a null model without predictors provides baseline statistics for comparison to subsequent models as predictors are added. Second, an unconditional growth model, with response scores for each time point nested within participants, examined fixed and random effects of time to determine, respectively, whether the probability of treatment response changed over time and whether significant variability in growth over time could be explained by other variables (e.g., rater type). As per recommendations by Heck and Thomas [20], our time metric (6, 12, 32) reflected the length of time (in weeks) following the start of treatment. Third, we evaluated a conditional growth model in which type of rater was examined as a within-subjects predictor. For ease of interpretation, we added two dichotomous rater variables created through dummy coding, with *adolescent* comparing adolescent and IE response ratings and *parent* comparing parents and IEs. Fourth, we examined a time × rater interaction to determine if differences between raters varied by time. Finally, we added treatment condition (SASS vs. SFL) as a between-subjects predictor, as well as examined three cross-level interactions: time × treatment, *adolescent* × reatment, and *parent* × treatment. Each cross-level interaction was examined by regressing the random slope of the relationship between CGI-I response and the within-subjects predictor (e.g., *adolescent*) on treatment condition. This allowed us to test whether differences in CGI-I response across time and rater varied by treatment condition. If significant interactions were observed, we examined models separately by treatment condition.

Chi square difference tests were performed utilizing log-likelihood values to compare the fit of each model (comparison) with the previous model (nested) in the sequence. A significant χ^2 demonstrates improvement of model fit and value of adding the predictors, when compared to the previous model. Pseudo R^2 statistics were also obtained to estimate improvement of fit from the null model to all subsequent models. For further model comparison, we examined Akaike's information criterion (AIC), Bayesian information criterion (BIC), and adjusted BIC, with lower values indicating better fit. Finally, to evaluate rater agreement among continuous CGI-I scores, we performed bivariate correlations between scores across raters for each time point.

Results

Analysis of the unconditional means model revealed significant effects for the fixed (estimate = −0.82, SE = 0.12, z = −6.66, $p < 0.001$) and random intercept (estimate = 1.12, SE = 0.26, z = 4.25, $p < 0.001$), indicating that mean CGI-I response differed from zero and that variance in response rates could be explained by other predictors (e.g., time,

rater type; AIC = 1509.05, BIC = 1519.29, adjusted BIC = 1512.94).

Does Probability of Treatment Response Change Over Time?

In the unconditional growth model, the fixed effect of time was significant (estimate = 0.06, SE = 0.01, z = 5.82, $p < 0.001$), however the random effect of time was non-significant (estimate = 0.00, SE = 0.00, z = 0.77). Therefore, while the probability of CGI-I responder status increased over time, little variability was present in the increased probability of responder status over time, suggesting that interactions between time and other predictors (e.g., rater type) may not be observed in subsequent models. Nonetheless, the unconditional growth model was an improvement upon the null model ($\chi^2(3) = 110.51$, $p < 0.05$, McFadden's pseudo-$R^2 = 0.04$, AIC = 1450.69, BIC = 1476.32, adjusted BIC = 1460.43).

Does Type of Rater Predict Probability of Treatment Response Over Time?

As expected, the conditional growth model, with type of rater added as a predictor, yielded further improvement in model fit ($\chi^2(2) = 57.57$, $p < 0.05$, McFadden's pseudo-$R^2 = 0.10$, AIC = 1362.38, BIC = 1398.25, adjusted BIC = 1376.01). Both dummy-coded rater variables, *adolescent* (estimate = 1.70, SE = 0.24, z = 7.11, $p < 0.001$) and *parent* (estimate = 0.88, SE = 0.22, z = 3.99, $p < 0.001$), were significant, though entering interactions between time and both rater variables did not improve model fit ($\chi^2(2) = 0.78$, McFadden's pseudo-$R^2 = 0.10$, AIC = 1365.64, BIC = 1411.76, adjusted BIC = 1383.17). Therefore, IE response rates differed from ratings by adolescents and parents, and level of agreement did not appear to improve over time.

Do Differences Between Raters Vary by Treatment Condition?

The addition of treatment condition and related interactions (time × treatment, rater × treatment) also resulted in improved fit over the prior models ($\chi^2(8) = 41.98$, $p < 0.05$, McFadden's pseudo-$R^2 = 0.13$, AIC = 1336.62, BIC = 1413.49, adjusted BIC = 1365.84). Treatment condition was a significant between-subjects predictor (estimate = 1.41, SE = 0.33, z = 4.23, $p < 0.001$), such that SASS was generally associated with higher probability of CGI-I responder status, when compared to the SFL condition. The relationship between treatment condition and responder status did not vary by time point. The *adolescent* × treatment interaction was also non-significant (estimate = −0.69, SE = 0.55, z = −1.27), however the *parent* × treatment interaction was significant (estimate = −1.11, SE = 0.54, z = −2.07, $p < 0.05$).

To explore this interaction, we tested the fourth model (with time × rater) separately by treatment condition. When testing this model in SFL participants, *adolescent* remained significant (estimate = 1.99, SE = 0.69, z = 2.89, $p < 0.01$), however *parent* was non-significant (estimate = 1.01, SE = 0.67, z = 1.51). Interactions between time and both rater variables remained non-significant. Among SASS participants, *adolescent* (estimate = 1.84, SE = 0.42, z = 4.40, $p < 0.001$) and *parent* (estimate = 1.84, SE = 0.42, z = 2.73, $p < 0.01$) remained significant, and the interaction between time and *parent* was non-significant but suggestive of an effect (estimate = −0.04, SE = 0.02, z = −1.65, $p < 0.10$). Post-hoc χ^2 tests within SASS participants indicated that parents reported higher response rates than IEs at midpoint ($\chi^2(1) = 6.23$, $p = 0.01$) but not post ($\chi^2(1) = 1.77$, $p = 0.18$) or follow-up ($\chi^2(1) = 0.23$, $p = 0.63$).

Across treatment conditions and time, adolescents reported consistently higher treatment response rates than IEs (Fig. 1). Parents also reported higher response rates than IEs, but this was only observed in the SFL condition. More specifically, parents whose children participated in SFL were more likely than IEs to endorse positive treatment response. Parents whose children participated in SASS more closely agreed with IEs regarding responder status, with stronger agreement observed by post and follow-up. Parents reported similar response rates across SASS and SFL conditions, whereas IEs (and, to some extent, adolescents) reported higher response rates for SASS participants. Correlations across raters over time can be found in Table 1.

Discussion

In line with our expectations, findings from the present study indicate that IEs were more conservative, compared to adolescents and parents, in reporting positive treatment response during a controlled trial comparing SASS, a school-based CBT program for social anxiety disorder, to SFL, a nonspecific school counseling program. While IEs were more conservative than adolescents across conditions, agreement between parents and IEs was stronger for SASS than SFL, with IE response rates for SASS participants increasing over time and converging with parent response rates by post-intervention and 5-month follow-up. This may indicate that IEs were initially slower to endorse responder status for SASS participants until they observed a more enduring effect. Alternatively, given that blinded and trained IEs are often considered the best available option for promoting objective evaluation in RCTs, and in

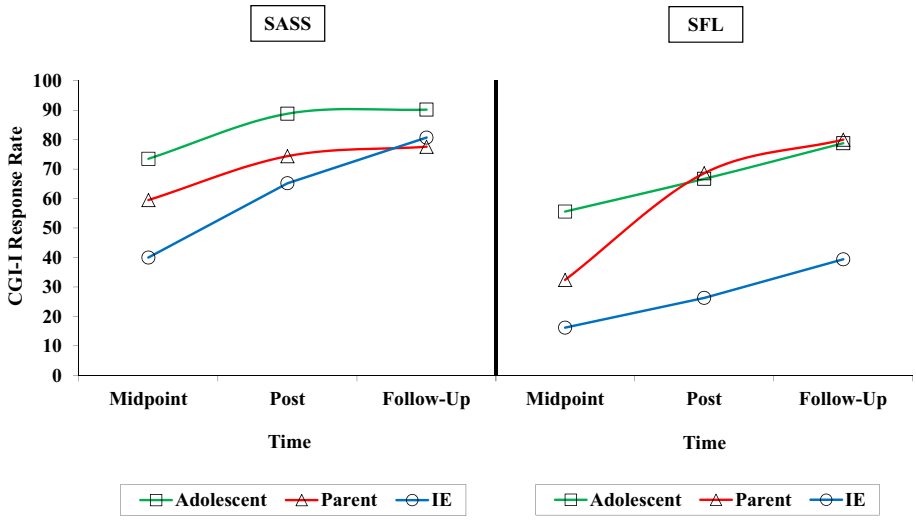

Fig. 1 CGI-I response rate (percentage of cases rated as "Improved," "Much Improved," or "Very Much Improved") reported by three raters (adolescent, parent, and independent evaluator) over three time points (midpoint of intervention, post-intervention, and 5-month follow-up). Response rates for SASS and SFL participants are displayed on the *left* and *right*, respectively

Table 1 Correlations between continuous CGI-I improvement scores by rater type and time

	SASS		SFL	
	Adolescent-IE	Parent-IE	Adolescent-IE	Parent-IE
Midpoint	0.32**	0.30**	0.46**	0.43*
Post-intervention	0.16	0.53***	0.32^{+}	0.57***
Follow-up (5 month)	0.24*	0.36**	0.46**	0.45*

IE independent evaluator

$^{+}p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

line with suggestions from Lewin et al. [11], this pattern of disagreement may underscore concerns about obtaining unbiased and accurate ratings from adolescents and parents on the CGI-I. For example, it could suggest that youth may over-report the benefits of treatment, though it is also possible that IEs did not consider some way in which youth perceived meaningful change. In addition, while parents may show greater agreement with IEs over the course of CBT, they (and, to a lesser extent, adolescents) exhibited relatively similar response rates between conditions, whereas IEs provided higher ratings for SASS compared to SFL participants. Therefore, parent improvement ratings may not differ between CBT and non-CBT conditions.

To better understand the variability between adolescents, parents, and IEs in rating treatment response, it may help to consider the role of potential differences in their perspectives. Although the CGI-I is a "global" rating, different informants may base ratings on different aspects of functioning [22, 23]. For example, IEs were trained to integrate information provided by the adolescent and parent and to consider behavioral changes in social anxiety symptoms and impairments (e.g., avoidance of social situations) assessed through clinical interviews. However, adolescents and parents may have been influenced by additional experiences, such as family functioning (e.g., family members getting along better) or school experiences (e.g., feeling more comfortable walking the school hallways). In addition, different informants may utilize different comparison groups when providing their ratings. Whereas adolescents may view themselves in relation to their peers, IEs may utilize their experiences assessing anxious teenagers and knowledge of behavioral indicators of mild, moderate, and severe anxiety to form their judgments. Therefore, it might be helpful to add behavioral descriptions to the CGI-I that fully operationalize anchors with a wider spectrum of meaningful improvements from various perspectives.

Moreover, adolescents and parents may be more vulnerable to biases (e.g., social desirability, halo, and placebo effects) that contributed to positive improvement ratings regardless of treatment condition. For instance, adolescents and parents may be influenced by both demand characteristics (e.g., believing that the researchers expected improvement) and personal motivations for treatment response due to the considerable time and investment (e.g., missed class time) spent in treatment. In addition, it has been shown that adolescents who exhibit higher social desirability are more likely to underreport levels of social avoidance [24]. As noted by Grills and Ollendick [7], adolescents may withhold reporting distress in clinical interviews due to stigma concerns and expectations that adults will respond negatively. Parents may share similar concerns about the effects of stigma on their child and may also not be fully informed

about their child's distress [25]. This may be especially problematic in rating improvement of adolescent SAD, as evidence points to greater disagreement in rating the severity of internalizing compared to externalizing disorders (see review by Grills and Ollendick [7]), as well as in adolescents compared to younger children [13]. If adolescents with SAD are indeed reluctant to acknowledge their distress due to the internal nature of SAD (e.g., worries), their desire for privacy, and their fears of negative evaluation, future studies should evaluate the level of social desirability in youth reports and investigate interview methods and other procedures (e.g., implicit measures, diary tracking of social behavior) that can facilitate more open and honest responses.

It is important to acknowledge a few limitations of the present study. First, additional points of CGI-I assessment during treatment would have increased our ability to investigate more subtle changes in rater agreement between sessions. Similarly, long-term follow-up assessment is needed to examine whether parent-IE agreement can be maintained well after the end of treatment. Second, as the present sample was predominantly comprised of white adolescents with SAD from families with moderate to high incomes, further study should investigate whether our findings generalize to racial/ethnic minority youth, families with lower socioeconomic status, younger children, and youth with other psychological disorders. For example, it may be expected that cross-informant agreement would be higher in both the present study and Lewin et al. [11], which examined improvement in anxiety-related disorders characterized by overt behavioral symptoms (SAD and OCD), compared to investigations of anxiety disorders featuring less observable symptoms (e.g., GAD). Third, we did not obtain CGI-I ratings from other potentially useful informants of adolescent social anxiety and clinical improvement, such as school counselors or teachers. As highlighted by Lewin et al. [11], treating clinicians might serve as objective raters and a more cost-effective option than using IEs in larger scale effectiveness research and community settings. Further study should examine the validity and utility of obtaining CGI-I ratings from school professionals (e.g., counselors, coaches) who have regular opportunities to observe school-based manifestations of social anxiety (e.g., talking with peers, class presentations) among youth enrolled in treatment.

Summary

Our findings suggest that adolescents and parents may disagree with IEs in rating clinical improvement of social anxiety disorder. Despite some promising findings showing agreement between parents and IEs over the course of our SASS treatment, ratings from adolescents and parents did not differ between CBT and a control intervention. These results are concerning because using trained and blinded IEs is likely not feasible and sustainable in the long run as we shift toward disseminating youth mental health interventions to community and school settings. Therefore, future studies should aim to design and evaluate strategies for improving cross-informant agreement on the CGI-I and other global measures of functioning. In addition to refining the CGI-I anchors and operationalizing behavioral descriptors of change, agreement could be increased by obtaining multiple ratings of improvement in different domains of social anxiety (e.g., by setting, such as school, home, etc.), rather than a global rating. An existing measure, the Social Anxiety Session Change Index [26], includes four questions regarding key facets of social anxiety (e.g., social avoidance, embarrassment concerns) and could be a useful alternative. Additionally, further study is needed to evaluate the feasibility of these approaches in schools, including whether they can be easily utilized by children, parents, and school personnel. If shown to be accurate and efficient, these approaches to assessing treatment progress may help therapists in community settings to develop and promote sustainable evidence-based treatments for youth.

Acknowledgements This research was supported by a federal grant awarded to Dr. Carrie Masia Warner from the National Institute of Mental Health, R01, MH081881.

References

1. Guy W (1976) ECDEU Assessment Manual for Psychopharmacology. US Department of Health, Education and Welfare Publication (ADM) 76–338. National Institute of Mental Health, Rockville
2. Klein RG, Koplewicz HS, Kanner A (1992) Imipramine treatment of children with separation anxiety disorder. J Am Acad Child Adolesc Psychiatry 31:21–28
3. Hedges DW, Brown BL, Shwalb DA (2009) A direct comparison of effect sizes from the clinical global impression-improvement scale to effect sizes from other rating scales in controlled trials of adult social anxiety disorder. Hum Psychopharmacol 24:35–40
4. Pediatric OCD Treatment Study Team (2004) Cognitive-behavior therapy, sertraline, and their combination for children and adolescents with obsessive–compulsive disorder. JAMA 292:1969–1976
5. The TADS Team (2007) The Treatment of Adolescent Depression Study (TADS). Arch Gen Psychiatry 64:1132–1144
6. Walkup JT, Albano AM, Piacentini J, Birmaher B, Compton SN, Sherrill JT et al (2008) Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. N Engl J Med 359:2753–2766
7. Grills AE, Ollendick TH (2002) Issues in parent–child agreement: the case of structured diagnostic interviews. Clin Child Fam Psychol Rev 5:57–83
8. Spinelli MG, Endicott J, Goetz RR (2015) Disagreement between therapist raters and independent evaluators in a controlled trial

of interpersonal psychotherapy for depressed pregnant women. J Psychiatr Pract 21:114–123
9. Fox JK, Herzig-Anderson K, Colognori D, Stewart, CE, Masia Warner C (2013) School-based treatment of anxiety in children and adolescents: new developments in transportability and dissemination. In: Weist MD, Lever NA, Bradshaw CP, Owens, JS (eds) Handbook of school mental health. Springer, New York
10. Masia Warner C, Klein RG, Dent HC, Fisher PH, Alvir J, Albano AM et al (2005) School-based intervention for adolescents with social anxiety disorder: results of a controlled study. J Abnorm Child Psychol 33:707–722
11. Lewin AB, Peris TS, De Nadai AS, McCracken JT, Piacentini J (2012) Agreement between therapists, parents, patients, and independent evaluators on clinical improvement in pediatric obsessive compulsive disorder. J Consult Clin Psychol 80:1103–1107
12. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B et al (2007) The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 45:S3–S11
13. van der Ende J, Verhulst FC, Tiemeier H (2012) Agreement of informants on emotional and behavioral problems from childhood to adulthood. Psychol Assessment 24:293–300
14. Masia Warner C, Colognori D, Brice C, Herzig K, Mufson L, Lynch C et al (2016) Can school counselors deliver cognitive-behavioral treatment for social anxiety effectively? A randomized controlled trial. J Child Psychol Psychiatry. doi:10.1111/jcpp.12550
15. Silverman WK, Albano AM (1996) The anxiety disorders interview schedule for DSM-IV: child and parent versions. Psychological Corporation, San Antonio
16. Ryan JL, Masia Warner C (2012) Treating adolescents with social anxiety disorders in schools. Child Adolesc Psychiatr Clin 21:105–118
17. Berk M, Ng F, Dodd S, Callaly T, Campbell S, Bernardo M, Trauer T (2008) The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. J Eval Clin Pract 14:979–983
18. Rupp Anxiety Group (2001) Fluvoxamine for the treatment of anxiety disorders in children. N Engl J Med 344:1279–1285
19. Muthén LK, Muthén BO (1998–2012) Mplus user's guide, 7th edn. Muthén & Muthén, Los Angeles
20. Heck RH, Thomas SL (2015) An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus, 3rd edn. Routledge, New York
21. Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley, New York
22. Beneke M, Rasmus W (1992) 'Clinical Global Impressions' (ECDEU): some critical comments. Pharmacopsychiatry 25:171–176
23. Zaider TI, Heimberg RG, Fresco DM, Schneier FR, Liebowitz MR (2003) Evaluation of the Clinical Global Impression Scale among individuals with social anxiety disorder. Psychol Med 33:611–622
24. DiBartolo PM, Albano AM, Barlow DH, Heimberg RG (1998) Cross-informant agreement in the assessment of social phobia in youth. J Abnorm Child Psychol 26:213–220
25. Bidaut-Russell M, Reich W, Cottler LB, Robins LN, Compton WM, Mattison RE (1995) The diagnostic interview schedule for children (PC-DISC v.3.0): parents and adolescents suggest reasons for expecting discrepant answers. J Abnorm Child Psychol 23:641–659
26. Hayes SA, Miller NA, Hope, DA, Heimberg RG, Juster HR (2008) Assessing client progress session by session in the treatment of social anxiety disorder: the Social Anxiety Session Change Index. Cogn Behav Pract 15:203–211