Title

Development and preliminary validation of the treatment adherence rating scale.

Authors

Gumport, Nicole B
Dong, Lu
Lee, Jason Y
et al.

Publication Date

2023-06-01

DOI

10.1016/j.jbtep.2022.101832

Peer reviewed

HHS Public Access

Author manuscript

J Behav Ther Exp Psychiatry. Author manuscript; available in PMC 2023 June 01.

Published in final edited form as:

J Behav Ther Exp Psychiatry. 2023 June ; 79: 101832. doi:10.1016/j.jbtep.2022.101832.

Development and Preliminary Validation of the Treatment Adherence Rating Scale

Nicole B. Gumport[a,*], **Lu Dong**[b,*], **Jason Y. Lee**[c], **Xin Zhao**[d], **Allison G. Harvey**[e]

[a]Stanford University School of Medicine, 401 Quarry Road, Stanford, CA, 94305.

[b]RAND Corporation, 1776 Main Street, Santa Monica, CA, 90407.

[c]University of California, Berkeley, 2121 Berkeley Way, Berkeley, CA 94720.

[d]University of California, Berkeley, 2121 Berkeley Way, Berkeley, CA 94720.

[e]University of California, Berkeley, Department of Psychology, 2121 Berkeley Way, Berkeley, CA 94720

Abstract

Background and Objectives: Patient adherence to treatment is an important barrier to the implementation of evidence-based psychological treatments (EBPTs). There is a need for simple and deployable measures of patient adherence to treatment. The Treatment Adherence Rating Scale (TARS) was developed and validated in two samples.

Methods: This study includes two samples: adults with Major Depressive Disorder who received Cognitive Therapy for depression (Sample 1; *N*=48, mean age=44.27 years), and at-risk adolescents who received either the Transdiagnostic Sleep and Circadian Intervention or Psychoeducation (Sample 2; *N*=176, mean age=14.77 years). Factor structure of the TARS scores was examined via Exploratory Factor Analyses (EFA) in Sample 1 and Confirmatory Factor Analyses (CFA) in Sample 2. Internal consistency, predictive validity, and construct validity of the TARS scores were examined.

Results: Results from EFA in Sample 1 supported a one-factor model. Results from CFA in Sample 2 suggested that a two-factor model (i.e., agreement and compliance) fit better than a one-factor model. TARS scores from both samples demonstrated adequate predictive validity with primary clinical outcomes and construct validity with treatment expectations.

Limitations: The sample was small with two specific populations. Future research should focus on other patient populations, a larger population, and other treatments. Future research examining patient ratings of these items are needed for further validation of the TARS.

Conclusions: Preliminary findings support the use of a two-factor model and highlight the potential utility of a simple measure of patient adherence to treatment across age and diagnostic groups.

Corresponding Author. Allison G. Harvey. University of California, Berkeley, Department of Psychology, 2121 Berkeley Way, Berkeley, CA 94720, aharvey@berkeley.edu.

*Co-first authors. NBG and LD contributed equally to this work

Keywords

treatment adherence; cognitive therapy; depression; factor analysis; scale development; sleep

Patient adherence to treatment recommendations has been identified as an important barrier to the implementation of evidence-based psychological treatments (EBPTs) (Harvey & Gumport, 2015; Kazdin, 2017). The World Health Organization (WHO) defined adherence as "the extent to which a person's behavior—taking medication, following a diet, and/or executing lifestyle changes—corresponds with agreed recommendations from a health care provider" (World Health Organization, 2003). The definition entails two components; namely, agreement with and compliance with the treatment recommendations. This interpretation is consistent with core components identified in various social behavioral models of medication adherence: 1) knowledge of treatment need, contents, and recommendations (agreement/understanding), and 2) motivation and skills to act on adherence behaviors (compliance) (Amico et al., 2018). This definition of patient adherence has been used in other empirical investigations and reviews of psychological treatment (Cho et al., 2021; Leeuwerik et al., 2019; Meis et al., 2022) and is broad as to capture a range of patient problems and interventions. Lack of patient adherence to treatment causes substantial economic burden to healthcare systems across many health problems (Levensky, 2006).

In the psychological treatment literature, homework completion/compliance appears to be the most commonly used indicator of patient adherence to treatment. There is compelling meta-analytic evidence that greater homework compliance during EBPTs is associated with better clinical outcomes (Kazantzis et al., 2006, 2010). Recent work offer supplemental evidence that homework compliance is associated with better outcomes (Dobson, 2021; Kazantzis, 2021). Other measures of patient adherence to treatment may involve treatment-specific evaluations, such as deriving the level of patient adherence to treatment to sleep recommendations using sleep diaries in cognitive behavior therapy for insomnia (Tremblay et al., 2009). However, such methods can be time-consuming and difficult to implement in certain clinical settings because they are lengthy and burdensome to patients and/or providers. Assessing patient understanding of, acceptance of, and agreement with treatment contents is also important. These constructs, rated by therapists, have been identified as essential elements of learning while receiving EBPTs or interacting with a treatment provider, and are predictive of positive clinical outcomes such as symptom reduction and functional improvement (Abramowitz et al., 2002; Jarrett et al., 2011; Lasalvia et al., 2008; Street et al., 2009; Strunk et al., 2010).

Two scales have been developed to measure patients' comprehension and usage of cognitive therapy (CT) skills. The scores of these scales have demonstrated predictive validity of clinical outcomes for depression (Jarrett et al., 2011; Strunk et al., 2014). For example, the Skills of Cognitive Therapy (SoCT) is an eight-item patient- or therapist-reported scale developed to measure "patients' understanding and use of basic CT skills," such as the ability to recognize and change the relations between thoughts, mood, and behaviors (Jarrett et al., 2011). The Competencies of Cognitive Therapy Scale (CCTS) also has patient-reported (30 items) and therapist-rated (9 items) versions and measures the mastery

and skillful use of CT skills. Both measures were associated with changes in depressive symptoms, and other measures of CT skills. However, these measures, while useful in assessing the understanding and usage of CT skills, do not directly address agreement/acceptance of the treatment content. They are also specific to CT and perhaps less generalizable to other psychological treatments.

Scales have been developed and used to measure patients' use of therapy skills. The Homework Rating Scale –Revised Version (HRS-II) (Kazantzis et al., 2005) is a 12-item measure that has both patient-reported and therapist-reported versions. It assesses homework compliance as well as client beliefs about and consequences of homework engagement. This measure has demonstrated that increased engagement with and quality of homework completion has been predictive of outcomes (McDonald & Morgan, 2013; Sachsenweger et al., 2015). However, while the HRS-II offers data on the role of homework in treatment and can be used across disorders and treatments, like the SoCT and CCTS, it does not measure agreement/acceptance of treatment contents. Disorder specific homework adherence measures have also been developed. One example is the Patient EX/RP Adherence Scale (PEAS), which can be used to measure adherence to exposure therapy homework (Simpson et al., 2011). However, these treatment and disorder specific measure are limited to these specific populations.

Taken together, patient adherence to treatment has been recognized as an important factor influencing treatment outcome, and has been identified as a patient-level, modifiable barrier to the implementation of EBPTs (Harvey & Gumport, 2015; Kazdin, 2017). However, there are currently few validated and easily deployable measures of patient adherence to treatment that can be readily used across various types of psychological treatments. Although simple, deployable, and adaptable indices of patient adherence to treatment (i.e., single item rating scales of patient understanding, patient acceptance/agreement, or homework compliance) are predictive of clinical outcomes (e.g., Bissonnette, 2008; Dong, Soehner, Bélanger, Morin, & Harvey, 2018; Lasalvia et al., 2008; Mausbach, Moore, Roesch, Cardenas, & Patterson, 2010), the reliability and validity of the compilation of these item scores have yet to be established. Moreover, few scales include measures of patient agreement/understanding. These constructs are often captured on separate measures of therapeutic alliance rather than as a metric of patient adherence to treatment (Fluckiger et al., 2021). A critical gap in the field exists for a comprehensive and psychometrically-validated measure of patient adherence to treatment during psychological treatments that is easily-administered, and deployable, and considers both agreement (i.e., acceptance/agreement of the treatment contents) and behavioral compliance (i.e., homework completion, and usage of skills). Such a measure of patient adherence to treatment could be used across different EBPTs to further characterize and understand the phenomenon of patient adherence/non-adherence to treatment with the goal of improving clinical outcomes.

The present study focuses on two samples. The first sample includes adults with Major Depressive Disorder (MDD) receiving cognitive therapy (CT) for depression. We chose this focus as depression is one of the most prevalent mental disorders and a leading cause of disability worldwide (Murray et al., 2012; Vos et al., 2012). The majority of patients who recover from depression will relapse (Eaton et al., 2008). Although CT is a widely

studied, evidence-based, and frontline treatment for depression (Cuijpers, Berking, et al., 2013; Cuijpers, Hollon, et al., 2013), room still remains for improvement (Bockting et al., 2015; Jarrett et al., 2013). The second sample incudes adolescents with an evening circadian preference, or "night owls," who are at-risk for multiple health-related problems. Individuals with an evening circadian preference have a preference for going to bed later and waking later (Carskadon et al., 1993; Roenneberg et al., 2004). Approximately 40% of adolescents experience a shift towards an evening circadian preference, which coupled with early school start times, contributes to a cycle of insufficient sleep during adolescence (Crowley et al., 2018). An evening circadian preference is associated with an increase in risk for a host of negative outcomes such as affective problems such as depression, anxiety, and suicidality (Goldstein et al., 2008; Gregory & Sadeh, 2012); substance use and impulsivity (Adan et al., 2010; Hasler et al., 2016), aggressive and antisocial behavior (Díaz-Morales et al., 2014; Schlarb et al., 2014), poor academic performance (Short et al., 2013), and obesity (Asarnow et al., 2017). These youth are particularly high-risk for negative outcomes. Both samples received an EBPT for their respective psychological problems.

The overall goal of this study is to report on the initial development of the Treatment Adherence Rating Scale (TARS) and assess its psychometric properties in the context of psychological treatments for adults with MDD and community-dwelling youth with an evening circadian preference and at-risk for multiple health-related problems. The present study has two specific aims. The first aim is to establish the internal consistency and structural validity of the TARS. We hypothesize that (a) the TARS scores will demonstrate adequate internal consistency and (b) that the five TARS items will be mapped onto a two-factor model (i.e., agreement and compliance), based on the WHO definition of adherence (World Health Organization, 2003). The second aim is to examine the predictive validity of TARS scores. We hypothesize that TARS scores will demonstrate predictive validity as evidenced by: 1) negative correlations with depressive symptoms in Sample 1 and an evening circadian preference and average bedtime in Sample 2 at post-treatment and follow-up and 2) positive correlations with functional improvement scores in Sample 1 and total sleep time in Sample 2 at post-treatment and follow-up. The third aim is to begin the process of establishing the construct validity of the TARS scores. We hypothesize that TARS scores will demonstrate convergent validity as evidenced by 1) positive correlations with treatment evaluation at post-treatment and follow-up in Sample 1 and Sample 2 and 2) positive correlations with a behavioral indicator of homework completion (i.e., average number of days sleep that was recorded on sleep diary and number of sleep diaries completed over the course of treatment) in Sample 2.

Methods

Development of the TARS

TARS items were derived based on the WHO definition of adherence (World Health Organization, 2003), which postulates that adherence has two subcomponents, agreement and compliance. Adherence is commonly indicated by behavioral compliance to treatment recommendations (e.g., Mausbach et al., 2010). In the medical literature, collaborative agreement between patients and providers agreement has also been identified as an

important aspect of adherence (Bissonnette, 2008; Vermeire et al., 2001). Therefore, TARS items incorporate both the agreement and the compliance aspects of adherence. The general principles used in generating the TARS items were to construct items that are efficient, easily-administered, and easily-adaptable across various EBPTs. The wordings and responses of the TARS items are presented in Supplement 1. Items 1 and 2 in TARS are intended to capture patient's agreement and understanding of in-session treatment contents. Items 3–5 in TARS are to capture out-of-session compliance with treatment recommendations. The intended use of the TARS is for the therapist to rate each client's treatment adherence on the five TARS items on a scale of 0% to 100% with 10% increments at the end of each weekly treatment session.

Participants and Procedures

Sample 1.—Participants were 48 adults with MDD (mean age = 44.27 years, 29 women), who received 14 weekly, 50-minute sessions of Cognitive Therapy (CT) for depression. Participants were randomized to receive either CT-as-usual according to published manuals (Beck, 1979), or CT plus a Memory Support Intervention (CT+Memory Support). The Memory Support Intervention was designed as an add-on to treatments-as-usual such as CT with the aim of improving patient memory of treatment contents (Harvey et al., 2014, 2016). The study was approved by the Institutional Review Board. Full details are available elsewhere (Harvey et al., 2016).The Memory Support Intervention did not lengthen or modify the CT content (Harvey et al., 2016). Treatment condition (CT-as-usual versus CT+Memory Support) did not differ significantly on the TARS scores. Therefore, the two conditions were combined for the purpose of the current study.

The inclusion criteria for Sample 1 were: 1) ≥ 18 years of age; 2) able and willing to give informed consent; 3) diagnosis of MDD according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000); 4) minimum score of 26 or above on the Inventory of Depressive Symptomatology, Self-Report (IDS-SR; Rush, Gullion, Basco, Jarrett, & Trivedi, 1996); 5) minimum scores of 24 or above on the Inventory of Depressive Symptomatology, Clinician Report (IDS-C; Rush et al., 1996); and 6) stable medication regimen for the past month if participants were taking psychotropic medications.

Participants in Sample 1 were excluded if they met any of the following exclusion criteria: 1) history of bipolar affective disorder; 2) history of psychosis or psychotic features (including schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, or psychotic organic brain syndrome); 3) current non-psychotic Axis I disorder that constitutes the principal diagnosis requiring treatment other than that offered within the study; 4) history of substance dependence in the past six months; 5) IQ below 80; 6) evidence of any medical disorder or condition that could cause depression, or preclude participation in CT or that is associated with memory problems; or 7) current suicidal risk sufficient to preclude treatment on an outpatient basis.

Sample 2.—Participants were 176 adolescents ages 10–18 years old with an evening circadian preference, who received six weekly, 50-minute sessions of individual therapy.

Participants were randomized to receive either the Transdiagnostic Sleep and Circadian Intervention for Youth (TranS-C; $n = 89$, mean age = 14.76 years, 49 women) or Psychoeducation (PE) ($n = 87$, mean age = 14.78 years, 53 females). The derivation and contents of TranS-C are documented in greater detail elsewhere (Harvey, 2015; Harvey & Buysse, 2017). Briefly, TranS-C targets modifiable psychosocial, behavioral, and cognitive processes. It was designed to treat sleep and circadian problems in youth and improve functional impairment. PE about sleep and health is an active control associated with sleep improvement (Harvey et al., 2015). A key difference between TranS-C and PE is that the emphasis of PE was only on providing information rather than facilitating behavior change. The study was approved by the Institutional Review Board at the University of California Berkeley. All participants provided informed consent or assent. Details are described elsewhere ((Harvey et al., 2018).

The inclusion criteria for Sample 2 was: 1) between 10 and 18 years of age, 2) residing with a parent or guardian, and going to school or work by 9am at least 3 days per week; 3) proficiency in English; 4) a minimum score of 27 on the Children's Morningness-Eveningness Preferences Scale; 5) a late sleep onset time for at least 3 nights per week according to a 7-day self-reported sleep diary (i.e., 10:40pm or later for 10–13 year olds, 11pm or later for 14–16 year olds, and 11:20pm or later for 17–18 year olds); and 6) evidence of being 'at risk' on one or more of the five domains of functioning (i.e., behavioral, cognitive, emotional, physical, social). Participants ceased taking medications related to sleep 30 days prior to the intake assessment (two weeks for melatonin).

Youths were excluded if they 1) experienced a physical or neurological degenerative condition related to sleep disturbance; 2) obstructive sleep apnea, restless legs syndrome, or periodic limb movement disorder, pervasive developmental disorder; 3) were diagnosed with bipolar disorder, schizophrenia, or another Axis I disorder or condition, which are associated with high risks of harm; or 4) experienced substance dependence in the past six months.

Measures

For both samples, we selected the primary clinical outcomes for each clinical trial as the outcome variables used to evaluate the validity of TARS scores. In addition to clinical outcomes, we have included a measure of treatment expectancy. Specific measures for each sample are described below.

Sample 1.

Inventory of Depressive Symptoms-Self Report (IDS-SR).: Participants' symptoms of depression were measured at post-treatment and 6-month follow-up using the IDS-SR. The IDS-SR is a widely-used, 30-item measure with established reliability and validity (Rush et al., 1996). Items were rated on a 4-point scale. In the current sample, the Cronbach's α for the IDS-SR items was 0.89 and 0.92 at post-treatment and 6-month follow-up respectively. A IDS-SR score for each time point was generated based on Rush's (1996) scoring instruction with higher scores indicating greater depression symptom severity.

Global Assessment of Functioning (GAF).: Functional impairment of participants measured using the GAF, which is an assessor-rated scale that ranges from 1 to 100 with lower scores indicating more severe impairment (American Psychiatric Association, 2000b). Excellent inter-rater reliability has been reported for the GAF between trained clinicians as well as between researchers (Hilsenroth et al., 2000; Startup et al., 2002; Vatnaland et al., 2007). There is also evidence supporting its validity for indicating global psychopathology/illness severity and change over time (Skodol et al., 1988). The GAF was rated at post-treatment and at 6-month follow-up.

Credibility and Expectancy Questionnaire (CEQ).: We examined credibility of therapy and expectancy of improvement using the Credibility and Expectancy Questionnaire (CEQ), which is a 5-item measure with satisfactory psychometric properties (Devilly & Borkovec, 2000). At the end of the second weekly session, post-treatment assessment, and follow-up assessments, participants rated credibility of treatment (item 1 to 4) on a 9-point Likert scale and circled the percentage of improvement in sleep symptoms they expected to occur (item 5). A sum score of the five items was generated for each participant with higher scores indicating higher credibility and expectancy. The CEQ was administrated to the participants at the end of the first treatment session.

Sample 2.

Children's Morningness-Eveningness Preference (CMEP).: Participants' circadian preference was examined via CMEP at baseline, post-treatment, 6-month follow-up, and 12-month follow-up assessments. The CMEP (Carskadon et al., 1993) is 10-item multiple-choice scale ranging from 10 to 43, with higher scores indicating a morning preference and lower scores indicating an evening circadian preference (Carskadon et al., 1993).

Sleep diary.: Trained undergraduate research assistants made a scripted phone call to the participants each morning to collect their 7-day sleep diary at baseline, post-treatment, 6-month follow-up, and 12-month follow-up assessments. Two sleep outcomes (Carney et al., 2012) were derived from the collected sleep diary in a manner that is consistent with standard guidelines (Buysse et al., 2006). Bedtime (BT) was calculated by averaging the bedtime reported on the sleep diary on weeknights. Total sleep time (TST) was calculated by subtracting time to fall asleep, sum of the duration of each awakening after sleep onset, and terminal wakefulness (i.e., time getting out of bed for the day minus time of the final awakening) from duration of time in bed. The average of TST on weeknights was calculated.

Completion of sleep diary.: Over the course of treatment, participants in the TranS-C condition were required to keep a daily sleep diary each morning to track their sleep. Participants in PE did not complete a sleep diary during the treatment because one key distinction between TranS-C and PE was that only TranS-C provided coaching for sleep change. The number of nights the sleep diary was completed by the youth between two adjacent treatment sessions was extracted for each participant. The average number of nights sleep diary was recorded during treatment served as a proxy for homework completion. If a participant partially or fully completed a sleep diary, the sleep diary was deemed complete.

The number of sleep diaries completed during treatment was served as a second proxy for homework completion.

CEQ.: The CEQ (Devilly & Borkovec, 2000) was administrated to participants at the end of the second session, post-treatment, and 6-month and 12-month follow-up.

Data Analysis

All data analyses were conducted in Stata 14 (StataCorp, 2015)and M*plus* 8 (Muthén & Muthén, 2017). The internal consistency of the TARS scores was assessed by calculating the inter-item correlations of TARS items and Cronbach's α. To identify the latent constructs associated with the ratings on the TARS items, we conducted exploratory factor analyses (EFAs) using the data from Sample 1, based on Fabrigar and colleagues' (1999) recommendations. We conducted confirmatory factor analyses (CFAs) with maximum likelihood estimation and fixed factor variance using data from Sample 2. As the TARS was administrated weekly to patients over the course of the treatment in both studies, we accounted for the non-independence among observations nested in individuals (i.e., within-class clustering) by controlling for clustering in our analyses using M*plus*. For both samples, we elected to use a single-level EFA/CFA and did not nest by therapist as the recommendations for clustering suggest having a minimum of 20 to include a variable as a cluster (Cameron & Miller, 2015; McNeish et al., 2017). In Sample 1 we had 3 therapists and in Sample 2 we have 18 therapists, which is lower than these recommendations.

EFA using Sample 1.—The appropriateness of EFA for Sample 1 was determined based on the goal of this study, which was to identify the factor structure of a newly developed measure with limited evidence to specify a prior factor model (Fabrigar et al., 1999). Sample size recommendations for EFA suggest a minimum sample of at least 100–300 cases (Ferguson & Cox, 1993), with 500 cases considered "very good" (Comrey & Lee, 1992). We include 552 scales in the analyses for Sample 1. We extracted one-factor and two-factor solutions applying oblique Geomin rotation, which allows the items to be correlated. The pairwise correlations between items ranged from 0.35 to 0.83. Simulation studies have demonstrated that Geomin oblique is suitable when the number of factors are below three (Browne, 2001). Fabrigar et al. (1999) listed iterated principal factor, principal factor, and maximum likelihood as common and reasonable methods for EFA model fitting. Maximum likelihood (ML) is recommended if the distributions of measured variables there is no presence of severe non-normality. We detected non-normality in ratings of TARS item 1 in Sample 1; However, in *Mplus* using maximum likelihood with robust standard errors and chi-square test statistic (MLR) (Yuan & Bentler, 2000) is robust to non-normality and non-independence of observations when complex data is specified (Muthén & Muthén, 2017).

Multiple criteria were used to determine the number of factors retained based on expert recommendations (e.g., Fabrigar et al., 1999; Henson & Roberts, 2006). Three criteria were used for the present study: 1) Kaiser's (1960) criterion, which involved assessing rating scores of TARS items, such that factors with eigenvalues above one were retained; 2) Cattell's (Cattell, 1966) scree test, which involved evaluating the rating scores of TARS

items, inspecting a graph of the observed eigenvalues ordered from largest to smallest, looking for natural break or drop-off point where the curve flattens off, and using the number of data points above the drop-off point as an indicator of number of factors to retain; and 3) Horn's (1965) parallel analysis, which involved comparing the observed eigenvalues to the "expected" eigenvalues obtained by a simulated distribution of eigenvalues from random samples that parallel the observed data in terms of the number of items and sample size, and using the number of observed eigenvalues exceeding expected eigenvalues as an indicator of factors to retain.

CFA using Sample 2.—To assess the validity of the proposed conceptual model from Sample 1, we performed a set of CFAs using data from Sample 2. CFAs offered parameter estimates, standard errors, goodness-of-fit statistics, and modification indices that can assist in model comparisons and scale refinements. A sample size of at least 150–315 cases is recommended for CFA (Boomsma & Hoogland, 2001; Muthén & Muthen, 2002). We include 919 TARS scales from Sample 2 in the CFA. ML was also used to conduct CFA, as the distributions of all TARS items in Sample 2 did not appear to violate the normality assumption (kurtosis < 7). We compared two models: 1) a one-factor model that includes all five items under a general factor of adherence, as indicated by EFA, and 2) a two-factor model (i.e., items 1 and 2 loading onto the "agreement/understanding" factor, items 3–5 loading onto the "compliance" factor), consistent with the WHO (2003) definition of adherence. Absolute fit criteria included comparative fit index (CFI) equal to or greater than 0.95 (Hu & Bentler, 1999), the Tucker-Lewis index (TLI) greater than 0.90 (Bentler, 1990), root mean square error of approximation (RMSEA) equal to or less than 0.08 (Browne & Cudeck, 1993; Schreiber et al., 2006). Comparative fit statistics included Akaike information criterion (AIC), Bayesian information criterion (BIC), and sample size adjusted BIC.

Predictive validity was assessed by examining the correlation coefficients between TARS scores and clinical outcome scores in Sample 1 (i.e., IDS-SR, GAF scores) and the correlation coefficients between TARS scores and primary clinical outcome scores in Sample 2 (i.e., CMEP, average weeknight bedtime, weeknight TST). Construct validity was assessed by examining the correlation coefficients between TARS scores and CEQ scores in Sample 1 and Sample 2 correlation coefficients between TARS scores and average number of completed sleep diary during treatment in Sample 2. In Sample 2, agreement/understanding scores were generated by summing ratings of items 1 and 2; compliance scores were calculated by summing ratings of item 3 to 5 of TARS. Average ratings on each available TARS items across sessions were used in this analysis. Data from both treatment groups in Sample 1 and data from the Trans-C group in Sample 2 were used in these analyses.

Results

Descriptive Statistics

Demographic information for Sample 1 is presented in Supplement 2. Demographic information for Sample 2 is presented in Supplement 3. The descriptive statistics (*M*s and *SD*s) of the TARS items are presented in Table 1.

Internal Consistency of the TARS Scores

Inter-item correlations of TARS are presented in Table 1. In both samples, the inter-item correlation coefficients were in the medium to large range based on Cohen's (1992) guidelines. Based on Nunnally's (Nunnally, 1978) guidelines, the internal consistency coefficients were considered "good" for both samples. Cronbach's α of TARS scores in Sample 1 was 0.87. Cronbach αs were 0.90 for agreement/understanding and 0.95 for compliance in the two-factor CFA model in Sample 2. The mean inter-item correlation coefficient was 0.83 for Sample 1 and the mean inter-item correlation was 0.87 in Sample 2.

Factor Structure

EFA using Sample 1.—Ratings of the TARS items during weekly treatment sessions in Sample 1 were used for EFAs. A total of 552 TARS scales were used in the EFA. Geomin rotated factor loadings from one- and two-factor solutions for TARS scores are presented in Table 2. For the TARS items, only one eigenvalue (3.36) was greater than one, suggesting a one-factor solution using Kaiser's (1960) criterion. Cattell's (1966) scree test suggested a two-factor solution, as two factors had eigenvalues above a visually determined drop-off point. Horn's (1965) parallel analysis suggested a two-factor solution, as two factors had observed eigenvalue exceeding its corresponding expected eigenvalues. As evident in Table 2, however, the two-factor solution was a Heywood case (i.e., item 4 had a negative unique variance), possibly due to too many common factors extracted given the relatively small total number of items (Dillon et al., 1987). Therefore, a one-factor solution was adopted.

CFA using Sample 2.—Table 3 shows the parameter estimates of both one-factor and two-factor models using Sample 2. Fit statistics are presented in Supplement 4. A total of 919 TARS scales were used in the CFA. Results suggested that the two-factor model fit better than the one-factor model. The two-factor model exhibited loadings that ranged between 0.78 to 0.96, with satisfactory model fit statistics, CFI=0.98, TLI=0.95, and RMSEA=0.05 (95% CI, 0.02–0.78). The one-factor CFA model loading 5 items was suboptimal due to poor absolute model fit, CFI=0.60, TLI=0.20, and RMSEA=0.18 [95% CI, 0.16–0.21], as well as poorer comparative fit statistics, AIC=26979, BIC=27051, adjusted BIC=27003, with respect to the two-factor model, AIC=26578, BIC=26655, adjusted BIC=26604. Thus, the two-factor CFA model was selected.

Predictive validity

As shown in Table 4, we examined the correlations between TARS scores and outcomes measures, including scores of IDS-SR and GAF in Sample 1. As displayed in Table 5, we examined the correlations between TARS scores and scores of primary outcomes, including CMEP, weeknight TST and BT. TARS scores were significantly and meaningfully correlated with scores from the relevant constructs in both samples. In Sample 1, TARS scores were negatively correlated with IDS-SR scores at post-treatment, and were positively correlated with GAF scores at post-treatment and 6-month follow-up. In Sample 2, compliance was positively correlated with CMEP scores at 6-month and 12-month follow-ups and was negatively correlated with TST at post-treatment. Agreement was not associated with

outcomes. For both samples, the corresponding effect sizes were in the small to medium range according to Cohen's (1992) guidelines.

Construct validity

Results for Sample 1 are displayed in Table 4 and results for Sample 2 are presented in Table 5. In Sample 1, the TARS was not correlated with the CEQ when it was administered after Session 1. In Sample 2, agreement was positively correlated with CEQ scores at 6-month and 12-month follow-ups, agreement was positively correlated with weekly average of nights sleep information that was recorded during treatment, and compliance was positively correlated with CEQ scores at post-treatment and 6-month follow-up. The corresponding effect sizes were in the small to medium range according to Cohen's (1992) guidelines.

Discussion

The overarching goal was to report on the development of the Treatment Adherence Rating Scale (TARS) and to provide preliminary results of its psychometric properties in the context of two psychological treatments. The first aim was to develop and assess the internal consistency and establish the factor structure of the TARS items. As predicted, the results indicated that the TARS demonstrated adequate internal consistency (Nunnally, 1978). An EFA was first conducted because there was no prior evidence regarding the potential factor structure. EFA results indicated a one-factor solution for the TARS. CFA supported a two-factor model, which mapped onto the two core components described in the WHO (2003) definition of adherence. It is worth noting that the two-factor solution for TARS scores also mapped onto the agreement and compliance dimensions, although this model was not selected because of the presence of a Heywood case.

The second aim was to evaluate the predictive validity of the TARS. In Sample 1, TARS scores were expected to be negatively correlated with depression symptom scores and positively correlated with functional impairment. Indeed, TARS scores were moderately negatively correlated with depressive symptoms scores at post-treatment. TARS scores were also positively correlated with functioning improvement at both post-treatment and 6-month follow-up. In Sample 2, higher TARS scores were hypothesized to be correlated with better sleep and circadian outcomes. Greater compliance scores were indeed positively correlated with less of an evening circadian preference (i.e., higher scores indicate a lower evening circadian preference) at the two follow-up time points. The findings from both samples are consistent with other studies examining patient adherence to treatment for depression and treatment for sleep problems that have demonstrated that a therapist-rated measure of adherence is associated with treatment outcome (Matthews et al., 2013; Vincent & Hameed, 2003). Surprisingly, total sleep time at post-treatment was significantly negatively correlated with compliance scores, which is divergent from prior research that has demonstrated a robust relationship between homework completion and treatment outcome (Kazantzis et al., 2010; Mausbach et al., 2010). Future research should replicate and clarify this unexpected relationship. Overall, results from Sample 1 and Sample 2 strongly support the predictive validity of TARS scores in that TARS scores are predictive of most relevant clinical outcomes.

The third aim was to begin the process of establishing convergent validity of the TARS. In Sample 1, TARS scores were expected to be significantly correlated with higher treatment expectancy and in Sample 2 agreement and compliance were expected to be significantly correlated with higher treatment expectancy and better homework completion (i.e., weekly completion of sleep diary). In Sample 1, TARS scores were not significantly associated with treatment expectancy. This may be because expectancy was assessed prior to most of the TARS ratings being completed. Future research is needed to unpack the relationship between expectancy and patient adherence to treatment recommendations. In Sample 2, both agreement and compliance were positively associated with treatment expectancy at 6-month follow-up. Compliance was positively correlated with treatment expectancy at post-treatment, while agreement was positively correlated with treatment expectancy at 12-month follow-up. Agreement was positively correlated with the average number of nights recorded on sleep dairy during treatment. Overall, the results offer preliminary support for the construct validity of the TARS as TARS scores were correlated with the included process constructs.

One strength of the TARS is that it is short and easily administered via a paper and pencil assessment that can be used as part of the routine procedures. Hence, it can be used during treatment sessions to monitor weekly treatment adherence. Second, the TARS items are derived from the definition of adherence provided by the WHO. As indicated, the items captured both understanding/agreement and behavioral compliance. Therefore, the TARS items appear to be more comprehensive than measuring treatment adherence using single indices such as homework completion. Third, the psychometric properties of the TARS were evaluated using two samples —first examined in Sample 1 (adults with MDD receiving CT), then verified using CFAs in Sample 2 (at-risk adolescents receiving psychosocial treatment for sleep problems). Moreover, as demonstrated in the current study, the contents of the TARS items are not specific to any particular EBPT or patient population and can be easily adapted to be used across patient populations, age groups, and types of EBPTs.

There are several limitations to the present study. First, we examined the psychometric properties of the TARS in adults with MDD receiving CT and adolescents with sleep and circadian problems receiving a psychosocial treatment. Future research should expand the utility of the TARS to other patient populations and other treatments to examine the generalizability of the TARS. Such research would be particularly important given that the TARS is intended to be deployable and adaptable to a variety of treatments and patient populations. Future studies should consider measuring skill use using validated measures (e.g., SoCT or CCTS) to further examine the discriminant validity of the TARS. Additionally, the evidence surrounding GAF should be interpreted with caution, as it was removed from the DSM-5 (Gold, 2014). Future studies should replicate this study and examine the latent construct with larger sample sizes to rework the items as needed. Future research should consider examining the relationship between TARS and session-by-session outcomes to further examine the impact of patient adherence to treatment recommendation on treatment outcomes. Without examining if the TARS as a predictor of subsequent symptom change, we are unable to disentangle if the results observed in this study were due to a relationship between prior symptoms change and TARS scores (i.e. patients who were already improving had better adherence). In addition, the TARS is a therapist-reported scale.

Therapist-rated scales are subject to biases, including social desirability. Future research examining patient ratings of these items are needed for further validation of the TARS. Note that therapist level intraclass correlation ranged from 0.011 to 0.164 in Study 1 and 0.106 to 0.195 in Study 2 and we did not include therapist as a cluster variable in our multilevel analyses due to relatively few number of clusters in both samples; future studies with more clusters (i.e., therapists in the study) and larger sample sizes should include therapist as a cluster variable or utilize bootstrapping to address potential biases in an unbalanced few cluster design (Cameron & Miller, 2015).

In sum, the current study provided preliminary results supporting the reliability and validity of TARS scores as a measure of patient's treatment adherence. This therapist-rated measure can be used to assess patient treatment adherence during weekly therapy sessions. The TARS has the potential to be readily used to monitor treatment adherence in a variety of EBPTs given that it is intended to be highly generalizable and not specific to adults with MDD or adolescents with an evening circadian preference receiving a psychosocial treatment. The development of the TARS provides an initial step to measure patient treatment adherence to EBPTs. Future studies are needed to further establish validity and reliability of the TARS scores and to apply this measure to other EBPTs and patient populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Abramowitz JS, Franklin ME, Zoellner LA, & DiBernardo CL (2002). Treatment compliance and outcome in obsessive-compulsive disorder. Behavior Modification, 26(4), 447–463. 10.1177/0145445502026004001 [PubMed: 12205821]

Adan A, Natale V, Caci H, & Prat G (2010). Relationship between circadian typology and functional and dysfunctional impulsivity. Chronobiology International, 27(3), 606–619. 10.3109/07420521003663827 [PubMed: 20524804]

American Psychiatric Association. (2000a). Diagnostic and statistical manual of mental disorders (4th ed., t). Author. 10.1176/appi.books.9780890423349

American Psychiatric Association. (2000b). Diagnostic and statistical manual of mental disorders (4th ed., text rev.).

Amico KR, Mugavero M, Krousel-Wood MA, Bosworth HB, & Merlin JS (2018). Advantages to Using Social-Behavioral Models of Medication Adherence in Research and Practice. Journal of General Internal Medicine, 33(2), 207–215. 10.1007/s11606-017-4197-5 [PubMed: 29204968]

Asarnow LD, Greer SM, Walker MP, & Harvey AG (2017). The Impact of Sleep Improvement on Food Choices in Adolescents With Late Bedtimes. Journal of Adolescent Health, 60(5), 570–576. 10.1016/j.jadohealth.2016.11.018

Beck AT (1979). Cognitive Therapy of Depression. Guilford Press.

Bentler P (1990). Comparative Fit Indices in Structural Models Permalink. Psychological Bulletin, 107(2), 238–246. [PubMed: 2320703]

Bissonnette JM (2008). Adherence: A concept analysis. Journal of Advanced Nursing, 63(6), 634–643. 10.1111/j.1365-2648.2008.04745.x [PubMed: 18808585]

Bockting CL, Hollon SD, Jarrett RB, Kuyken W, & Dobson K (2015). A lifetime approach to major depressive disorder: The contributions of psychological interventions in preventing relapse

and recurrence. Clinical Psychology Review, 41, 16–26. 10.1016/J.CPR.2015.02.003 [PubMed: 25754289]

Boomsma A, & Hoogland JJ (2001). The Robustness of LISREL Modeling Revisited. Structural Equation Modeling Present and Future, 139–168.

Browne MW (2001). An Overview of Analytic Rotation in Exploratory Factor Analysis. Multivariate Behavioral Research, 36(1), 111–150. 10.1207/S15327906MBR3601_05

Browne MW, & Cudeck R (1993). Alternative ways of assessing model fit. In Bollen KA & Long JS (Eds.), Testing structural equation models (pp. 136–162). SAGE Publications.

Buysse D, Ancoli-Israel S, Edinger J, Lichstein K, & Morin C (2006). Recommendations for a standard research assessment of insomnia. Sleep, 29(9), 1155–1173. 10.1093/sleep/29.9.1155 [PubMed: 17040003]

Cameron C, & Miller D (2015). A Practitioner's Guide to Cluster-Robust Inference (Robust Covariance for OLS). Journal of Human Resources, 50(2), 317–372.

Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, & Morin CM (2012). The Consensus Sleep Diary: Standardizing Prospective Sleep Self-Monitoring. Sleep, 35(2), 287–302. 10.5665/sleep.1642 [PubMed: 22294820]

Carskadon MA, Vieira C, & Acebo C (1993). Association between puberty and delayed phase preference. Sleep, 16(3), 258–262. 10.1093/sleep/16.3.258 [PubMed: 8506460]

Cattell RB (1966). The Scree Test For The Number Of Factors. Multivariate Behavioral Research, 1(2), 245–276. 10.1207/s15327906mbr0102_10 [PubMed: 26828106]

Cho JH, Kremer S, & Young J (2021). Who to Refer to a Behavioral Insomnia Clinic? — Recommendations Based on Treatment Rationale and Response Prediction. Current Sleep Medicine Reports, 7(4), 213–220. 10.1007/s40675-021-00220-3 [PubMed: 34804765]

Cohen J (1992). A power primer. Psychological Bulletin, 112(1), 155–159. 10.1037/0033-2909.112.1.155 [PubMed: 19565683]

Comrey AL, & Lee HB (1992). A First Course in Factor Analysis. Lawrence Eribaum Associates.

Crowley SJ, Wolfson AR, Tarokh L, & Carskadon MA (2018). An update on adolescent sleep: New evidence informing the perfect storm model. Journal of Adolescence, 67, 55–65. 10.1016/J.ADOLESCENCE.2018.06.001 [PubMed: 29908393]

Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, & Dobson KS (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Canadian Journal of Psychiatry, 58(7), 376–385. 10.1177/070674371305800702 [PubMed: 23870719]

Cuijpers P, Hollon SD, Van Straten A, Bockting C, Berking M, & Andersson G (2013). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. BMJ Open, 3(4), 1–8. 10.1136/bmjopen-2012-002542

Devilly GJ, & Borkovec TD (2000). Psychometric properties of the credibility/expectancy questionnaire. Journal of Behavior Therapy and Experimental Psychiatry, 31(2), 73–86. 10.1016/S0005-7916(00)00012-4 [PubMed: 11132119]

Díaz-Morales JF, Escribano C, Jankowski KS, Vollmer C, & Randler C (2014). Evening adolescents: The role of family relationships and pubertal development. Journal of Adolescence, 37(4), 425–432. 10.1016/J.ADOLESCENCE.2014.03.001 [PubMed: 24793390]

Dillon WR, Kumar A, & Mulani N (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. Psychological Bulletin, 101(1), 126–135. 10.1037/0033-2909.101.1.126

Dobson KS (2021). A Commentary on the Science and Practice of Homework in Cognitive Behavioral Therapy. Cognitive Therapy and Research, 45(2), 303–309. 10.1007/s10608-021-10217-5

Dong L, Soehner AM, Bélanger L, Morin CM, & Harvey AG (2018). Treatment agreement, adherence, and outcome in cognitive behavioral treatments for insomnia. Journal of Consulting and Clinical Psychology, 86(3), 294–299. 10.1037/ccp0000269 [PubMed: 29265834]

Eaton WW, Shao H, Nestadt G, Lee BH, Bienvenu OJ, & Zandi P (2008). Population-Based Study of First Onset and Chronicity in Major Depressive Disorder. Archives of General Psychiatry, 65(5), 513. 10.1001/archpsyc.65.5.513 [PubMed: 18458203]

Fabrigar LR, Wegener DT, MacCallum RC, & Strahanm EJ (1999). Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 4, 272–2999.

Ferguson E, & Cox T (1993). Exploratory factor analysis : a user ‘ s guide Exploratory Factor Analysis : A Users ‘ Guide. International Journal of Selection and Assessment, 1(2), 84–94.

Fluckiger C, Del Re A, Wlodasch D, Horvath AO, Solomonov N, & Wampold BE (2021). Assessing the alliance-outcome association adjusted for patient characteristics and treatment processes: A meta-analytic summary of direct comparisons. Journal of Counseling Psychology, 67(6), 706–711. 10.1037/cou0000424

Gold LH (2014). DSM-5 and the assessment of functioning: The world health organization disability assessment schedule 2.0 (WHODAS 2.0). Journal of the American Academy of Psychiatry and the Law, 42(2), 173–181. [PubMed: 24986344]

Goldstein TR, Bridge JA, & Brent DA (2008). Sleep Disturbance Preceding Completed Suicide in Adolescents. Journal of Consulting and Clinical Psychology, 76(1), 84. 10.1037/0022-006x.76.1.84 [PubMed: 18229986]

Gregory AM, & Sadeh A (2012). Sleep, emotional and behavioral difficulties in children and adolescents. Sleep Medicine Reviews, 16(2), 129–136. 10.1016/j.smrv.2011.03.007 [PubMed: 21676633]

Harvey AG (2015). A Transdiagnostic Intervention for Youth Sleep and Circadian Problems. Cognitive and Behavioral Practice. 10.1016/j.cbpra.2015.06.001

Harvey AG, & Buysse DJ (2017). Treating Sleep Problems: A Transdiagnostic Approach. Guilford Press.

Harvey AG, & Gumport NB (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. Behaviour Research and Therapy, 68, 1–12. 10.1016/j.brat.2015.02.004 [PubMed: 25768982]

Harvey AG, Hein K, Dolsen MR, Dong L, Rabe-Hesketh S, Gumport NB, Kanady J, Wyatt JK, Hinshaw SP, Silk JS, Smith RL, Thompson MA, Zannone N, & Blum DJ (2018). Modifying the Impact of Eveningness Chronotype (“Night-Owls”) in Youth: A Randomized Controlled Trial. Journal of the American Academy of Child & Adolescent Psychiatry, 57(10), 742–754. 10.1016/J.JAAC.2018.04.020 [PubMed: 30274649]

Harvey AG, Lee J, Smith RL, Gumport NB, Hollon SD, Rabe-Hesketh S, Hein K, Dolsen MR, Hamen K, Kanady JC, Thompson MA, & Abrons D (2016). Improving Outcome for Mental Disorders by Enhancing Memory for Treatment. Behaviour Research and Therapy, 81, 35–46. 10.1016/j.brat.2016.03.007 [PubMed: 27089159]

Harvey AG, Lee J, Williams J, Hollon SD, Walker MP, Thompson MA, & Smith R (2014). Improving Outcome of Psychosocial Treatments by Enhancing Memory and Learning. Perspectives on Psychological Science, 9(2), 161–179. 10.1177/1745691614521781 [PubMed: 25544856]

Harvey AG, Soehner AM, Kaplan KA, Hein K, Lee J, Kanady J, Li D, Rabe-Hesketh S, Ketter TA, Neylan TC, & Buysse DJ (2015). Treating insomnia improves mood state, sleep, and functioning in bipolar disorder: A pilot randomized controlled trial. Journal of Consulting and Clinical Psychology, 83(3), 564–577. 10.1037/a0038655 [PubMed: 25622197]

Hasler BP, Soehner AM, & Clark DB (2016). Sleep and circadian contributions to adolescent alcohol use disorder. Alcohol, 49(4), 377–387. 10.1016/j.alcohol.2014.06.010.Sleep

Henson RK, & Roberts JK (2006). Use of Exploratory Factor Analysis in Published Research. Educational and Psychological Measurement, 66(3), 393–416. 10.1177/0013164405282485

Hilsenroth MJ, Ackerman SJ, Blagys MD, Baumann BD, Baity MR, Smith SR, Price JL, Smith CL, Heindselman TL, Mount M. k, & Holdwick DJ (2000). Reliability and validity of DSM-IV axis V. American Journal of Psychiatry, 157(3), 1858–1863. [PubMed: 11058486]

Horn JL (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30(2), 179–185. 10.1007/BF02289447 [PubMed: 14306381]

Hu LT, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55. 10.1080/10705519909540118

Jarrett RB, Minhajuddin A, Kangas JL, Friedman ES, Callan JA, & Thase ME (2013). Acute phase cognitive therapy for recurrent major depressive disorder: Who drops out and how much do

patient skills influence response? Behaviour Research and Therapy, 51(4–5), 221–230. 10.1016/j.brat.2013.01.006 [PubMed: 23485420]

Jarrett RB, Vittengl JR, Clark LA, & Thase ME (2011). Skills of Cognitive Therapy (SoCT): A new measure of patients' comprehension and use. Psychological Assessment, 23(3), 578–586. 10.1037/a0022485 [PubMed: 21319902]

Kaiser HF (1960). The Application of Electronic Computers to Factor Analysis. Educational and Psychological Measurement, 20(1), 141–151. 10.1177/001316446002000116

Kazantzis N (2021). Introduction to the Special Issue on Homework in Cognitive Behavioral Therapy: New Clinical Psychological Science. Cognitive Therapy and Research, 45(2), 205–208. 10.1007/s10608-021-10213-9

Kazantzis N, Deane FP, & Ronan KR (2006). Homework Assignments in Cognitive and Behavioral Therapy: A Meta-Analysis. Clinical Psychology: Science and Practice, 7(2), 189–202. 10.1093/clipsy.7.2.189

Kazantzis N, Deane FP, Ronan KR, & L'Abate L (Eds.). (2005). Using Homework Assignments in Cognitive Behavior Therapy. Routledge Taylor & Francis Group.

Kazantzis N, Whittington C, & Dattilio F (2010). Meta-Analysis of Homework Effects in Cognitive and Behavioral Therapy: A Replication and Extension. Clinical Psychology: Science and Practice, 17(2), 144–156. 10.1111/j.1468-2850.2010.01204.x

Kazdin AE (2017). Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. Behaviour Research and Therapy, 88, 7–18. 10.1016/j.brat.2016.06.004 [PubMed: 28110678]

Lasalvia A, Bonetto C, Tansella M, Stefani B, & Ruggeri M (2008). Does staff-patient agreement on needs for care predict a better mental health outcome? A 4-year follow-up in a community service. Psychological Medicine, 38(1), 123–133. 10.1017/S0033291707000785 [PubMed: 17537280]

Leeuwerik T, Cavanagh K, & Strauss C (2019). Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: A systematic review and meta-analysis. Journal of Anxiety Disorders, 68(August), 102135. 10.1016/j.janxdis.2019.102135 [PubMed: 31704633]

Levensky ER (2006). Nonadherence to treatment. In Fisher JE & O'Donohue WT (Eds.), Practitioner's guide to evidence-based psychotherapy. Springer.

Matthews EE, Arnedt JT, McCarthy MS, Cuddihy LJ, & Aloia MS (2013). Adherence to cognitive behavioral therapy for insomnia: A systematic review. Sleep Medicine Reviews, 17(6), 453–464. 10.1016/j.smrv.2013.01.001 [PubMed: 23602124]

Mausbach BT, Moore R, Roesch S, Cardenas V, & Patterson TL (2010). The relationship between homework compliance and therapy outcomes: An updated meta-analysis. Cognitive Therapy and Research, 34(5), 429–438. 10.1007/s10608-010-9297-z [PubMed: 20930925]

McDonald BR, & Morgan RD (2013). Enhancing Homework Compliance in Correctional Psychotherapy. Criminal Justice and Behavior, 40(7), 814–828. 10.1177/0093854813480781

McNeish D, Stapleton LM, & Silverman RD (2017). On the unnecessary ubiquity of hierarchical linear modeling. Psychological Methods, 22(1), 114–140. 10.1037/met0000078 [PubMed: 27149401]

Meis LA, Polusny MA, Kehle-Forbes SM, Erbes CR, O'Dougherty M, Erickson EPG, Orazem RJ, Burmeister LB, & Spoont MR (2022). Making sense of poor adherence in PTSD treatment from the perspectives of veterans and their therapists. Psychological Trauma: Theory, Research, Practice, and Policy. 10.1037/tra0001199

Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, … Lopez A (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet, 380(9859), 2197–2223. 10.1016/S0140-6736(12)61689-4 [PubMed: 23245608]

Muthén LK, & Muthen B (2002). How To Use a Monte Carlo Study. Structural Equation Modeling, 9(4), 599–620. http://www.statmodel.com/bmuthen/articles/Article_096.pdf

Muthén LK, & Muthén BO (2017). Mplus user's guide (8th ed.).

Nunnally JC (1978). Psychometric theory (2nd ed.). McGraw-Hill.

Roenneberg T, Kuehnle T, Pramstaller PP, Ricken J, Havel M, Guth A, & Merrow M (2004). A marker for the end of adolescence. Current Biology, 14(24), 1038–1039. 10.1016/j.cub.2004.11.039

Rush AJ, Gullion CM, Basco MR, Jarrett RB, & Trivedi MH (1996). The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychological Medicine, 26, 477–486. 10.1017/S0033291700035558 [PubMed: 8733206]

Sachsenweger MA, Fletcher RB, & Clarke D (2015). Pessimism and Homework in CBT for Depression. Journal of Clinical Psychology, 71(12), 1153–1172. 10.1002/jclp.22227 [PubMed: 26768682]

Schlarb AA, Sopp R, Ambiel D, & Grünwald J (2014). Chronotype-related differences in childhood and adolescent aggression and antisocial behavior - A review of the literature. Chronobiology International, 31(1), 1–16. 10.3109/07420528.2013.829846 [PubMed: 24147657]

Schreiber JB, Stage FK, King J, Nora A, & Barlow EA (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. Journal of Educational Research, 99(6), 323–338. 10.3200/JOER.99.6.323-338

Short MA, Gradisar M, Lack LC, Wright HR, & Dohnt H (2013). The sleep patterns and well-being of Australian adolescents. Journal of Adolescence, 36(1), 103–110. 10.1016/J.ADOLESCENCE.2012.09.008 [PubMed: 23088812]

Simpson HB, Maher MJ, Wang Y, Bao Y, Foa EB, & Franklin M (2011). Patient Adherence Predicts Outcome from Cognitive-Behavioral Therapy in Obsessive-Compulsive Disorder. Journal of Consulting and Clinical, 79, 247–252.

Skodol AE, Link BG, Shrout PE, & Horwath E (1988). The revision of axis V in DSM-III-R: Should symptoms have been included? American Journal of Psychiatry, 145(7), 825–829. 10.1176/ajp.145.7.825 [PubMed: 3381926]

Startup M, Jackson MC, & Bendix S (2002). The concurrent validity of the Global Assessment of Functioning (GAF). British Journal of Clinical Psychology, 41(4), 417–422. 10.1348/014466502760387533 [PubMed: 12437796]

StataCorp. (2015). Stata Statistical Software: Release 14. StataCorp LP.

Street RL, Makoul G, Arora NK, & Epstein RM (2009). How does communication heal? Pathways linking clinician–patient communication to health outcomes. Patient Education and Counseling, 74(3), 295–301. 10.1016/j.pec.2008.11.015 [PubMed: 19150199]

Strunk DR, Brotman MA, & DeRubeis RJ (2010). The process of change in cognitive therapy for depression: Predictors of early inter-session symptom gains. Behaviour Research and Therapy, 48(7), 599–606. 10.1016/j.brat.2010.03.011 [PubMed: 20362978]

Strunk DR, Hollars SN, Adler AD, Goldstein L a, & Braun, J. D. (2014). Assessing Patients' Cognitive Therapy Skills: Initial Evaluation of the Competencies of Cognitive Therapy Scale. Cognitive Therapy and Research, 38(5), 559–569. 10.1007/s10608-014-9617-9 [PubMed: 25408560]

Tremblay V, Savard J, & Ivers H (2009). Predictors of the effect of cognitive behavioral therapy for chronic insomnia comorbid with breast cancer. Journal of Consulting and Clinical Psychology, 77(4), 742–750. 10.1037/a0015492 [PubMed: 19634966]

Vatnaland T, Vatnaland J, Friis S, & Opjordsmoen S (2007). Are GAF scores reliable in routine clinical use? Acta Psychiatrica Scandinavica, 115(4), 326–330. 10.1111/j.1600-0447.2006.00925.x [PubMed: 17355524]

Vermeire E, Hearnshaw H, Van Royen P, & Denekens J (2001). Patient adherence to treatment: three decades of research. A comprehensive review. Journal of Clinical Pharmacy and Therapeutics, 26(5), 331–342. 10.1046/j.1365-2710.2001.00363.x [PubMed: 11679023]

Vincent NK, & Hameed H (2003). Relation Between Adherence and Outcome in the Group Treatment of Insomnia. Behavioral Sleep Medicine, 1(3), 125–139. 10.1207/S15402010BSM0103_1 [PubMed: 15600217]

Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, … Memish ZA (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet, 380(9859), 2163–2196. 10.1016/S0140-6736(12)61729-2 [PubMed: 23245607]

World Health Organization. (2003). The World Health Report 2003: Shaping the Future.

Yuan KH, & Bentler PM (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. Sociological Methodology, 30(1), 165–200.

Table 1.

Descriptive Statistics and Inter-item of the Treatment Adherence Rating Scale (TARS) items

	Descriptive Statistics			Inter-item Correlations				
	N	***M***	***SD***	**1**	**2**	**3**	**4**	**5**
Sample 1								
1. Understanding	552	83.42	14.66	-				
2. Accept/agree with content	550	78.62	16.72	0.71	-			
3. Complete practice at home	504	68.99	24.75	0.35	0.41	-		
4. Adherence to instruction	503	71.73	19.24	0.48	0.57	0.80	-	
5. Overall mastery of skills	498	67.29	20.38	0.48	0.60	0.62	0.83	-
Sample 2								
1. Understanding	919	88.81	13.44	-				
2. Accept/agree with content	918	85.90	15.87	0.73	-			
3. Complete practice at home	397*	79.40	23.87	0.46	0.51	-		
4. Adherence to instruction	562*	77.63	23.55	0.40	0.50	0.82	-	
5. Overall mastery of skills	544*	76.98	23.39	0.46	0.55	0.81	0.91	-

Note. *N* = no. of observations. *M* = Mean. *SD* = Standard Deviation. For correlation coefficients, all $p < 0.001$ for correlation coefficients.

*The Psychoeducation group in Sample 2 did not have homework assigned, thus leading to a smaller sample size for item 3–5.

Table 2.

Results of Exploratory Factor Analyses on the Treatment Adherence Rating Scale (TARS) scores.

	One-Factor Solution		Two-Factor Solution		
	Factor Loadings	Unique Variance	Factor Loadings		Unique Variance
			Factor 1	Factor 2	
1. Understanding	0.52 *	0.73	**0.67** *	0.08	0.49
2. Accept/agree with content	0.62 *	0.62	**0.99** *	< 0.01	0.02
3. Complete practice at home	0.81 *	0.35	−0.04	**0.80** *	0.39
4. Adherence to instruction	0.98 *	0.04	0.00	**1.02** *	−0.05
5. Overall mastery of skills	0.85 *	0.28	0.21 *	**0.69** *	0.32

Note. Bold indicates rotated factor loadings on each factor. Results from oblique Geomin rotations were used.

* The two-factor solution for the TARS scores was a Heywood case, and was therefore not adopted. Factor 1 = Agreement/understanding. Factor 2 = compliance.

Table 3.

Confirmatory Factor Analysis Results for One-factor and Two-factor Models.

	One-Factor Model		Two-Factor Model	
	Loadings *(SE)*	**Communalities**	**Loadings *(SE)***	**Communalities**
1. Understanding	0.59 (0.06)	0.35	**0.78 *(0.05)**	**0.61**
2. Accept/agree with content	0.05 (0.05)	0.00	**0.93 *(0.04)**	**0.86**
3. Complete practice at home	0.87 (0.03)	0.76	0.86 (0.03)	0.74
4. Adherence to instruction	0.94 (0.03)	0.88	0.94 (0.03)	0.88
5. Overall mastery of skills	0.97 (0.01)	0.94	0.96 (0.02)	0.92

Note.

* Item 1 and Item 2 mapped onto Factor 1 = agreement/understanding. Item 3, Item 4, and Item 5 mapped on to Factor 2 = compliance.

Table 4.

Correlations between Treatment Adherence Rating Scale (TARS) Scores and Validity Measures in Sample 1.

	TARS total
IDS-SR	
Post-treatment	−0.39*
6-month follow-up	−0.24
GAF	
Post-treatment	0.33*
6-month follow-up	0.36*
CEQ	
After session 1	0.28

Notes. IDS-SR = Inventory of Depressive Symptoms- Self-Report. GAF = Global Assessment of Functioning. CEQ = Credibility Expectancy Questionnaire.

* $p < 0.05$

** $p < 0.01$.

Table 5.

Correlations between Treatment Adherence Rating Scale (TARS) Scores and Validity Measures in Sample 2.

	Agreement/Understanding	Compliance
CMEP		
Post-treatment	0.18	0.23
6-month follow-up	0.10	0.25*
12-month follow-up	0.23	0.28*
Average weeknight total sleep time		
Post-treatment	−0.09	−0.29*
6-month follow-up	−0.02	−0.09
12-month follow-up	−0.11	−0.06
Post-treatment	−0.12	−0.05
6-month follow-up	0.02	0.11
12-month follow-up	−0.01	< 0.01
CEQ		
Post-treatment	0.12	0.30*
6-month follow-up	0.37**	0.33*
12-month follow-up	0.26*	0.17
Completion of Sleep Diary (homework)		
Nights of sleep recorded during treatment	0.26*	0.22
Sleep diary completed during treatment	0.14	0.22

Notes. CMEP = Children's Morningness—Eveningness Preferences Scale. CEQ = Credibility and Expectancy Questionnaire.

* $p < 0.05$

** $p < 0.01$.