Available online at www.sciencedirect.com

ScienceDirect

Behavior Therapy 45 (2014) 541–552

Behavior Therapy

www.elsevier.com/locate/bt

Assessing Treatment Integrity in Cognitive-Behavioral Therapy: Comparing Session Segments With Entire Sessions

Florian Weck
Florian Grikscheit
Volkmar Höfling
Ulrich Stangier
University of Frankfurt

The evaluation of treatment integrity (therapist adherence and competence) is a necessary condition to ensure the internal and external validity of psychotherapy research. However, the evaluation process is associated with high costs, because therapy sessions must be rated by experienced clinicians. It is debatable whether rating session segments is an adequate alternative to rating entire sessions. Four judges evaluated treatment integrity (i.e., therapist adherence and competence) in 84 randomly selected videotapes of cognitive-behavioral therapy for major depressive disorder, social anxiety disorder, and hypochondriasis (from three different treatment outcome studies). In each case, two judges provided ratings based on entire therapy sessions and two on session segments only (i.e., the middle third of the entire sessions). Interrater reliability of adherence and competence evaluations proved satisfactory for ratings based on segments and the level of reliability did not differ from ratings based on entire sessions. Ratings of treatment integrity that were based on entire sessions and session segments were strongly correlated ($r = .62$ for adherence and $r = .73$ for competence). The relationship between treatment integrity and outcome was comparable for ratings based on session segments and those based on entire sessions. However, significant relationships between therapist competence and therapy outcome were only found in the treatment of social anxiety disorder. Ratings based on segments proved to be adequate for the evaluation of treatment integrity. The findings demonstrate that session segments are an adequate and cost-effective alternative to entire sessions for the evaluation of therapist adherence and competence.

Keywords: adherence; assessment; competence; session segments; treatment integrity

Treatment integrity refers to the degree to which a treatment is delivered as intended (Yeaton & Sechrest, 1981). In psychotherapy research, treatment integrity prevails when the therapist demonstrates a high level of adherence and competence. Accordingly, adherence is defined as the extent to which a therapist employs interventions as described in the treatment manual, and competence is defined as the extent to which the therapist implements these interventions in a skillful manner (Waltz, Addis, Koerner, & Jacobson, 1993).

A high level of treatment integrity is a precondition for ensuring the experimental validity of psychotherapy treatment studies (Perepletchikova & Kazdin, 2005; Schlosser, 2002; Weck, Bohn, Ginzburg, & Stangier, 2011b). Thus, internal validity is ensured when changes in the dependent variable (treatment outcome) can be attributed to the independent variable (the intervention), which the not the case when treatment is not implemented as intended. Furthermore, external validity is ensured when the findings are generalizable, which does not occur when the independent variable (the intervention) is

This research was supported by Grant WE 4654/4-1 from the German Research Foundation. We would like to thank Dr. Evelyn Behar for her detailed and helpful suggestions regarding style and content in an earlier draft of this manuscript.

Address correspondence to Florian Weck, Ph.D., Department of Clinical Psychology and Psychotherapy, University of Frankfurt, Varrentrappstrasse 40–42, Frankfurt D-60486, Germany; e-mail: weck@psych.uni-frankfurt.de.

0005-7894/45/541-552/$1.00/0

not replicable because it was not implemented as intended.

In contrast to the general importance of treatment integrity, its evaluation has been neglected in psychotherapy research. The assessment of treatment integrity was found to be inadequate in 89% of randomized controlled trials (RCTs) published in the six most influential psychological and psychiatric journals from 2000 to 2004 (Perepletchikova, Treat, & Kazdin, 2007). An adequate assessment of treatment integrity would involve the evaluation of both therapist adherence and therapist competence, an independent evaluation of therapists' competence in applying the treatment protocol to the clients, and the presentation of data on the validity and reliability of used integrity measures. The considerable time input and high costs associated with the assessment process are particularly important barriers to assessing treatment integrity (Perepletchikova, Hilt, Chereji, & Kazdin, 2009).

For the evaluation of treatment integrity, direct assessment methods, such as information from audiotapes or videotapes evaluated by independent judges, are preferable to indirect methods, such as therapist self-reports, because therapists tend to overestimate their level of treatment integrity (Martino, Ball, Nich, Frankforter & Carrol, 2009). Judges evaluating treatment integrity must also be experienced therapists to provide adequate assessment (Weck, Hilling, Schermelleh-Engel, Rudari, & Stangier, 2011). Moreover, it has been shown that 5–10 treatment sessions per client are needed to achieve sufficient generalizability for ratings of adherence and competence (Dennhag, Gibbons, Barber, Gallop, & Crits-Christoph, 2012). This scenario leads to the high costs associated with the evaluation of treatment integrity.

It is questionable whether it is necessary to consider entire treatment sessions, from start to finish, or whether a segment of the treatment session would be sufficient to ensure an adequate evaluation of treatment integrity. Earlier studies have shown that process variables (e.g., empathy, amount of transference) assessed on the basis of segments were less reliable and valid than process variables assessed on the basis of entire sessions (Bachrach et al., 1981; Mintz & Luborsky, 1971). However, those studies used only very short segments (4- and 5-minute segments) for the evaluation of process variables and do not refer to treatment integrity at all.

One recent study directly compared the reliability and validity of adherence and competence ratings based on session segments with those based on entire sessions (Weck, Bohn, Ginzburg, & Stangier, 2011a). Accordingly, session segments entailed the first-, middle-, and final third of the entire session, with all of those thirds lasting a mean of 20 minutes. Therapist adherence and competence in 34 videotaped therapy sessions of cognitive therapy for social anxiety disorder were evaluated by four independent judges. Two judges provided ratings based on entire therapy sessions and two judges based their ratings on all three session segments. The aggregated adherence and the aggregated competence evaluation based on the second segments (middle third of the entire sessions) yielded an interrater reliability ($ICC_{adherence}$ = .81, $ICC_{competence}$ = .71) and correlation with therapy outcome ($r_{adherence}$ = .55, $r_{competence}$ = .45) that were comparable to those of the entire sessions. Moreover, the evaluation based on the second segment correlated strongly with those based on entire sessions ($r_{adherence}$ = .65, $r_{competence}$ = .71). In contrast, ratings based on the first and third segment showed partly lower interrater reliability, lower or mostly insignificant relationships with therapy outcome, and lower correlations with the ratings based on entire sessions. The results imply that ratings based on the second segment seem to constitute a good approximation of ratings based on entire sessions. However, the generalizability of the findings was limited, because only the treatment of social anxiety disorder was examined, and only 10 therapists were employed. Moreover, judges who evaluated the second segment also evaluated the first and the final segments (the segments were presented in permuted order). Therefore, the reliability and validity of ratings based on the second segment could sometimes be overestimated, because judges could remember, in two thirds of the cases, information about the therapeutic process in the other segments. Therefore, further research that addresses these previous limitations and uses a larger sampling of treatment outcome investigations is necessary to demonstrate the suitability of session segments for the evaluation of treatment integrity.

The aim of the current study was to compare adherence and competence ratings based on session segments (i.e., the middle third segment of the entire session) to ratings based on entire sessions both in terms of (a) reliability and (b) relationship to treatment outcome. As such, cognitive-behavioral therapy (CBT), with a cognitive focus on three different disorders from three major diagnostic categories, were considered: mood disorder (i.e., major depressive disorder), anxiety disorder (i.e., social anxiety disorder), and somatoform disorder (i.e., hypochondriasis). We hypothesized that ratings of therapist adherence and competence that were based on segments would (a) be equivalent to and (b) strongly correlated ($r > .5$; according to Cohen, 1992) with ratings of therapist adherence and competence based on entire sessions (Hypothesis 1). We also hypothesized that ratings of