

Psychometric Evaluation of the Brief Adjustment Scale-6 (BASE-6): A New Measure of General  
Psychological Adjustment

Alexandra Paige Peterson

A thesis submitted in partial fulfillment of the  
requirements for the degree of  
Master of Science - Psychology

University of Washington

2015

Committee:

Corey Fagan

Ronald Smith

Program Authorized to Offer Degree:

Psychology

©Copyright 2015  
Alexandra Paige Peterson

University of Washington

**Abstract**

Psychometric Evaluation of the Brief Adjustment Scale-6 (BASE-6): A New Measure of General Psychological Adjustment

Alexandra Paige Peterson

Chair of the Supervisory Committee:

Corey Fagan, Ph.D.

Psychology

The Brief Adjustment Scale-6 (BASE-6) is a six-item self-report measure of general psychological adjustment developed as a shorter alternative to the most commonly used instrument for overall distress and functioning, the Outcome Questionnaire-45 (OQ-45.2). The OQ-45.2 is most frequently used to monitor client status throughout psychotherapy, referred to as routine outcome monitoring. This practice highlights client improvement and the need for therapeutic adjustments when clients are not improving as expected, resulting in faster symptom reduction, reduced likelihood of client dropout, and improved outcomes. I investigated the psychometric properties of the BASE-6, using (a) two different online community samples, each of which comprised a subsample in psychotherapy as well as a subsample not in psychotherapy and (b) a general university sample. In all samples, I found high correlations with the more extensive OQ-45.2, as well as very strong psychometric qualities. Compared with the OQ-45.2, respondents rated the measure as equally useful for measuring adjustment but viewed it as significantly easier to use, and expressed greater willingness to complete it on a weekly basis. Results indicate the BASE-6 holds promise as a brief and acceptable alternative when frequent administration of the OQ-45.2 is not feasible

## Initial Psychometric Evaluation of the Brief Adjustment Scale-6 (BASE-6): A New Measure of General Psychological Adjustment

Measuring client status in psychotherapy is a growing practice in the behavioral health field. Insurance companies, governing psychological organizations and policy makers are increasingly encouraging clinicians to use validated measures to assess how clients are doing in hopes of improving both client outcomes and provider accountability as well as decreasing healthcare expenditures. In 2011, the Centers for Medicaid and Medicare Services (CMS) developed the Physician Quality Reporting System (PQRS) and offered higher reimbursement rates to clinicians who provided evidence-BASE-6d data on service quality. In 2015, CMMS began financially penalizing practitioners by reducing reimbursement for those who failed to provide such data. Fortunately, in light of the shift towards data-driven care, many brief, well-validated and disorder-specific psychological assessment instruments are already accessible to behavioral health care providers (Beidas et al., 2015). However, there remains a need for valid, time-efficient measures that address general psychological adjustment, a construct broadly defined as an individual's subjective sense of distress and ability to function in daily life.

Measures of general psychological adjustment are typically transtheoretical and designed to capture symptoms, wellbeing, and functioning within a single instrument. They are also frequently used to monitor client status throughout psychotherapy, a practice called routine outcome monitoring (ROM). ROM is associated with improved client outcomes, faster, more accurate diagnoses and improved identification of clients at risk for treatment failure or dropping out of treatment altogether (Overington & Ionita, 2012; Shimokawa, Lambert, & Smart, 2010; Carlier et al., 2012). Thus far, general psychological adjustment has been primarily measured with the Outcome Questionnaire 45 (OQ-45.2), a well-validated, 45-item self-report instrument

that assesses symptom distress, interpersonal problems, and social role functioning (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005). Unfortunately, the length and content complexity of the OQ-45.2 negatively impacts completion and compliance. Clients report measure fatigue when given the OQ-45.2 on a weekly basis (Miller, Duncan, Brown, Sparks, & Claud, 2003), and lengthy measures are often unattractive to behavioral healthcare providers. Clinicians in a managed care setting responded so negatively to the length and frequent administration of a 30-item outcome questionnaire introduced into daily practice that the State Psychological Association President stated “I have never seen such a negative reaction from providers” (Hanlon 2005, pg. 11). Clinicians’ ratings of how practical (e.g., time efficient, affordable, burdensome) a standardized assessment tool is significantly predicts use (Jensen-Doss & Hawley, 2010) and providers assert any measure that takes longer than five minutes to fill out, score, and interpret is unattractive for frequent administration (Brown et al., 1999). Thus, short and simple measures are optimal in order to make regular assessment of general psychological adjustment palatable to both clients and busy clinicians.

In an attempt to address this problem, Miller et al., (2003) created an ultra-brief substitute for the OQ-45.2, the four-item Outcome Rating Scale (ORS). The ORS translates the three dimensions of the OQ-45.2 and a question addressing overall adjustment into a visual analogue format (scored with a ruler) that takes about a minute to complete. The feasibility and appeal of shorter measures like the ORS was demonstrated by comparing completion rates, with the ORS achieving 89% and the OQ-45.2 only reaching 25% over a 12 month period (Miller et al., 2003). Though this suggests the length of the ORS is ideal, the measure demonstrates moderate concurrent validity with the gold-standard OQ-45.2 and lacks information that could be beneficial, such as communicating the nature of the client’s distress (e.g. anxiety, depressed

mood, irritability). Additionally, one item in the ORS encompasses work and friendships, two areas of functioning that could elicit very different reports.

The Brief Adjustment Scale-6 (BASE-6) was created to address the still unmet need for a brief, comprehensive, and valid measure of general psychological adjustment for use in psychotherapy. I aimed for an acceptable alternative to the OQ-45.2 that would provide clinical detail beyond what the ORS offers, yet would be robust enough to stand on its own or as a complement to disorder-specific measures. In this article, I report on a study designed to evaluate the psychometric properties of this new measure. It was hypothesized that the BASE-6 would demonstrate: (1) acceptable internal consistency and test-retest reliability; (2) high concurrent validity with the OQ-45.2 in the form of positive correlations between the BASE-6 total score and OQ-45.2 total and subscale scores; and (3) higher acceptability than the OQ-45.2 in terms of ease of use, utility, and participant willingness to complete the measure on a weekly basis.

### **Method**

**Participants.** Following institutional approval of the research, 302 adults were recruited for the pilot sample (Mturk 1) through Amazon Mechanical Turk (Mturk), a crowd-sourced online participant pool. Eligibility requirements included US residency and fluency in written and spoken English. Twenty-five participants were dropped due to missing data that precluded creating summary scores on measures (final  $n = 277$ ). Participants had a mean age of 30.1 ( $SD = 10.1$ ) and the majority were male (56%), had a graduate or professional degree (52%), and elected not to provide their ethnicity (73%). The sample was split into a psychotherapy group (Tx1;  $n = 114$ ; 41%) and a no-psychotherapy group (NoTx1;  $n = 163$ ; 58%), BASE-6d on whether participants reported current participation in psychotherapy (question: “Are you currently participating in mental health counseling/therapy?”). On average, participants in Tx1

( $M = 66.40$ ,  $SD = 25.25$ ) scored in the clinical range on the OQ-45.2 (greater than 63) while those in NoTx1 scored below ( $M = 61.15$ ,  $SD = 25.04$ ) (Lambert et al., 1999).

Participants in the second sample (Mturk 2) included 601 adults recruited through Mturk. Eligibility requirements were identical to those of the pilot with the addition of a high-reputation Mturk identifier (work approval rating of 95% or higher). Research has indicated that using reputation criteria results in higher quality data (Peer, Vosgerau, & Acquisti, 2014). Sixty-five participants were excluded from analyses for failing to provide consent ( $n = 28$ ) or careless responding as indicated by bogus control items ( $n = 37$ ), (final  $n = 536$ ). The majority of participants were Caucasian (73.1%), female (51.3%) and college educated (some college, 36.4%; undergraduate degree, 35.6%; some graduate school, 6.2%; graduate/professional degree, 12.5%), with a mean age of 33.7 years ( $SD = 12.4$ ). This sample was also split into a psychotherapy group (Tx2;  $n = 77$ ; 14%) and a no-psychotherapy group (NoTx2;  $n = 459$ ; 86%), participants in Tx2 scoring above the OQ-45.2 clinical cut-off ( $M = 77.10$ ,  $SD = 25.83$ ) and those in NoTx2 scoring below ( $M = 54.93$ ,  $SD = 27.04$ ).

The third sample (Student) included 267 undergraduate students from a large university in the Northwestern U.S. Students were recruited through the university psychology subject pool and obtained course credit in exchange for taking part in the study. To participate, participants were 18 or older and fluent in written and spoken English. Twenty-three were excluded for careless responding as indicated by bogus items (final  $n = 244$ ). Students had a mean age of 18.94 ( $SD = 1.82$ ) and were predominantly Asian (50.8 %) or European-American (34.4%), and female (73.4%). Due to attrition, 151 students provided data for test-retest reliability.

**Procedure.** Participants in each sample were provided a link where they accessed the questionnaire battery. Upon accessing the link, they were provided with an Institutional Review

Board approved electronic consent form. At this point they chose to participate or opt out. Those who chose to participate were asked to complete the full questionnaire battery sequenced as follows: demographics, adjustment measure, acceptability questionnaire, adjustment measure, and acceptability questionnaire. Adjustment measures (BASE-6 and OQ-45.2) were counterbalanced to ensure scores on measures were not due to the order in which they were presented. After this sequence the participants concluded with an electronic debriefing statement. Participants in the Student sample also completed the adjustment measures at an additional time point to investigate test-retest reliability. Students were given access to the first set of questionnaires online in Week 1 of the study; the questionnaires were available for students to complete for six days. Students obtained access to the second set of questionnaires Week 2 of the project (also available for 6 days).

### **Measures.**

**Outcome Questionnaire 45.2 (OQ-45.2).** As the criterion measure of psychological adjustment, I used the Outcome-Questionnaire-45, a 45-item self-report measure (Lambert et al., 1996). The OQ-45.2 was designed to provide information on a client's progress in psychotherapy through repeated administration, and produces a total distress score and three subscale scores: Symptom Distress, Interpersonal Relations, and Social Role Performance. Each item is rated on a 5-point scale (0 = never, 1 = rarely; 2 = sometimes; 3 = frequently; 4 = almost always), with higher scores indicating endorsement of pathology (Lambert et al., 1999). Completion takes approximately 5 minutes. The OQ-45.2 has been researched extensively, and has demonstrated good concurrent and construct validity across a wide range of patient populations, adequate internal consistency, test-retest reliability, and sensitivity to change over short time periods

(Lambert et al., 1996; Umphress et al., 1997; Boswell et al., 2013; Vermeersch et al., 2000; Vermeersch et al., 2004).

**Brief Adjustment Scale-6 (BASE-6).** The second measure of adaptive functioning, and the primary focus of this validation study, was the Brief Adjustment Scale-6. The BASE-6 is a brief self-report measure of general psychological adjustment developed for stand-alone use and as part of a measurement feedback system for ROM (Smith et al., 2011; currently called OwlOutcomes). The item pool contains six items and assesses individuals' perceptions of emotional distress (three items: depression, anxiety and, anger) and related interference (three items: self-esteem, personal relationships, and occupational functioning). Sample items include "To what extent have you felt unhappy, discouraged, and/or depressed this week?" and "How much has emotional distress interfered with feeling good about yourself this week?" All six items were on a seven-point Likert scale (1 = *not at all*, 4 = *somewhat*, 7 = *extremely*), higher scores indicating lower general psychological adjustment. Completion of the BASE-6 required approximately one minute.

**Acceptability Questionnaire.** I evaluated acceptability of the OQ-45.2 and the BASE-6 using participants' ratings on four items that assessed whether participants found each measure frustrating and difficult to complete, the degree to which they would be willing to complete the measure weekly in psychotherapy, and the degree to which they believed the measure would be useful in psychotherapy. Each item was rated on a 7-point Likert scale (1 = *not at all*, 4 = *somewhat*, 7 = *extremely*).

**Bogus Items.** I inserted three bogus items, designed to assess careless responding, at random within the full questionnaire battery taken by the Mturk2 and Student samples. Bogus items had an unambiguous correct answer (e.g., "I have never brushed my teeth"), and incorrect

answers indicated a high likelihood of dishonest or inattentive responding (Meade & Craig, 2012).

### Results

**Reliability.** To address measure reliability I examined internal consistency (via Cronbach's alpha) in all samples and test-retest reliability (via an intraclass correlation coefficient) in the Student sample. BASE-6 scores exhibited excellent internal consistency, ranging from  $\alpha = 0.87$  to  $\alpha = 0.92$ . Test-retest reliability was in the moderate range, with the BASE-6 exhibiting an ICC (2, 2) of 0.76, 95% CI = [0.68, 0.83].

**Correlations between BASE-6 and OQ-45.2.** I next assessed concurrent validity by examining correlation coefficients between the BASE-6 and the OQ-45.2 total score and subscale scores. Product-moment correlations and correlations corrected for attenuation attributable to instrument unreliability are shown in Table 1. Inspection of the table reveals that BASE-6 total scores exhibited strong positive correlations with OQ-45.2 total scores in each group (Tx1, Tx2, NoTx1, NoTx2 and Student). Additionally, the BASE-6 demonstrated positive correlations of medium to large magnitude with the OQ-45.2 subscales. Correlations between the BASE-6 and OQ-45.2 Social Role Performance and Interpersonal Relations subscales are slightly lower than for Symptom Distress; this may be because each subscale is encompassed in a single item on the BASE-6. Overall, these respectable, positive correlations suggest that the BASE-6 measures similar characteristics as the OQ-45.2.

**Acceptability.** Finally, I tested whether there were differences in acceptability between the BASE-6 and OQ-45.2 (see Table 2). In all samples, participants rated the BASE-6 as easier to use than the OQ-45.2 and reported they would be more willing to fill it out on a weekly basis. There was no difference in how useful participants thought the BASE-6 and OQ-45.2 would be

for quickly communicating information about their adjustment to a therapist. These results suggest that the BASE-6 is more appealing than the longer OQ-45.2 and remains just as useful from the perspective of the individual completing the measure.

### **Discussion**

The results of the present study provided moderate to strong support for the reliability, concurrent validity, and acceptability of the BASE-6, suggesting it may be a viable and acceptable alternative to the OQ-45.2 for measuring general psychological adjustment, as well as an alternative to other brief measures that do not account for type of distress (e.g., ORS). Clinician and client preference for brief measures has been well documented and the BASE-6 has the potential to fit the need for an easy to use measure of general distress and functioning (Brown et al., 1999; Jensen-Doss & Hawley, 2010; Miller et al., 2003).

There are multiple future directions that will further the validation process of this new measure. Firstly, the factor structure of the BASE-6 has yet to be established; a hypothesized unidimensional factor structure should be investigated. It would also be beneficial to examine BASE-6 properties through an Item Response Theory perspective, which would indicate whether BASE-6 items perform differently depending on variables such as gender, which items contribute the most to BASE-6 total score and whether the full 7-point Likert response scale is necessary. Furthermore, the empirical investigation of the BASE-6 as a tool for routine outcome monitoring is still in its infancy. The strong correlations between the OQ-45.2 and the BASE-6 and the extent to which participants viewed the BASE-6 as easy to use and acceptable for weekly administration suggest that it has potential as a tool for session-by-session monitoring. Future research should focus on evaluating the BASE-6 as a measure of weekly therapeutic progress, ideally validated with non-self-report criteria. This would entail evaluating how sensitive the

BASE-6 is to change, as well as establishing criteria that would indicate when change was clinically meaningful. Related to this goal would be establishing a BASE-6 clinical cut-off that would accurately discriminate between individuals who fall into the clinical range on other indices of distress and functioning (e.g., OQ-45.2) and those who do not. Finally, though this study clearly shows the BASE-6 is more acceptable to respondents than the OQ-45.2, it would be beneficial to evaluate BASE-6 acceptability from the perspective of a clinician. It is possible the BASE-6 may be viewed as even more acceptable by this group due to how simple the measure is to score.

**Limitations.** A limitation in this study is the lack of diversity of our samples (primarily educated in all three samples; largely indeterminate ethnicity in Mturk 1, Caucasian in Mturk 2, and Asian in the Student sample), therefore further investigation of the BASE-6 in diverse populations is warranted. Additionally, it is unclear whether participants reporting they were currently in psychotherapy fit any particular diagnostic criteria. Though the BASE-6 was designed to cut across diagnoses, it is possible the measure could be more useful for some groups than others. Therefore, evaluating the performance of the BASE-6 in further clinical samples is necessary.

Overall, the BASE-6 may provide an additional option for assessing general psychological adjustment among psychotherapy clients, and has notable potential to be integrated into routine outcome monitoring due to its brevity and acceptability.

## References

- Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., Mandell, D. S. (2015). Free, brief, and validated: standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice, 22*(1), 5–19.  
<http://doi.org/10.1016/j.cbpra.2014.02.002>
- Boswell, D.L., White, J.K., Sims, W.D., Harrist, R., & Romans, J.C. (2013). Reliability and validity of the Outcome Questionnaire. *Psychological Reports, 112*(3), 689-693.  
doi:10.2466/02.08.PR0.112.3.689-693.
- Brown, J., Dreis, S., & Nace, D. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M. Hubble, B. Duncan, & S. Miller (Ed.), *The heart and soul of change* (pp. 389-406). Washington, DC: APA Press.
- Carlier, I. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. A., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *Journal of Evaluation In Clinical Practice, 18*(1), 104-110.  
doi:10.1111/j.1365-2753.2010.01543.x
- Hanlon, P. (2005). PacifiCare screening tool, policies raise concerns. *New England Psychologist, 13*, 11–12.
- Jensen-Doss, A., & Hawley, K. M. (2010). Understanding barriers to evidence-BASE-6d assessment: clinician attitudes toward standardized assessment tools. *Journal of Clinical Child & Adolescent Psychology, 39*(6), 885–896.
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy, 3*(4), 249-258.

Lambert, J. L., Hansen, N. B., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W.

(1999). Administration and scoring manual for the OQ-45.2 (2<sup>nd</sup> ed.). Jupiter, FL:

American Professional Credentialing Services.

Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing

feedback to psychotherapists on their patients' progress: clinical results and practice suggestions. *Journal of Clinical Psychology, 61*(2), 165-174.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data.

*Psychological Methods, 17*, 3, 437-55.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., Claud, D. A. (2003). The Outcome Rating

Scale: a preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*(2), 91-100.

Overington, L., & Ionita, G. (2012). Progress monitoring measures: a brief guide. *Canadian*

*Psychology, 53*(2), 82-92. doi:10.1037/a0028017

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality

on Amazon Mechanical Turk. *Behavior Research Methods, 46* (4), 1023-1031.

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of

patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology, 78*(3), 298-311. doi:10.1037/a0019247

Smith, R., Fagan, C., Wilson, N., Chen, J., Corona, M., Nguyen, H., Racz, S., & Shoda, Y.

(2011). Internet-based approaches to collaborative therapeutic assessment: new opportunities for professional psychologists. *Professional Psychology-Research and Practice, 42*(6), 494-504.

- Umphress, V., Lambert, M., Smart, D., Barlow, S., & Clouse, G. (1997). Concurrent and construct validity of the outcome questionnaire. *Journal of Psychoeducational Assessment, 15*(1), 40-55.
- Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome Questionnaire-45: Item sensitivity to change. *Journal of Personality Assessment, 74*, 242-261.
- Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome Questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology, 51*(1), 38-49.

Table 1. Product-Moment Correlations and Coefficients Corrected for Attenuation between the BASE-6 and OQ-45.2

Group		OQ-45.2 Total	OQ-45.2 Symptom	OQ-45.2 Interpersonal	OQ-45.2 Social Role
Tx1	<i>r</i>	0.72	0.72	0.41	0.64
( <i>n</i> =114)	<i>att. r</i>	0.79	0.79	0.54	0.89
Tx2	<i>r</i>	0.74	0.72	0.58	0.60
( <i>n</i> =77)	<i>att. r</i>	0.82	0.80	0.69	0.75
NoTx1	<i>r</i>	0.75	0.73	0.58	0.68
( <i>n</i> =163)	<i>att. r</i>	0.81	0.89	0.71	0.85
NoTx2	<i>r</i>	0.81	0.80	0.68	0.69
( <i>n</i> =459)	<i>att. r</i>	0.86	0.86	0.76	0.83
Student	<i>r</i>	0.66	0.61	0.54	0.57
( <i>n</i> =244)	<i>att. r</i>	0.74	0.69	0.67	0.77

\*All correlations significant  $p < .001$

Table 2. Differences in Evaluative Ratings between the BASE-6 and the OQ-45.2

Sample	<i>Mean</i> BASE-6	<i>Mean</i> OQ-45.2	<i>t</i>	<i>p</i>	95% <i>CI</i>
<b>Ease of Use*</b>					
Mturk 1	3.79	4.27	$t(274) = -3.86$	<.001	[-0.71, -0.23]
Mturk 2	2.78	3.21	$t(535) = -5.91$	<.001	[-0.57, -0.29]
Student	2.92	3.37	$t(243) = -3.81$	<.001	[-0.69, -0.22]
<b>Suitability for Weekly Use</b>					
Mturk 1	4.86	4.65	$t(276) = 2.41$	<.05	[0.04, 0.38]
Mturk 2	5.03	4.69	$t(535) = 5.49$	<.001	[0.22, 0.47]
Student	4.59	4.26	$t(243) = 3.60$	<.001	[0.15, 0.51]
<b>Utility</b>					
Mturk 1	4.94	4.81	$t(274) = 1.59$	<i>ns</i>	[-0.03, 0.29]
Mturk 2	5.02	5.05	$t(535) = -0.57$	<i>ns</i>	[-0.13, 0.07]
Student	4.52	4.59	$t(243) = -1.02$	<i>ns</i>	[-0.19, 0.06]

\*Lower value indicates greater ease of use.