Canadian Journal of Counselling and Psychotherapy /
Revue canadienne de counseling et de psychothérapie
ISSN 0826-3893 Vol. 55 No. 4 © 2021 Pages 437–461
https://doi.org/10.47634/cjcp.v55i4.71356

Progress Monitoring Measures in Psychotherapy and Mental Health: Examination of a Structural Equation Model
Indicateurs de suivi de progrès en psychothérapie et en santé mentale : étude d'un modèle d'équations structurelles

Caissie Bourque
Université de Moncton
Andréanne Charbonneau
Université de Moncton
Geneviève Bouchard
Université de Moncton
Mylène Lachance-Grzela
Université de Moncton

ABSTRACT
Progress monitoring measures are psychometric assessments designed to monitor treatment progress by evaluating global ratings of mental health. Their use in clinical practice is related to increases in the effectiveness of therapy and decreases in negative outcomes. Yet, clinicians often report struggling with limited understanding of the differences between the numerous measures available, and only a small percentage of clinicians report using them. In order to assist clinicians in measure selection, the current study documented and compared the ability of five measures (i.e., the ORS, the OQ-45, the BASIS-24, the CORE-OM, and the TOP) to assess two aspects of mental health: psychological well-being and psychological distress. Data from a clinical sample of 53 French-speaking individuals were analyzed using structural equation modelling. Results showed strong convergent validity between the measures and indicated that they evaluated mental health accurately. Further analysis specified that, among the five measures, the OQ-45 and the CORE-OM were the best at assessing mental health.

RÉSUMÉ
Les indicateurs de suivi de progrès sont des instruments psychométriques conçus pour mesurer le progrès thérapeutique en s'appuyant sur des évaluations globales de la santé mentale. Leur utilisation en pratique clinique est liée à des hausses de l'efficacité de la thérapie et à des diminutions des effets négatifs. Pourtant, les cliniciens rapportent

qu'ils ont souvent une compréhension limitée des différences entre les nombreux indicateurs disponibles, et seul un faible pourcentage d'entre eux dit les utiliser. Afin d'aider les cliniciens à choisir parmi les indicateurs de suivi de progrès, l'étude actuelle a documenté et comparé l'aptitude de cinq instruments (c.-à-d., l'ORS, l'OQ-45, le BASIS-24, le CORE-OM, et le TOP) à mesurer deux aspects de la santé mentale : le bien-être psychologique et la détresse psychologique. Les données provenant d'un échantillon clinique de 53 francophones ont été analysées au moyen de la modélisation d'équations structurelles. Les résultats montrent une forte validité convergente entre les indicateurs et révèlent qu'ils évaluent la santé mentale avec exactitude. Une analyse plus poussée a précisé que parmi les indicateurs, l'OQ-45 et le CORE-OM étaient les meilleurs instruments pour évaluer la santé mentale.

Despite the empirically validated effectiveness of psychotherapy for treating various psychological problems (Howard et al., 1986; Lambert & Bergin, 1994), research indicates that not all clients see improvements in their mental health. Data from a Canadian national population study revealed that nearly half of the total number of clients choose to end psychotherapy prematurely for various reasons, including dissatisfaction with the treatment or the therapist and the perception of psychotherapy as being unhelpful (Westmacott & Hunsley, 2010). Evidence also suggests that approximately 5% to 10% of clients deteriorate during treatment (Lambert & Ogles, 2004). Compounding these problems are clinicians' tendencies to overestimate clients' progress as well as their difficulties in predicting client drop-out (Hannan et al., 2005; Hatfield et al., 2010). Studies have shown that, when relying solely on their clinical judgment, clinicians seem to lack the ability to detect the deterioration of their clients rapidly and accurately, and instead they tend to estimate that their clients are progressing at a rate that is without basis in the literature (Hansen et al., 2002; Walfish et al., 2012). In light of these issues, the current study provides clinicians with information on psychometrically sound instruments, referred to as progress monitoring measures, that can be integrated into clinical practice to provide systematic feedback to clinicians about treatment responses.

The Canadian Psychological Association (CPA, 2018; Tasca et al., 2019) recommended the implementation of progress monitoring measures in clinical contexts to supplement clinical judgment. These measures are typically self-reported, atheoretical brief psychometric assessments of a client's "vital signs" of psychological functioning (Overington & Ionita, 2012). They are meant to be administered at regular intervals or at each session (CPA, 2018). They centre mainly on three aspects of a client's mental health: symptoms, well-being, and functioning (Overington & Ionita, 2012). The continuous administration of progress monitoring measures allows clinicians to monitor treatment progress closely by comparing clients' current and previous levels of psychological functioning (CPA, 2018; Tasca et al., 2019). As such, progress monitoring provides clinicians with information on outcome changes, alerts them to deviations from expected

responses to treatment, and allows them to adjust treatment interventions if necessary (Lambert & Shimokawa, 2011; Overington & Ionita, 2012). Research has demonstrated that progress monitoring decreases the chances of negative results and increases the effectiveness of therapy (Lambert, 2007; Lambert et al., 2005; Lambert et al., 2018). Evaluating progress can also help clients become more aware of their symptoms and of their progress (Lambert & Harmon, 2018), and clinicians gain information regarding their own effectiveness (Muir et al., 2019).

Despite the proven effectiveness of progress monitoring measures in providing valid information to clinicians during therapy, few professionals report using them in their clinical practice (Hatfield & Ogles, 2004; Ionita & Fitzpatrick, 2014). Ionita et al.'s (2020) research on barriers to the use of progress monitoring measures indicated that many Canadian clinicians report not using these measures because of their limited knowledge and understanding of the differences between them. Clinicians in Ionita et al.'s (2020) research also reported concerns about burdening their clients, additional work and time, and the cost of these measures. Hence, the main objective of the current study is to explore and compare various progress monitoring measures in order to help clinicians choose the one that may best suit their clinical practice needs.

Common Progress Monitoring Measures

Among the various progress monitoring measures available for use with adult clients, the *Partners for Change Outcome Management System* (PCOMS; Duncan, 2012; Miller et al., 2005) and the *Outcome Questionnaire-45* (OQ-45; Lambert et al., 1996) are the most highly researched in regard to effectiveness (Lambert & Harmon, 2018). The PCOMS is composed of two subscales: the *Outcome Rating Scale* (ORS) and the *Session Rating Scale* (SRS). Prior studies have shown that the ORS is an accurate measure of psychological distress and that the SRS has moderate concurrent validity with measures of the therapeutic alliance (Duncan, 2012). The OQ-45, for its part, provides an index of mental health functioning (Lambert & Harmon, 2018). As such, the PCOMS focuses on increasing collaboration in the therapeutic relationship (in particular through the SRS), whereas the OQ-45 focuses on predicting clients who are at risk of deteriorating in therapy (Lambert & Harmon, 2018). The OQ-45 and its French translation, along with the ORS, have been shown to be associated positively with measures of psychological distress and associated negatively with measures of psychological well-being (Brosseau-Liard et al., 2020; Campbell & Hemsley, 2009; Lambert et al., 1996).

There are also several other progress monitoring measures that have been studied less, such as the *Behavior and Symptom Identification Scale* (BASIS-24; Eisen et al., 1994), the *Clinical Outcomes in Routine Evaluation-Outcome Measure* (CORE-OM; Evans et al., 2002), and the *Treatment Outcome Package* (TOP; Kraus et al., 2005). The BASIS-24 is a brief measure of psychological distress and functioning

that was demonstrated to be sensitive to change (Cameron et al., 2007). For its part, the CORE-OM assesses subjective well-being, problems/symptoms, life functioning, and risk to self and to others, and it is able to differentiate individuals with clinical and non-clinical symptoms (Evans et al., 2002). Finally, the TOP is a multi-dimensional outcome measure designed to assess a broad range of clients' difficulties and resources (Boswell et al., 2015; Youn et al., 2012). It can distinguish between clinical and non-clinical samples and has good convergent validity with similar measures of psychological distress (Kraus et al., 2005).

Progress Monitoring Measure Selection

Given the wide variety of measures available to monitor clients' progress, selecting one can be challenging for clinicians, especially since they can all be used regardless of a client's symptoms and the type of psychotherapy followed (Drapeau, 2012). As mentioned by Wampold (2015, p. 460), "They all come highly recommended and seem to have more than adequate credentials." Yet, they also differ from one another in many ways (see Table 1 for a summary of the key characteristics of the five progress monitoring measures reviewed). Some measures assess the client's psychological functioning without focusing directly on symptoms of specific disorders (e.g., the ORS; Campbell & Hemsley, 2009), whereas the items of others were derived from specific symptoms to measure the full spectrum of pathology (e.g., the TOP; Kraus et al., 2005).

Progress monitoring measures also have different lengths and degrees of complexity (see Table 1). For instance, the ORS subscale of the PCOMS is an extremely brief, four-item visual analogue outcome measure (Campbell & Hemsley, 2009). It can be completed in about 1 minute and provides a general score indicating the client's perceived functioning and well-being. The TOP, for its part, consists of 58 items assessing 12 clinical and functional domains (Boswell et al., 2015). Both shorter and longer instruments have advantages and drawbacks, which makes it difficult to determine the appropriate length of a progress monitoring measure. On the one hand, the length of administration and methodological complexity may discourage psychotherapists from using some measures as frequently and routinely (Duncan & Reese, 2013). Indeed, many clinicians report being concerned about time requirements and burdening clients (Ionita et al., 2020; Ionita et al., 2016). On the other hand, longer instruments may provide greater information and measure clients' psychological functioning more adequately (Duncan & Reese, 2013).

Yet, further empirical investigation is required to determine whether the differences between each measure's ability to assess mental health are significant (Duncan & Reese, 2013; Halstead et al., 2013). Even though all progress monitoring measures previously mentioned have demonstrated their ability to provide a global rating of mental health functioning, to the best of our knowledge there has not been a study that offers a direct comparison of these measures. Testing

Table 1
Key Characteristics of the Five Selected Progress Monitoring Measures

Progress monitoring measure	Time requirements	Administration	Scoring	Price	Clientele (alternate versions available)
BASIS-24[1,2,3,4] *https://www.ebasis.org*	5–15 minutes (24 items)	Self-report paper data collection or administered in an interview (face-to-face or telephone)	Hand-scored weighted sum; optional online scoring and reporting application via WebScore	USD$395 (CAD$525.85) for an annual single-site licence and USD$95 (CAD$126.47) for additional site licences; optional purchase of WebScore	Adolescents (pilot version)
CORE-OM[1,2,5,6] *https://www.coreims.co.uk*	5–10 minutes (34 items)	Self-report paper data collection in service setting	Hand-scored; optional automated reports from the CORE Information Management Systems (CORE-IMS)	Paper version free of charge; annual licence available for additional services (costs depend on usage); trial packs available for £15 (CAD$25.81)	Youth and adolescents; people with learning difficulties
OQ-45[1,2,7,8] *https://www.oqmeasures.com*	5–10 minutes (45 items)	Ideally administered using a self-report electronic version (a paper-and-pencil form is also available)	Online scoring and automated real-time clinical reports from the OQ-Analyst software	USD$200 (CAD$266.25) annually for the OQ-Analyst licence	Groups; youth and adolescents

Progress monitoring measure	Time requirements	Administration	Scoring	Price	Clientele (alternate versions available)
ORS[1,2,9,10] *https://www.fit-outcomes.com*	1 minute (4 items)	Self-report paper or web-based data collection prior to session	Hand-scored; optional automated reports from FIT-Outcomes when using the electronic version	None for paper version; €134 (CAD$209.02) or more annually for a single-user licence of the outcome management system (FIT-Outcomes)	Young children; children; adolescents; adults
TOP[1,2,11] *http://www.outcomereferrals.com*	5–15 minutes (58 items)	Self-report paper or web-based data collection, in service setting or at home	Scoring and real-time clinical reports via https://wellnesscheck.net/	Free basic service through https://wellnesscheck.net/; additional costs depend on additional services	Children; adolescents; people with substance abuse issues

Note. See the following references: [1]Drapeau (2012), [2]Overington & Ionita (2012), [3]McLean Hospital (2006), [4]McLean Hospital (2019), [5]Evans et al. (2002), [6]CORE IMS (n.d.), [7]Lambert (2015), [8]OQ Measures (n.d.), [9]Duncan (2012), [10]FIT-Outcomes (n.d.), [11]Outcome Referrals (2020).

which measure is best able to assess mental health is an important next step in the literature, as each one of these measures was designed to provide a global rating of mental health and to help clinicians monitor a client's functioning. By documenting each measure's level of convergent validity with a global index of mental health functioning, the current study will hopefully provide expanded resources for clinicians seeking to implement progress monitoring in clinical practice.

The Current Study

Consequently, the main purpose of the present study is to assist clinicians in making informed choices about measure selection, and it will meet this goal by evaluating and comparing five different measures. Given that all progress monitoring measures are presumed to provide clinicians with a client's "vital signs" of psychological functioning, the first aim of this study was to evaluate whether the ORS, the OQ-45, the BASIS-24, the CORE-OM, and the TOP are in fact measuring the same construct. In accordance with the progress monitoring literature, we predicted good convergent validity between the five measures. The second aim of the current study was to investigate the overall ability of the five measures to assess clients' mental health. It is important to document the relationship between progress monitoring measures and mental health in order to address the concerns that some clinicians express about the usefulness and validity of these measures (Hatfield & Ogles, 2007). For this study, we defined overall mental health as being composed of both psychological well-being (i.e., life satisfaction, quality of life, and self-esteem) and psychological distress (i.e., depression, anxiety, anger, and cognitive disturbance). Accordingly, we expected progress monitoring measures to evaluate these two components of mental health accurately. Finally, the third and main goal of this study was to determine which measure assesses clients' overall mental health best. Although to the best of our knowledge this is the first study to compare progress monitoring measures psychometrically using data from a unique sample, previous research suggests that longer measures such as the TOP and the OQ-45 are likely to be better at assessing mental health (Duncan & Reese, 2013; Halstead et al., 2013).

Method

Participants

The original study sample comprised 53 French-speaking Canadian individuals starting psychotherapy. Five participants failed to complete all measures and were thus excluded from the analyses. The final sample consisted of 48 individuals, 17% of whom were men and 83% of whom were women. Participants were required to be at least 17 years old (M = 26.79, SD = 9.91, range = 17–64), to be able to read French, and to be in the process of beginning individual psychotherapy. They did not need to be starting psychotherapy for the first time, however: we

included individuals who were starting therapy again after a prolonged hiatus from their previous time in psychotherapy. We accepted all participants involved in psychotherapy regardless of their diagnoses or the reasons for which they were receiving treatment. This was done to be able to reflect the reality of problems seen in psychotherapy and a variety of baseline levels of mental health. We also accepted participants who were receiving psychotherapy from psychologists (95%) or from social workers (5%).

Participants were mostly Canadian (80%) or French (14%). The highest level of education most commonly completed was high school (47%), followed by a bachelor's degree (33%), a college certificate (12%), and a master's degree (8%). The vast majority of participants were students (56%) or part-time or full-time workers (30%). The median annual income ranged from $10,000 to $19,999.

Procedure

Because participants were recruited through partnerships with psychology services, the research ethics boards of the university and of the mental health centres at which the research was conducted reviewed and approved the research proposal. Recruitment through these partnerships consisted mainly of clinicians presenting information about the study and handing an informational pamphlet to clients who were attending their first session of psychotherapy. Interested participants were instructed to call or to send an email to the research team associated with this project. After an initial contact from interested participants, a research assistant asked follow-up questions to determine their eligibility to participate in the study. Other recruitment strategies included the use of posters and informational pamphlets placed around the university campus, in the waiting rooms of mental health centres, and in public areas in the city, in addition to informational Facebook ads and information booths held on the university campus.

After providing informed consent, participants were asked to complete several questionnaires, starting with the CORE-OM, the BASIS-24, the OQ-45, the ORS, and the TOP, followed by measures of psychological distress and psychological well-being. Finally, they reported on multiple demographic variables. French translations of all questionnaires were used in this study. Participants could choose to answer the questionnaires in a pencil-and-paper format or an online format through Survey Monkey. Participants received $20 in exchange for their participation in this study.

Participants took part in the study after their first (78%) or second (22%) session of psychotherapy. A multivariate analysis was conducted to assess differences in variables of interest between clients who participated after their first session of therapy and those who participated after their second session. Results of the MANOVA revealed no significant differences between the two groups on life satisfaction, quality of life, self-esteem, psychological distress, or progress monitoring measures, $F(9, 35) = .50$, $p = .87$.

Measures

Demographic Questionnaire

Participants were asked to report their age, their gender, their level of education, their annual salary, their nationality, and the date of their first session of psychotherapy. They also provided information regarding their marital status and whether or not they had children.

Progress Monitoring Measures

Although some of the following progress monitoring measures include subscales, scores were computed for overall scales only. Moreover, despite the availability of computerized scoring systems for some of the measures, results were all scored by hand for consistency reasons. The TOP and the CORE-OM were translated from English to French, due to the fact that a French version of these measures was not available when this research project was being conceptualized.[1] To achieve linguistic and conceptual equivalence, the original English versions of these two measures were translated into French and back-translated into English by senior and junior researchers. The translation and back-translation process was followed by a discussion of the resulting discrepancies, during which the items were finalized.

The Clinical Outcomes in Routine Evaluation-Outcome Measure. The CORE-OM (Evans et al., 2002) is a 34-item measure that assesses a client's mental state during the previous week by covering four domains: subjective well-being, problems/symptoms, life functioning, and risk to self and others. Each item is measured on a scale from 0 ("Not at all") to 4 ("Most or all the time"). The total score is calculated by the average of all responses after reverse coding certain scores when necessary. The higher the total score, the worse a person's mental state. The psychometric properties of the English-language version of this questionnaire are adequate. The internal consistency varied between .75 and .94 for all subscales for a clinical sample (Evans et al., 2002). For our French translation, the Cronbach alphas were between .68 and .89 for all subscales and .94 for the global scale.

The Behavior and Symptom Identification Scale. The BASIS-24 (Eisen et al., 1994; McLean Hospital, 2006) contains 24 items and six subscales to evaluate a client's mental state during the previous week. The six subscales are depression/functioning, problems in interpersonal relationships, self-harm, emotional liability, psychotic symptoms, and substance abuse. Items are measured on a 5-point scale. Response options reflect either the degree of difficulty the participant has been experiencing (0 = "No difficulty" and 4 = "Extreme difficulty") or the frequency at which the participant experienced a symptom/problem (0 = "None of

1 A French translation of the CORE-OM is now available (Clinical Outcomes in Routine Evaluation [and CST], 2020). The psychometric exploration of the translation is in progress. It should be noted that our French translation of the CORE-OM is broadly similar to the version recently published by the CORE System Trust.

the time"/"Never" to 4 = "All of the time"/"Always"). The total score, which varies from 0 to 4, is a weighted sum calculated by multiplying the score for each item by its weight (available in the BASIS-24 Instruction Guide; McLean Hospital, 2006) and totalling the weighted ratings for all items. The higher the score, the more difficulty or distress the person is reporting in general. The internal consistency for the BASIS-24 overall score for in-patients and outpatients from different ethnic groups was reported to vary between .87 and .91 (Eisen et al., 2006). To the best of our knowledge, there are no psychometric data for the French translation of the BASIS-24. The Cronbach alpha for this study was .85.

The Outcome Questionnaire-45. The OQ-45 (Flynn et al., 2002; Lambert & Bergin, 1994) comprises 45 items and three subscales, including symptom distress, interpersonal relations, and social role. Items are scored on a scale from 0 ("Never") to 4 ("Almost always"). Certain questions have been reframed due to grammatical errors in the existing French translation of the OQ-45. The total score represents the sum of all responses and was calculated after reverse coding appropriate items. The higher the score, the more the individual is experiencing distress. The French version of the OQ-45 (Mesure d'impact; MI-45) has an internal consistency of .91 for symptom distress, .81 for interpersonal relations, and .65 for social role for a clinical population (Brosseau-Liard et al., 2020). In the current study, the Cronbach alphas were .93 for symptom distress, .80 for interpersonal relations, .65 for social role, and .94 for the global scale.

The Outcome Rating Scale and the Session Rating Scale. The ORS and the SRS are the two measures comprised in the PCOMS (Duncan, 2012; Miller et al., 2005). The ORS assesses a client's perceived level of global distress and functioning, whereas the SRS assesses elements of the therapeutic alliance. Since the purpose of the current study is to evaluate the ability of progress monitoring measures to assess mental health, only the data provided by the ORS were analyzed.

Comprised of 4 items, the ORS is a brief measure that can be answered and scored in under a minute. It asks individuals to evaluate how they have been in the previous week or since their last session regarding their individual, interpersonal, social, and general well-being. Respondents indicate their answer by tracing a line on four 10 cm visual scales, each corresponding to one item. The score is calculated by measuring to the nearest millimetre where the line was traced on each of the four visual scales. The total score is the sum of all 4 items. For the purpose of the current study, the total score was reversed so that a high score indicates more difficulty or distress. The internal consistency of the ORS was .85 for a clinical population (Duncan, 2012). To the best of our knowledge, there is no psychometric data for the French translation of the ORS. The scale's internal reliability for the current sample was .75.

The Treatment Outcome Package. The *TOP* (Boswell et al., 2015; Kraus et al., 2005) is composed of 58 items evaluating 12 subscales: work functioning,

sexual functioning, social functioning, depression, panic, psychosis, suicidal ideation, violence, mania, sleep, substance abuse, and quality of life. Participants are asked to answer how many times during the previous 2 weeks they have engaged in or agreed with each item. Items are scored on a scale from 1 ("All") to 6 ("None"). The total score is represented by the sum of responses to all items and was reversed for the purpose of the current study. More precisely, higher scores indicate more problematic functioning and more clinical symptoms. The internal consistency varied from .53 to .93 for all subscales (Kraus et al., 2005). The Cronbach alpha for our translated version of the questionnaire was .93 for the global scale.

Mental Health Measures

Psychological Distress. The *Psychiatric Symptom Index* was developed by Ilfeld (1976) and is designed to evaluate symptoms of depression, anxiety, anger, and cognitive disturbance. Although specific psychiatric disorders are not measured by the Psychiatric Symptom Index, its items screen for a broad range of symptoms and its continued use in research has been previously supported (Okun et al., 1996). Moreover, the French version of this instrument has been used successfully in several other studies and in previous health surveys conducted by Santé Québec (Boyer & Villa, 2011). The index includes 29 items that are evaluated on a 4-point scale from 0 ("Not at all") to 3 ("Very often"). All items were totalled to produce an overall score, with scores higher than 20 indicating high levels of psychiatric symptoms, scores from 10 to 19 indicating moderate levels of psychiatric symptoms, and scores from 0 to 9 indicating low levels of psychiatric symptoms. The scale has an internal consistency of .91 (Ilfeld, 1976). The French translation of this measure was reported to have an internal consistency of .89 (Kovess et al., 1985). Reliability analyses indicated an alpha of .93 for our sample.

Psychological Well-Being. The *Quality of Life Index* (Ferrans & Powers, 1985) is comprised of two 18-item questionnaires that evaluate respondents' perceptions of their quality of life. More precisely, this measure assesses quality of life by evaluating participants' satisfaction in different areas of life as weighted by their importance for the person. Thus, the first questionnaire asks individuals to evaluate their satisfaction with several aspects of their lives and the second questionnaire asks individuals how important they consider each of these aspects in their lives (for example, their satisfaction with their health and the importance of their health). Both questionnaires are scored on a 6-point scale ranging from 1 ("Very dissatisfied"/"Very unimportant") to 6 ("Very satisfied"/"Very important"). This measure is comprised of four subscales representing various domains of life: 1) health and functioning, 2) social and economic, 3) psychological and spiritual, and 4) family.

In order to score the Quality of Life Index, the results from the importance scale were combined with the results from the satisfaction scale. This combination is accomplished by subtracting 3.5 from each item in the satisfaction scale

and multiplying the difference with the corresponding score from the importance scale. To obtain the final overall scores, which range from 0 to 30, 15 is added to the mean of all weighted scores. The higher the overall score, the higher the quality of life. The internal consistency of the scale was .93 for the original measure (Ferrans & Powers, 1985) and .92 for our sample.

The *Rosenberg Self-Esteem Scale* (Rosenberg, 1989; Vallières & Vallerand, 1990) is a 10-item instrument that evaluates self-esteem. In this case, self-esteem is defined as the degree to which an individual considers themselves to be a person of value who has several good qualities. The items are scored on a 4-point scale ranging from 1 ("Strongly disagree") to 4 ("Strongly agree"). The total score of self-esteem is obtained by the sum of responses to all 10 items. The higher the score, the better an individual's self-esteem. The internal consistency of the French version of this scale varied from .70 to .90 (Vallières & Vallerand, 1990). The alpha coefficient for the current study was .88.

The *Satisfaction With Life Scale* (Blais et al., 1989; Diener et al., 1985) is a 5-item scale that estimates individuals' global satisfaction with their lives. Each item is evaluated on a scale ranging from 1 ("Strongly disagree") to 7 ("Strongly agree"). The total score is the sum of all items. A higher score on this instrument indicates higher satisfaction with life. The French version of this scale was previously reported to have an internal consistency of .80, which corresponds to that of the original English version (.87; Blais et al., 1989). In the current study, the instrument demonstrated high internal consistency ($\alpha = .85$).

Results

Preliminary Analyses

Missing values were substituted by a participant's mean score when 80% of the items on the corresponding questionnaire were answered (20 occurrences; 0.2% of all data). Means, standard deviations, and correlations between the five progress monitoring measures and the various measures pertaining to mental health are reported in Table 2. As shown in this table, all five progress monitoring measures are associated in significant ways with measures of psychological well-being and psychological distress. Table 2 also reveals self-reported difficulties in functioning and psychological distress. Indeed, suggested cut-offs between dysfunctional and functional populations place the average progress monitoring score of the current sample in the dysfunctional range (for more information on guidelines on the interpretation of progress monitoring scores, see Barkham et al., 2006; Cameron et al., 2007; Lambert et al., 2013; Miller & Duncan, 2000). Ilfeld's (1976) cut-off points for total psychological distress scores also indicate that, on average, participants reported high levels of psychiatric symptoms (total scores higher than 20). These self-reported difficulties were expected with individuals starting psychotherapy.

Table 2
Correlations and Descriptive Statistics of Key Variables (N = 48)

	1	2	3	4	5	6	7	8	9
1. CORE-OM									
2. BASIS-24	.85								
3. OQ-45	.89	.88							
4. TOP	.71	.75	.72						
5. ORS	.78	.73	.75	.60					
6. Quality of life	-.72	-.69	-.74	-.55	-.75				
7. Life satisfaction	-.68	-.65	-.66	-.57	-.67	.59			
8. Self-esteem	-.72	-.67	-.72	-.66	-.61	.50	.66		
9. Psychological distress	.74	.76	.74	.73	.61	-.78	-.57	-.50	
Range	[0, 4]	[0, 4]	[0, 180]	[58, 348]	[0, 40]	[0, 30]	[5, 35]	[10, 40]	[0, 87]
Mean	1.40	1.46	65.66	132.25	17.94	18.13	21.65	27.84	34.26
Standard deviation	.62	.53	25.00	31.92	6.82	5.13	7.19	5.69	16.03

Note. All correlations are significant ($p < .001$).

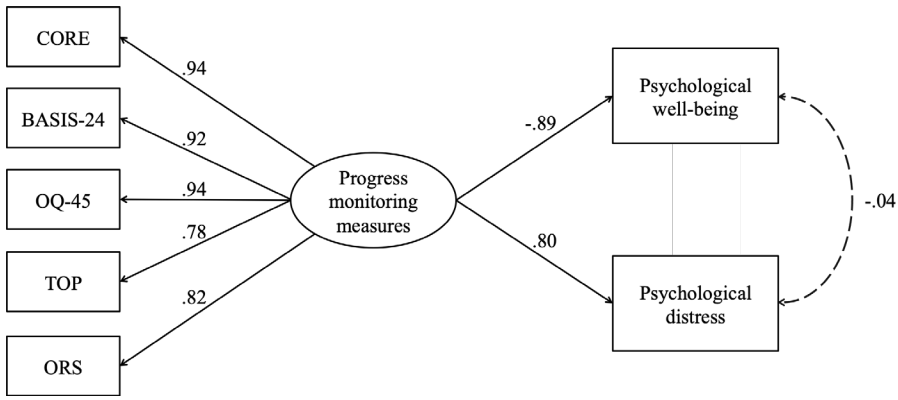
Inferential Analyses

The proposed structural regression model presented in Figure 1 was tested to provide evidence of convergent validity between the five progress monitoring measures (Hypothesis 1) and to evaluate the overall ability of these monitoring measures to assess mental health (Hypothesis 2). Structural regression models enable researchers to test hypotheses about both measurement and structural relations (Kline, 2016). More precisely, the measurement part of our model features the progress monitoring measures' latent construct, composed of five observed variables (i.e., the CORE-OM, the BASIS-24, the OQ-45, the TOP, and the ORS). The structural part of our model represents the hypothesis that progress monitoring measures have the ability to assess a client's psychological well-being and psychological distress. Psychological distress was evaluated using the overall scores of the Psychiatric Symptom Index, whereas psychological well-being was evaluated through a composite score comprised of the results from the Quality of Life Index, the Rosenberg Self-Esteem Scale, and the Satisfaction With Life Scale. To obtain this composite score, we averaged the Z-scores of the three questionnaires (see Maltais et al., 2019, for a similar procedure).

The model was tested with the EQS 6.3 structural equation program using the maximum likelihood robust estimation method. This method is appropriate for samples of 250 participants or fewer (Kline, 2016). To determine whether the overall model fit was acceptable, the following statistics were used: a non-significant Satorra-Bentler chi-square (S-B χ^2), a goodness-of-fit index (GFI) and a robust comparative fit index (CFI) superior to .90, and a value smaller than .08 and .07, respectively, on the standardized root mean square residual (SRMR) and the robust root mean square error of approximation (RMSEA; Hooper et al., 2008). Finally, to ensure the psychometric qualities of the data matrix, data were screened carefully and multicollinearity between variables of interest was assessed by inspecting squared multiple correlations between each variable and all the rest. Squared multiple correlations were all under .90, therefore indicating no problem of multicollinearity (Kline, 2016).

Following Kline's (2016) best practice recommendations for structural equation modelling, the structural regression model was evaluated using two-step modelling. First, we tested the measurement model of the latent construct to ensure that the variance of the five progress monitoring measures was adequately explained by the latent variable. The loading of the CORE-OM indicator was fixed to 1.0 to scale the factor. The measurement model provided a good fit for the data, with all global fit indices indicating acceptability (S-B $\chi^2(5) = 3.30$, $p = .65$; GFI = .98; Robust CFI = 1.00; SRMR = .01; Robust RMSEA = .00). Standardized factor loadings ranged from .77 to .95. Specifically, the progress monitoring factor explained 88% of the variance in the CORE-OM, 85% in the BASIS-24, 90% in the OQ-45, 65% in the ORS, and 60% in the TOP. These

Figure 1
Illustration of the Structural Regression Model of Mental Health



Note. All values represent standardized estimates. Direct effects are represented with single-sided arrow paths and covariances are represented with double-sided arrow paths. A solid line indicates a significant effect ($p < .05$), whereas a dashed line indicates a non-significant effect.

results provided support for Hypothesis 1 by confirming that the five progress monitoring measures are related to one another.

We then examined the structural part of the model to evaluate the hypothesized relationships between the progress monitoring measures, psychological well-being, and psychological distress. Error terms of the outcome variables were permitted to co-vary with one another. The proposed model linking the progress monitoring measures factor to psychological well-being and psychological distress provided an acceptable fit to the data for all global fit indices (S-B $\chi^2(13) = 13.70$, $p = .40$; GFI = .93; Robust CFI = 1.00; SRMR = .03; Robust RMSEA = .03). Moreover, all paths were significant, with the exception of the covariance between the error terms of the outcome variables (see Figure 1). The model accounted for 78% of the variance in psychological well-being and 64% of the variance in psychological distress. Taken together, results of the two-step modelling for our structural regression model confirmed Hypothesis 2. More precisely, the five progress monitoring measures, considered as a whole, assessed mental health accurately.

Further analysis was then conducted with the five progress monitoring measures individually to determine which one assesses clients' mental health best (Hypothesis 3). Specifically, five path models were performed to evaluate the predictive power of the CORE-OM, the BASIS-24, the OQ-45, the TOP, and the ORS. In each of these models, one of the progress monitoring measures was

Table 3
Fit Indices for All Path Models

Predictor variable	S-B χ^2 (p)	GFI	Robust CFI	SRMR	Robust RMSEA	Robust AIC	R^2	
							Well-being	Distress
CORE	4.18 (.04)	.95	.97	.04	.25	2.18	.73	.57
BASIS-24	6.85 (.01)	.94	.93	.05	.33	4.85	.65	.60
OQ-45	3.51 (.06)	.95	.97	.04	.22	1.51	.72	.60
TOP	7.30 (.01)	.90	.92	.07	.35	5.30	.53	.59
ORS	8.20 (.00)	.88	.88	.09	.39	6.20	.61	.37

Note. S-B χ^2 = Satorra-Bentler chi-square; GFI = goodness of fit index; CFI = comparative fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; AIC = Akaike information criterion.

entered as a predictive variable and psychological well-being and psychological distress were entered as the two outcome variables. Because the covariance between the error terms of the outcome variables was non-significant in the structural regression model, it was not included in the path models. Results of all path models are presented in Table 3. As shown in this table, RMSEA values for the five path models did not indicate acceptability, as they were all over .07. Yet, these results are not surprising given that the RMSEA tends to favour larger models and to penalize smaller models with few variables like those examined in the current study (Kline, 2016). Table 3 also reveals that all global fits, with the exception of the RMSEA, were considered to be acceptable for only one path model (i.e., the OQ-45).

Model fits of the five competing path models were compared using recommendations for Akaike information criterion (AIC; Burnham & Anderson, 2004). AIC is a predictive fit index combining statistical estimation and model selection in a single framework and can be used to compare and rank non-nested models (Burnham & Anderson, 2004; Kline, 2016). Specifically, lower AIC values represent better approximating models. In the current study, the OQ-45 model has the lowest AIC value, followed by the CORE-OM, the BASIS-24, the TOP, and the ORS. Moreover, differences in AIC (Δi) between two models can be interpreted as follows: models having $\Delta i < 2$ are similar in terms of approximating abilities, those in which $4 < \Delta i < 7$ have less similarity, and models having $\Delta i > 10$ have no similarity (Burnham & Anderson, 2004). Although the OQ-45 model has the lowest AIC value, its ability to assess mental health is similar to the CORE-OM model's ability. The BASIS-24 and the TOP's abilities to assess mental health are also somewhat similar to those of the OQ-45 and the CORE-OM, given that the differences in AIC values for these sets of models are between 2 and 4. Finally, both the OQ-45 and the CORE-OM are better than the ORS at assessing a client's mental health ($4 < \Delta i < 7$). These results partially confirmed Hypothesis 3, which suggested that longer measures would be better than shorter measures at assessing mental health.

Discussion

A growing body of research indicates that clinicians tend to report having limited knowledge about progress monitoring measures (Ionita et al., 2020; Tasca et al., 2019) and being concerned about both time requirements associated with monitoring progress and the validity of shorter measures (Ionita et al., 2016). The current study addressed these two ongoing challenges by giving new insights on how the ORS, the OQ-45, the BASIS-24, the CORE-OM, and the TOP differ from each other psychometrically. Results of the structural regression model show strong convergent validity between the five progress monitoring measures and demonstrate that the measures, considered as a whole, assess the two components

of mental health (i.e., psychological well-being and psychological distress) accurately. Our present findings also suggest that the OQ-45 and the CORE-OM are better than the BASIS-24, the TOP, and more importantly the ORS at assessing a client's mental health.

In order to help clinicians overcome their knowledge barrier, it is of paramount importance that they be provided with more information on progress monitoring measures, how they work, and how they differ from one another. In this regard, it is important to remember that despite their differences, all measures were designed to provide clinicians with a global rating of clients' mental health and thus of information on outcome changes. The presumption that the ORS, the OQ-45, the BASIS-24, the CORE-OM, and the TOP all measure the same construct was supported by the analysis of the measurement model of our structural regression model. According to Kline (2016), a measurement model is reasonably correct if all indicators have factor loadings higher than .70 on the specified factor. In the current study, standardized factor loadings range from .77 to .95, therefore indicating good convergent validity. Consistent with Hypothesis 1, these results confirm that the five progress monitoring measures that were theoretically designed to evaluate a common construct are in fact related.

As predicted by Hypothesis 2, the results of our structural regression model also provide evidence of the overall ability of the five progress monitoring measures to assess mental health accurately. These findings are in line with those of previous studies (see Drapeau, 2012, for a summary) indicating that progress monitoring measures correlate with global ratings of mental health and other well-known instruments used to measure key aspects of mental health (e.g., the Beck Depression Inventory, the Quality of Life Scale, and the Rosenberg Self-Esteem Scale). The documentation on the strong relationship between progress monitoring measures and mental health is important as a way to help ease the concern that some clinicians express about the usefulness of progress monitoring measures (Hatfield & Ogles, 2007). In this regard, Ionita and Fitzpatrick (2014) indicated that a number of clinicians report using instruments developed to measure specific aspects of mental health (e.g., the Depression Anxiety Stress Scale, the Symptom Checklist-90, and Beck Scales and Inventories) as a way to gather information about outcome changes instead of using measures that were specifically designed to monitor clients' progress. Our results indicate that progress monitoring measures possess the ability to assess accurately the underlying construct that they are intended to measure and thus to support their use by clinicians in clinical practice.

While the results of the structural regression model lend further support that the ORS, the OQ-45, the BASIS-24, the CORE-OM, and the TOP have much in common, these five measures differ importantly in terms of length and complexity. Researchers have argued that such distinctions may prevent clinicians from considering the use of some measures in practice, given that many

of them report being concerned about additional time requirements and about burdening their clients (Ionita et al., 2020). Yet, other researchers and clinicians question the reliability and validity of shorter measures (Halstead et al., 2013; Ionita et al., 2016). Our study makes a substantial contribution to the literature of progress monitoring by comparing the level of convergent validity between the five measures and mental health, in order to determine whether the characteristics that differentiate them are important psychometrically and thus by guiding clinicians toward the selection of one measure over another. Results of the five competing path models revealed that the OQ-45 and the CORE-OM are better than the other three measures at assessing clients' mental health. Although the differences between the abilities of the OQ-45 and the CORE-OM and between the BASIS-24 and the TOP are small, the differences demonstrated between the OQ-45, the CORE-OM, and the ORS seem particularly noteworthy. In fact, our results suggest that the ability of the ORS to assess mental health is not as accurate as the abilities of the OQ-45 and the CORE-OM. The differences between these measures appear to result from the fact that the ORS lacks the ability to assess psychological distress accurately (see Table 3). Yet, collecting information regarding both a client's psychological well-being and psychological distress likely increases the chances of clinicians detecting warning signs of potential problems in the therapeutic process.

Taken together, these findings support Halstead et al.'s (2013) argument that longer measures tend to be better than shorter measures at providing outcome information. Nevertheless, we concluded that Hypothesis 3 was only partially supported because the TOP, which is the longest of the five measures, is not among the ones with the best abilities to assess mental health. Although contrary to our hypothesis, these results are in line with Wampold's (2015) argument that longer instruments do not always provide greater information, especially when their factor structures are complex and difficult to replicate.

Implications

Our study supports the overall use of progress monitoring measures to assess mental health. Nevertheless, we acknowledge that the search for a measure that can be completed quickly while also being sufficiently informative to meet a clinician's monitoring needs can be difficult. In this regard, several studies underlined the need for researchers to document the differences between the wide variety of measures available in order to help clinicians make informed decisions and weigh the costs and benefits of choosing one measure over another. According to Ionita and Fitzpatrick's (2014) study, the most commonly used measure to monitor clients' progress by Canadian clinicians in 2012 was the ORS. One of the main benefits of the ORS is the fact that it involves only four items that are quick to answer. Yet, data from the current study reveal that using the shortest measure can have some drawbacks, given that the ability of the ORS to assess

mental health and more precisely psychological distress is not as accurate as the ability of other available measures. While these findings may seem a bit daunting to clinicians who are concerned about time requirements, our results also suggest that longer measures are not necessarily more valid and informative than shorter measures. For instance, the OQ-45 and the CORE-OM, composed of 45 and 34 items, respectively, seem to be better at assessing clients' mental health than the TOP, which consists of 58 items. Another example is the similarity in terms of approximating abilities of the TOP and the ORS. Such knowledge is important to help clinicians overcome their knowledge barrier and consequently improve the services they offer clients in psychotherapy.

Prior research found that users often learned about monitoring progress practices from journals, whereas non-users reported that workshops were their preferred source of practice knowledge (Ionita et al., 2020). In order to ensure that all clinicians are well-informed about progress monitoring practices, workshops and training programs should serve as opportunities to increase clinicians' awareness and knowledge by disseminating information about differences between existing measures. We believe that programs developed to promote evidence-based practice can help clinicians provide the best services to their clients by deepening their understanding of the pros and cons of each measure.

Limitations and Future Directions

The data analyzed in the current study were all collected at a single point in time and thus limit any possible conclusions about the directionality of the significant links in our models. Future research should collect longitudinal data to explore further the links between our variables of interest and to compare other important psychometric properties that could help guide decision-making for the selection of one measure over another (e.g., reliability to detect small changes over the course of psychotherapy). Although our study provides initial data about psychometric differences of progress monitoring measures, many questions and areas for investigation remain to be addressed to provide the guidance needed by clinicians to select reliable and valid measures that suit their needs. We hope that the current study will serve as a catalyst for additional research and that future studies will continue to build the evidence needed on progress monitoring measures.

The use of French versions of the progress monitoring measures may be both a strength and a limitation of the current study. On the one hand, translations do not have exactly the same psychometric properties as the original measures (Drapeau, 2012). Our findings may therefore differ in other languages, which limits the generalizability of our results. On the other hand, this study contributes to the progress monitoring literature by examining the ability of the French translations to assess mental health, given that the psychometric properties of the French-language versions remained understudied. This may explain, in part,

why French-speaking clinicians are less likely to report using progress monitoring measures than their English-speaking colleagues (Ionita et al., 2020; Ionita & Fitzpatrick, 2014). To encourage clinicians to engage in the use of these measures regardless of their language practices, it is important to offer translated versions of the measures and to conduct research on the translations. Future research with larger samples should explore further whether the French translations of the CORE-OM and the TOP are equivalent to the original English-language versions of the measures using confirmatory factor analyses. Unfortunately, the current study did not meet the sample size requirements to conduct such analyses.

Conclusion

Engaging in continuous progress monitoring continues to be challenging for many clinicians due to limited knowledge about the various measures available. The current study contributes to an unexplored area of research by addressing the need for detailed comparisons of the differences between progress monitoring measures. Overall, our results indicated good convergent validity between the five measures and mental health. Nevertheless, important differences in the measures' abilities to evaluate the two components of mental health were identified. Data reported in this study suggest that, among the five measures, the OQ-45 and the CORE-OM's abilities to assess mental health are the most accurate, whereas the ORS seems to have difficulty in assessing psychological distress accurately. We hope that these results will help reduce the practice–research gap that continues to exist in psychotherapy settings by providing initial guidance regarding the selection of a measure to those seeking to implement progress monitoring in their practice.

References

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, *6*(1), 3–15. https://www.doi.org/10.1080/14733140600581218

Blais, M. R., Vallerand, R. J., Pelletier, L. G., & Brière, N. M. (1989). L'échelle de satisfaction de vie : validation canadienne-française du « Satisfaction With Life Scale ». *Canadian Journal of Behavioural Science*, *21*(2), 210–223. https://doi.org/10.1037/h0079854

Boswell, J. F., Kraus, D. R., Castonguay, L. G., & Youn, S. J. (2015). Treatment outcome package: Measuring and facilitating multidimensional change. *Psychotherapy*, *52*(4), 422–431. https://doi.org/10.1037/pst0000028

Boyer, R., & Villa, J. (2011). *Répertoire d'instruments pour la surveillance des impacts psychosociaux des aléas climatiques*. Institut national de santé publique. https://www.inspq.qc.ca/pdf/publications/1329_RepertInstruSurvImpactsPsychoSocAleasClim.pdf

Brosseau-Liard, P. É., Vandette, M.-P., Jamshidi, P., Kogan, C. S., & Aubry, T. (2020). Propriétés psychométriques de la version française (Mesure d'impact; MI-45) du Outcome

Questionnaire-45 (OQ-45) en milieu clinique et universitaire. *Canadian Journal of Behavioural Science*, *52*(1), 78–83. https://doi.org/10.1037/cbs0000138

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Cameron, I. M., Cunningham, L., Crawford, J. R., Eagles, J. M., Eisen, S. V., Lawton, K., Naji, S. A., & Hamilton, R. J. (2007). Psychometric properties of the BASIS-24© (Behaviour and Symptom Identification Scale–Revised) mental health outcome measure. *International Journal of Psychiatry in Clinical Practice*, *11*(1), 36–43. https://doi.org/10.1080/13651500600885531

Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, *13*(1), 1–9. https://doi.org/10.1080/13284200802676391

Canadian Psychological Association. (2018). *Outcomes and progress monitoring in psychotherapy*. https://cpa.ca/docs/File/Task_Forces/Treatment%20Progress%20and%20Outcome%20Monitoring%20Task%20Force%20Report_Final.pdf

Clinical Outcomes in Routine Evaluation (and CST). (2020). *French*. https://www.coresystemtrust.org.uk/translations/french/

CORE IMS. (n.d.). *CORE IMS*. https://www.coreims.co.uk

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Drapeau, M. (Ed.). (2012). 10 outils pour évaluer l'efficacité des interventions en psychothérapie [Special issue]. *Cahier recherche et pratique*, *2*(2). https://www.ordrepsy.qc.ca/c/document_library/get_file?uuid=41f6de0d-f187-487f-87d6-e17384f1181b&groupId=26707

Duncan, B. L. (2012). The Partners for Change Outcome Management System (PCOMS): The Heart and Soul of Change Project. *Canadian Psychology*, *53*(2), 93–104. https://doi.org/10.1037/a0027762

Duncan, B. L., & Reese, R. J. (2013). Clinical and scientific considerations in progress monitoring: When is a measure too long? *Canadian Psychology*, *54*(2), 135–137. https://doi.org/10.1037/a0032362

Eisen, S. V., Dill, D. L., & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry*, *45*(3), 242–247. https://doi.org/10.1176/ps.45.3.242

Eisen, S. V., Gerena, M., Ranganathan, G., Esch, D., & Idiculla, T. (2006). Reliability and validity of the BASIS-24© mental health survey for Whites, African-Americans, and Latinos. *Journal of Behavioral Health Services and Research*, *33*(3), Article 304. https://doi.org/10.1007/s11414-006-9025-3

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, *180*(1), 51–60. https://doi.org/10.1192/bjp.180.1.51

Ferrans, C. E., & Powers, M. J. (1985). Quality of life index: Development and psychometric properties. *Advances in Nursing Science*, *8*(1), 15–24. https://doi.org/10.1097/00012272-198510000-00005

FIT-Outcomes. (n.d.). *FIT-Outcomes: Feedback informed treatment*. https://www.fit-outcomes.com

Flynn, R. J., Aubry, T. D., Guindon, S., Tardif, I., Viau, M., & Gallant, A. (2002). Validation d'une version française du *Outcome Questionnaire* et évaluation d'un service de counselling en milieu clinique [Validity of a French version of the Outcome Questionnaire and evaluation of a counselling program in a clinical setting]. *Canadian Journal of Program Evaluation*, *17*(3), 57–74. https://evaluationcanada.ca/system/files/cjpe-entries/17-3-057.pdf

Halstead, J., Youn, S. J., & Armijo, I. (2013). Scientific and clinical considerations in progress monitoring: When is a brief measure too brief? *Canadian Psychology*, *54*(1), 83–85. https://doi.org/10.1037/a0031324

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, *61*(2), 155–163. https://doi.org/10.1002/jclp.20108

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, *9*(3), 329–343. https://doi.org/10.1093/clipsy/9.3.329

Hatfield, D., McCullough, L., Frantz, S. H. B., & Krieger, K. (2010). Do we know when our clients get worse? An investigation of therapists' ability to detect negative client change. *Clinical Psychology and Psychotherapy*, *17*(1), 25–32. https://doi.org/10.1002/cpp.656

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice*, *35*(5), 485–491. https://doi.org/10.1037/0735-7028.35.5.485

Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administration and Policy in Mental Health and Mental Health Services Research*, *34*(3), 283–291. https://doi.org/10.1007/s10488-006-0110-y

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://academic-publishing.org/index.php/ejbrm/article/view/1224

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi.org/10.1037/0003-066X.41.2.159

Ilfeld, F. W., Jr. (1976). Further validation of a psychiatric symptom index in a normal population. *Psychological Reports*, *39*(3, Suppl.), S1215–S1228. https://doi.org/10.2466/pr0.1976.39.3f.1215

Ionita, G., Ciquier, G., & Fitzpatrick, M. (2020). Barriers and facilitators to the use of progress-monitoring measures in psychotherapy. *Canadian Psychology*, *61*(3), 245–256. https://doi.org/10.1037/cap0000205

Ionita, G., & Fitzpatrick, M. (2014). Bringing science to clinical practice: A Canadian survey of psychological practice and usage of progress monitoring measures. *Canadian Psychology*, *55*(3), 187–196. https://doi.org/10.1037/a0037355

Ionita, G., Fitzpatrick, M., Tomaro, J., Chen, V. V., & Overington, L. (2016). Challenges of using progress monitoring measures: Insights from practicing clinicians. *Journal of Counseling Psychology*, *63*(2), 173–182. https://doi.org/10.1037/cou0000122

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kovess, V., Murphy, H. B. M., Tousignant, M., & Fournier, L. (1985). *Évaluation de l'état de santé de la population des territoires des DSC de Verdun et de Rimouski*. Unité de recherche psychosociale du Centre hospitalier Douglas, MAS et Hôpital Rivière-des-Prairies.

Kraus, D. R., Seligman, D. A., & Jordan, J. R. (2005). Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The Treatment Outcome Package. *Journal of Clinical Psychology*, *61*(3), 285–314. https://doi.org/10.1002/jclp.20084

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1–14. https://doi.org/10.1080/10503300601032506

Lambert, M. J. (2015). Progress feedback and the OQ-system: The past and the future. *Psychotherapy*, *52*(4), 381–390. https://doi.org/10.1037/pst0000027

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 143–189). John Wiley and Sons.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.2)*. 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. https://doi.org/10.1002/jclp.20113

Lambert, M. J., & Harmon, K. L. (2018). The merits of implementing routine outcome monitoring in clinical practice. *Clinical Psychology: Science and Practice*, *25*(4), Article e12268. https://doi.org/10.1111/cpsp.12268

Lambert, M. J., Kahler, M., Harmon, C., Burlingame, G. M., Shimokawa, K., & White, M. M. (2013). *Administration and scoring manual. Outcome Questionnaire OQ®-45.2* (Rev. ed.). OQ Measures.

Lambert, M. J., & Ogles, B. (2004). The efficacity and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). John Wiley and Sons.

Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy*, *48*(1), 72–79. https://doi.org/10.1037/a0022238

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, *55*(4), 520–537. https://doi.org/10.1037/pst0000167

Maltais, M., Bouchard, G., & Saint-Aubin, J. (2019). Mechanisms of mindfulness: The mediating roles of adaptive and maladaptive cognitive factors. *Current Psychology*, *38*(3), 846–854. https://doi.org/10.1007/s12144-017-9665-x

McLean Hospital. (2006). *BASIS-24® Instruction Guide*.

McLean Hospital. (2019). *eBASIS*. https://www.ebasis.org

Miller, S. D., & Duncan, B. L. (2000). *The Outcome and Session Rating Scales: Administration and scoring manual.* Institute for the Study of Therapeutic Change.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology*, *61*(2), 199–208. https://doi.org/10.1002/jclp.20111

Muir, H. J., Coyne, A. E., Morrison, N. R., Boswell, J. F., & Constantino, M. J. (2019). Ethical implications of routine outcomes monitoring for patients, psychotherapists, and mental health care systems. *Psychotherapy*, *56*(4), 459–469. https://doi.org/10.1037/pst0000246

Okun, A., Stein, R. E. K., Bauman, L. J., & Silver, E. J. (1996). Content validity of the Psychiatric Symptom Index, CES-Depression Scale, and State-Trait Anxiety Inventory from the perspective of DSM-IV. *Psychological Reports*, *79*(3), 1059–1069. https://doi.org/10.2466/pr0.1996.79.3.1059

OQ Measures. (n.d.). *OQ measures: The measure of mental health vital signs*. https://www.oqmeasures.com

Outcome Referrals. (2020). *Outcome referrals*. http://www.outcomereferrals.com

Overington, L., & Ionita, G. (2012). Progress monitoring measures: A brief guide. *Canadian Psychology*, *53*(2), 82–92. https://doi.org/10.1037/a0028017

Rosenberg, M. (1989). *Society and the adolescent self-image* (Rev. ed.). Wesleyan University Press.

Tasca, G. A., Angus, L., Bonli, R., Drapeau, M., Fitzpatrick, M., Hunsley, J., & Knoll, M. (2019). Outcome and progress monitoring in psychotherapy: Report of a Canadian Psychological Association task force. *Canadian Psychology*, *60*(3), 165–177. https://doi.org/10.1037/cap0000181

Vallières, E. F., & Vallerand, R. J. (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi de Rosenberg [French-Canadian translation and validation of Rosenberg's Self-Esteem Scale]. *International Journal of Psychology*, *25*(2), 305–316. https://doi.org/10.1080/00207599008247865

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports*, *110*(2), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Wampold, B. E. (2015). Routine outcome monitoring: Coming of age—with the usual developmental challenges. *Psychotherapy*, *52*(4), 458–462. https://doi.org/10.1037/pst0000037

Westmacott, R., & Hunsley, J. (2010). Reasons for terminating psychotherapy: A general population study. *Journal of Clinical Psychology*, *66*(9), 965–977. https://doi.org/10.1002/jclp.20702

Youn, S. J., Kraus, D. R., & Castonguay, L. G. (2012). The treatment outcome package: Facilitating practice and clinically relevant research. *Psychotherapy*, *49*(2), 115–122. https://doi.org/10.1037/a0027932

About the Authors

Caissie Bourque is a Psy.D. candidate and a sessional lecturer at the Université de Moncton. Her research interests include progress monitoring and outcome measures as well as therapeutic alliance measures.

Andréanne Charbonneau completed her Ph.D. studies in the School of Psychology at the Université de Moncton and is currently working as an academic researcher and a sessional lecturer. Her areas of interest include family and romantic relationships, psychological well-being, and multivariate statistics. https://orcid.org/0000-0002-0128-7678

Geneviève Bouchard is a professor of psychology at the Université de Moncton. Her research interests include abnormal psychology, psychological well-being, couple psychology, and family psychology. She is also a registered psychologist in the province of Quebec. https://orcid.org/0000-0001-7542-704X

Mylène Lachance-Grzela is an associate professor of psychology at the Université de Moncton. Her research interests include interpersonal relationships, work–life balance, and psychological well-being. She is also a registered psychologist in the province of New Brunswick. https://orcid.org/0000-0001-6395-2007

The authors declare that they have no conflict of interest. This research was supported by a grant from the National Health Training Consortium Research Fund (#22761V). The authors would like to thank Michelle Arsenault and Isabelle Harrigan for their assistance with data collection and data entry.

Correspondence concerning this article should be addressed to Andréanne Charbonneau, School of Psychology, Université de Moncton, 18 Antonine-Maillet Avenue, Moncton, NB, Canada, E1A 3E9. Email: andreanne.plamondon@umoncton.ca