American Psychological Association

Psychiatric Rehabilitation Journal

1095-158X/19/$12.00

2019, Vol. 42, No. 4, 372–381
http://dx.doi.org/10.1037/prj0000368

Is the Wellness Recovery Action Plan (WRAP) Efficacious for Improving Personal and Clinical Recovery Outcomes? A Systematic Review and Meta-Analysis

Louise Canacott
University of Nottingham and Nottinghamshire Healthcare NHS Foundation Trust, Nottinghamshire, United Kingdom

Nima Moghaddam
University of Lincoln

Anna Tickle
University of Nottingham and Nottinghamshire Healthcare NHS Foundation Trust, Nottinghamshire, United Kingdom

Objective: The Wellness Recovery Action Plan (WRAP) is a structured approach to illness self-management that is widely used within mental health services. This systematic review identifies, appraises, and meta-analyzes quantitative evidence from experimental or quasi-experimental comparison group designs for effects of WRAP on measures reflecting personal recovery and clinical symptomatology. ***Method:*** A systematic strategy was used to search 6 electronic reference databases (Medline, CINAHL, EMBASE, PsycINFO, PsycARTICLES, The Cochrane Library) using full-text, keywords, and Medical Subject Headings (MeSh)/Thesaurus headings terms. Unpublished research was identified using the same strategy in the EThOS database. Controlled trials of WRAP were selected and assessed for quality. Meta-analyses were applied to quantitative data to establish the effects of WRAP on outcomes of interest. ***Results:*** Of 253 studies initially identified, 5 quantitative studies (reported in 6 papers) reporting controlled trials were included in the review. Meta-analyses revealed that, relative to inactive control conditions, WRAP was (a) superior for promoting self-perceived recovery outcomes (demonstrating a small-but-significant pooled effect), but (b) not superior for reducing clinical symptomatology—although restriction to randomized-controlled trials revealed 1 small effect favoring WRAP for reducing depression. ***Conclusions and Implications for Practice:*** Participation in WRAP has positive outcomes for participants, quantifiable using comprehensive measures of self-perceived recovery. Improvements were not sustained over time. Future research could explore this, as well as potential effects of follow-up WRAP sessions. The results support a move to broader measurement of outcomes within mental health, away from a reliance on clinical outcome measures. Recommendations for further research are made.

Impact and Implications
The current article pooled data from studies into the effectiveness of the Wellness Recovery Action Plan in adult mental health settings. The findings provide support for a broader conceptualization of outcomes than clinical symptom measurement: demonstrating significant growth in self-identified personal recovery in the absence of significant symptom reduction.

Keywords: recovery, Wellness Recovery Action Plan, illness management, mental health, meta-analysis

Recovery-oriented health care services depart from traditional understandings of recovery as a clinical concept, instead understanding recovery as a personal journey (Slade, 2009). Traditionally, within a biomedical model, recovery has been viewed as a reduction in clinical symptomatology. This was challenged by a disparity with subjective accounts of personal recovery (e.g., Deegan, 1996), which highlighted power imbalances within mental health services; limitations on autonomy and the subjugation of service-user views; and the central roles of hope, identity, empowerment, and independent living in the process of recovery (Lester & Glasby, 2010). Such accounts led to recovery being commonly understood as "a way of living a satisfying, hopeful, and contributing life even within the limitations caused by illness" (Anthony, 1993, p. 527).

Transdiagnostic recovery principles of empowerment, personal control, the development of positive identities, connectedness, hope for the future, and finding meaning in life (Tew et al., 2012) have now been adopted into mainstream approaches in mental

This article was published Online First May 27, 2019.

Louise Canacott, Trent Doctorate of Clinical Psychology, Division of Applied Psychiatry and Applied Psychology, University of Nottingham, and Nottinghamshire Healthcare NHS Foundation Trust, Nottinghamshire, United Kingdom; Nima Moghaddam, Trent Doctorate of Clinical Psychology, School of Psychology, University of Lincoln; Anna Tickle, Trent Doctorate of Clinical Psychology, Division of Applied Psychiatry and Applied Psychology, University of Nottingham, and Nottinghamshire Healthcare NHS Foundation Trust.

Correspondence concerning this article should be addressed to Anna Tickle, University of Nottingham, B Floor, YANG Fujia, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, UK. E-mail: Anna.Tickle@nottingham.ac.uk

health services (Slade, 2009) and recovery is viewed as a core component of meaningful working with individuals with mental illness (Shepherd, Boardman, & Slade, 2008).

The Wellness Recovery Action Plan

The Wellness Recovery Action Plan (WRAP; Copeland, 1997) is a structured peer-based, group program approach to illness self-management that may be applied to the management of physical and mental well-being (Copeland, 2001), based on Copeland's personal experiences. WRAP is facilitated by peers in recovery: Trained in WRAP by the Copeland Centre for Wellness and Recovery and receiving mentoring from advanced WRAP facilitators (Copeland, 1997). Key objectives are for participants to identify internal and external resources for facilitating recovery through development of a personalized wellness plan (Copeland, 1997). A typical WRAP series comprises eight to 10 weekly sessions of group education to enable participants to improve their ability to take responsibility for their own wellness, manage mental health symptoms using self-help strategies, and identify and use sources of support (Copeland, 1997, 2004).

WRAP may be used as a standalone intervention or integrated into existing services to enhance mainstream treatment. For example, a concerted effort to promote the use of WRAP in Scotland has seen it adopted by mental health self-help and mutual support groups across a range of organizations and integrated into Care Pathways for Depression as a suggested routine activity (Scottish Centre for Social Research & Pratt, 2010). Although WRAP might enhance mainstream services, there is also evidence that its use can reduce the use of mainstream services by individuals while continuing to improve on measures of psychiatric symptoms and self-perceived recovery (Cook et al., 2013).

Given that individuals with chronic mental health conditions may become frequent users of services (Naylor et al., 2012), there is a need to consider alternative self-management approaches to living with mental illness. WRAP is considered the most popular self-management strategy for maintaining mental health (Higgins et al., 2012). Preliminary evaluations suggest potentially beneficial outcomes of WRAP in multiple areas, including increased self-advocacy, reductions in service need and usage, reduced symptomatology, increased recovery-knowledge, and increased hopefulness and quality of life in individuals with persistent mental illness (Copeland Centre for Wellness & Recovery, 2014). Its implementation is expanding internationally: Studies report positive outcomes in Canada (Allott, Loganathan, & Fulford, 2002), China (Zhang, Li, Yeh, Wong, & Zhao, 2007), United States (Buffington, 2003), Ireland (Higgins et al., 2012), New Zealand (Doughty, Tse, Duncan, & McIntyre, 2008), Scotland (Pratt, MacGregor, Reid, & Given, 2013), and the United Kingdom (Davidson, 2005). WRAP has been recognized by the U.S. Substance Abuse and Mental Health Services Administration as an evidence-based program, despite no meta-analytic review of available literature. In view of this, and the growing adoption of WRAP training, it is timely to evaluate and critically appraise the evidence base for its effectiveness—with a focus on outcome data from robust experimental or quasi-experimental comparison group designs.

Objectives

The review aimed to answer the questions, "Does the Wellness Recovery Action Plan (WRAP) contribute to quantitatively measured personal and clinical recovery?" and "What is the quality of the existing literature evaluating WRAP?" The objectives were to

1. Systematically identify and assess the quality of papers reporting experimental or quasi-experimental comparison group designs investigating quantitative recovery outcomes for individuals with mental health difficulties receiving WRAP training.

2. Use meta-analysis to investigate quantitative evidence for the effects of WRAP on self-perceived recovery and clinical symptomatology for individuals with mental health difficulties.

Method

This review followed PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman & the PRISMA Group, 2009).

Searching

Six electronic reference databases (Medline, CINAHL, EMBASE, PsycINFO, PsycARTICLES, The Cochrane Library) were searched on July 9, 2017 using full-text, keywords, and Medical Subject Headings (MeSh)/Thesaurus headings terms as follows: ((well* OR health*) (recover* OR recuperate*) (action* OR plan*)); "Wellness Recovery Action Plan"; "WRAP"; exp. Mental Health; 3 AND 4; 1 OR 2 OR 5. No date restrictions were set. The EThOS database was searched using the same strategy to identify unpublished research. The references of all selected studies were hand-searched, as was the *Psychiatric Rehabilitation Journal*, as it is known to frequently publish articles pertinent to the topic of recovery in mental health.

Study Selection

The primary outcome of interest of this review was personal recovery as an outcome of WRAP training, assessed using comprehensive quantitative measures (i.e., measures designed to capture the broad meaning of recovery, across multiple domains considered important to mental health and wellness; Sklar, Groessl, O'Connell, Davidson, & Aarons, 2013). The second outcome of interest was clinical symptomatology, most reliably measured using standardized assessment tools. Qualitative studies regarding the perceived effects of WRAP training were considered in a complementary review currently in preparation.

To be included, studies must

- be published in the English language in peer-reviewed journals or "gray literature";
- investigate WRAP as delivered to adults with mental health difficulties (e.g., depression, anxiety, psychosis), but not restricted to those who have acquired a specific formal diagnosis;
- assess the effects of WRAP on clinical symptoms or personal recovery, using self-report, quantitative measures;

- use an experimental or quasi-experimental comparison group design to investigate WRAP outcomes (waiting list comparison group, treatment as usual comparison group, or randomized controlled trials).

The exclusion criteria were

- indistinguishable results from multiple stakeholders (where self-report data from adults with mental health difficulties could not be extracted);
- noninclusion of a validated outcome measure of either personal recovery or clinical symptomatology. Moreover, consistent with Sklar et al. (2013) only comprehensive measures of personal recovery were considered suitable: studies assessing singular components of recovery or recovery-related construct were excluded.

Each paper identified for inclusion in the review was read in detail and key data extracted.

Quality Appraisal

It is important to assess the quality of included studies, to support consideration of the weight that should be accorded to findings and inform the quality of future studies through critical appraisal of existing research. There are several standardized tools available to facilitate this, but research has demonstrated a low-level of interrater reliability in quality assessments using such tools (Jüni et al., 1999). The Quality Assessment Tool for Quantitative Studies developed by the Effective Public Health Practice Project (1998) was used to determine the quality of quantitative papers because it can be used to assess the quality of RCTs, Quasi-experimental, case control, and prospective studies—that is, the tool is applicable across the various designs eligible for inclusion in this review. Quality appraisal was undertaken independently by Louise Canacott and Nima Moghaddam, with a plan to discuss any discrepancies and agree final consensus ratings.

Analysis

For each outcome of interest, meta-analyses were completed as a means of pooling outcomes across studies to increase power and improve the precision of establishing whether the intervention had a significant effect on outcomes of interest. RevMan 5.3 (Cochrane Collaboration, 2014) was used to conduct meta-analyses. Standardized mean differences were obtained with 95% confidence intervals (CIs) for continuous outcomes to obtain a measure of treatment effect. The I^2 statistic was used to statistically assess for heterogeneity (Higgins, Whitehead, Turner, Omar, & Thompson, 2001): This statistic describes the percentage of total variation in effect estimates that is due to heterogeneity (between-study differences) versus random error (chance). Analyses in which the I^2 statistic exceeded 50% were interpreted with additional caution (all such analyses are highlighted in the discussion section, with conclusions qualified accordingly): Clinical and methodological diversity may underlie such statistical heterogeneity and limit confidence in combined estimates of effects (which could vary considerably according to as-yet unidentified moderator variables).

For purposes of meta-analysis, all relevant outcome data were extracted from original study-reports as unadjusted (crude) descriptive statistics (means and standard deviations, with respective sample sizes). As recommended by Higgins et al. (2001), data obtained from ordinal outcome measures were treated as continuous data. Standardized mean difference was used as a summary statistic as studies used a range of outcome measures to assess symptomatology (depression and anxiety) and personal recovery. Because of the assumed heterogeneity among studies (e.g., in terms of sampled populations and applied outcome measures), random-effects models were used within RevMan: Applying adjustments to study weights according to the degree of between-study inconsistency in observed effects. Exceptionally, fixed-effects models were applied for any analyses based on fewer than three studies: In such cases, it is difficult to estimate between-study heterogeneity with precision, and a fixed-effects model is arguably more apt (Chen, Fang, & Wilson, 2017).

Sensitivity analyses were conducted to test whether findings were robust to decisions made about study inclusion. Specifically, it was considered apt to test whether overall effect estimates (and study homogeneity) were influenced by the inclusion of non-RCTs (quasi-experimental designs; which may be considered to produce a lower grade of evidence/control over threats to internal validity): This was accomplished by rerunning analyses (for each outcome variable) with restriction to data from RCTs alone. Essentially, these sensitivity analyses explored the impact of including data from non-RCT studies (which have a greater risk of bias)—if results appear similar (in informal comparisons) across analyses, this would indicate that findings are not contingent on decisions about including/excluding such data.

For studies reporting multiple measures of (or conceptually relating to) recovery, a single measure was selected for data-extraction and analysis by prioritizing comprehensiveness (capture of recovery as a broad, multidomain construct; following Sklar et al., 2013).

Results

Search Results

Figure 1 outlines the selection process, which identified five studies (reported across six articles) for inclusion. Articles ex-

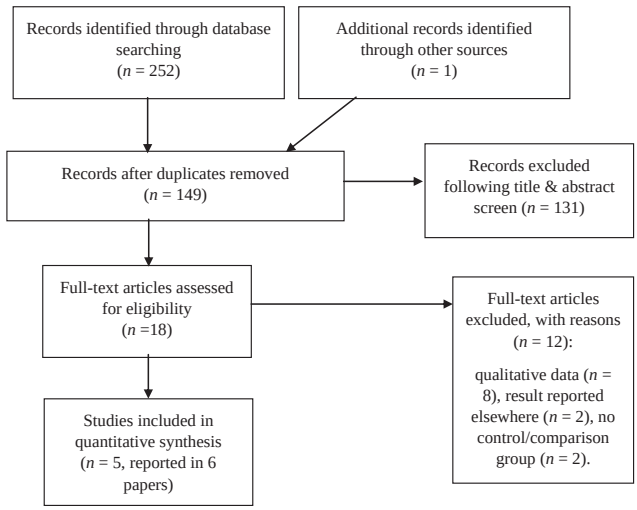

Figure 1. PRISMA diagram outlining selection process.

cluded at the title/abstract-screening stage ($n = 131$) were clearly ineligible for inclusion, primarily because they did not (a) pertain to WRAP intervention (e.g., studies of 'recovery' from surgery), (b) provide primary data on outcome effects (e.g., editorials), or (c) include a mental health population.

Data Abstraction and Analysis

Table 1 outlines the key characteristics of selected studies. Each study is assigned a number (1–5), used for brevity within the results section. Please note that the study by Cook, Copeland, Floyd, et al. (2012; Cook, Copeland, Jonikas, et al., 2012) is referred to as 1a and 1b, reflecting the two papers that report different aspects of the study.

Study Characteristics and Key Findings

Five studies from the United States ($n = 2$), Republic of Ireland ($n = 2$), and China ($n = 1$) were selected for quantitative analyses. Two studies (Studies 1a/1b and 4; combined $n = 555$) were randomized controlled trials (RCTs), the remainder (Studies 2, 3, and 5; combined $n = 290$) used quasi-experimental methods (matched- or nonequivalent comparison groups). Comparison/control samples were either receiving treatment as usual (TAU; $n = 2$; Studies 2 and 5), on a waiting list to receive WRAP ($n = 2$; Studies 3, 4), or receiving TAU while on a waiting list to receive WRAP ($n = 1$; Study 1a/1b).

The total number of participants across the five studies was 845, comprising 37% males and 63% females; the mean age of participants across studies was 45.1 ($SD = 10.7$). One RCT produced two papers (1a and 1b) and the number of participants in this study is included only once in the above total figure. Sample size ranged from 36 (4) to 519 (1a and 1b). The *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., revised (*DSM–IV–TR*) was used in two studies (Studies 1a/1b and 2) to specify diagnostic criteria. The International Classification of Diseases-10 was used in one study (Study 4), and two studies provided demographic diagnostic information without specifying diagnostic criteria Studies (3 and 5). Diagnoses included depression, anxiety, obsessive–compulsive disorder, schizophrenia, psychosis, bipolar disorder, schizoaffective disorder, and others (as classified by the *DSM–IV–TR*); and psychoactive substance use disorders, schizophrenia, schizotypal, and delusional disorders, affective disorders, and neurotic and somatoform disorders (as classified by the ICD-10). In most studies, the majority of participants presented with a diagnosed mood disorder (Studies 1a/1b, 2, and 4); Mak et al. (Study 5) reported on a sample wherein the majority of participants (54%) had a diagnosis of schizophrenia; one study (Study 3) did not provide separate frequencies for participants with "mental health" diagnoses (inclusive of individuals with diagnosed mood or schizophrenia-spectrum disorders). In studies that reported on concurrent medication-use (Studies 1a/1b and 5), most participants were in receipt of psychotropic mediation.

WRAP intervention was delivered in eight to 12 weekly sessions, except in Study 4, wherein the intervention was delivered over a 2-day workshop schedule. Study 4 was also notably divergent in that WRAP intervention was not delivered with the involvement of service-user/peer facilitators (although the facilitators—all allied health professionals—were suitably trained and accredited and were reportedly encouraged to focus on personal vs. professional experiences when in their facilitative role). All studies included a postintervention assessment-point, with all but one (Study 3) including follow-up assessments (although there was limited consistency/comparability between studies in these follow-up assessments, precluding pooled analysis or further interpretation here). Three reported follow-up assessments for both (experimental [WRAP] and comparison) groups, at either 3- (Study 5) or 6-months (Study 1a/1b and 2) postintervention. One study (Study 4) only reported (6-month) follow-up data for the experimental (WRAP) group (precluding comparative analysis). Clinical symptomatology was assessed using the Brief Symptom Inventory ($n = 1$; Study 1a/1b), Colorado Modified Symptom Index ($n = 2$; Studies 2 and 5) and the Hospital Anxiety and Depression Scale ($n = 2$; Studies 3 and 4).

Quality Appraisal

Table 2 presents the quality appraisal of included studies. Based on independent appraisals, Louise Canacott and Nima Moghaddam agreed on all ratings (precluding the need to resolve discrepancies by consensus).

Most samples appeared to be representative of adults receiving community-based adult mental health services, comprising various diagnoses and educational backgrounds, with inclusion in some studies of participants from various minority groups. The sample in one study (Study 4) consisted of inpatients and outpatients and results were not differentiated by inpatient status. Representativeness is limited in each study due to the nature of voluntary participation.

Participants were not blind to the intervention in any of the studies, which would be practically difficult to achieve given the nature of WRAP, but assessors were reported to be blind to the intervention in three studies (Studies 1a/1b, 2, and 4). Randomized controlled trials (Studies 1a/1b and 4) were considered to have strong study designs: reducing allocation bias through appropriate and clearly stated random allocation methods in addition to concealed allocation. All papers that compared samples prior to intervention reported no significant baseline differences between experimental and comparison groups (controlling for observed confounders). Four studies reported participant attrition; and from these studies, between 80 and 100% of participants completed the intervention and outcome assessments. Intervention fidelity was assessed in three studies (Studies 1a/1b, 4, and 5) and deemed to be sufficient (86–100%). However, fidelity assessment was based on facilitator self-evaluation in one of these studies (Study 4), and another (Study 5) observed that facilitators did not always have time to explain core components (including crisis and postcrisis planning)—potentially raising questions about the sufficiency of fidelity (and robustness of fidelity assessment).

Meta-Analytic Results

Recovery

Five studies assessed recovery using standardized questionnaires (Studies 1a, 2, 3, 4, and 5; see Figure 2). Three studies used the Mental Health Recovery Measure (Studies 3, 4, and 5), one study used the Recovery Assessment Scale (Study 1) and one

Table 1
Key Characteristics of Selected Studies

Author(s), year and location	Design	Study sample N = total (female: male)	Research aims	Clinical outcome measures	Recovery outcome measures	Summary points and key findings
1(a). Cook, Copeland, Floyd, et al., 2012, USA	RCT (treatment as usual [TAU] waiting list control group)	N = 519 (342:177), outpatient sample, with documented DSM-IV-TR diagnosis of mental disorder	To compare symptomatology of anxiety and depression and self-perceived recovery among those receiving WRAP versus those receiving TAU	Brief Symptom Inventory depression and general anxiety subscales (Piersma, Reaume, & Boes, 1994)	Recovery Assessment Scale (Giffort, Schmook, & Wood, 1995)	• WRAP participants reported significantly greater symptom reduction in anxiety (r = .44*). and depression (r = .48***) scales • WRAP participants reported significantly greater self-perceived recovery (r = .52*)
1(b). Cook, Copeland, Jonikas, et al., 2012, USA	As for 1(a)	As for 1(a)	To determine the efficacy of WRAP (versus TAU) for symptom reduction, increased hopefulness, and increased quality of life	Global Symptom Severity Index of the Brief-Symptom Inventory (Piersma et al., 1994). Positive Symptom Total of the Brief-Symptom Inventory (Piersma et al., 1994)		• WRAP participants reported significantly greater symptom reduction (r = .2*) • WRAP participants reported significantly increased rates of hope (r = .2*,) and quality of life (r = .1***) versus controls • Decreased symptomatology was associated with increased number of WRAP sessions received
2. Fukui et al., 2011, USA	Quasi-experimental (matched comparison group)	N = 114 (71:43), community sample with severe and persistent mental illness	To examine the effects of WRAP participation on psychiatric symptoms, hope, and recovery outcomes	Modified Colorado Symptom Index, (Conrad et al., 2001)	Recovery Markers Questionnaire (Ridgway, Press, Ratzlaff, Davidson, & Rapp, 2003)	• Significant improvements for WRAP group in symptomatology (r = .33***) and hope (r = .58*) • No significant interaction effect reported on recovery outcome measure (no figures provided)
3. O'Dwyer, 2014, Republic of Ireland	Quasi experimental (non-equivalent, waiting list comparison group)	N = 57 (35:23[a]), Participants with mental illness or acquired brain (AB) injury receiving support from community service	To examine the effects of WRAP on anxiety, depression and recovery in clients an adult mental health population and an acquired brain injury population	Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983)	The Mental Health Recovery Measure (Young & Bullock, 2003)	• Statistically significant differences in anxiety and depression scores for mental health group compared to wait-list counterparts • More pronounced differences for mental health group than AB group • Statistically significant increases in recovery knowledge in both groups
4. O'Keeffe et al., 2016, Republic of Ireland	RCT (waiting list control group)	N = 36 (19:17), Participants from inpatient and outpatient mental health settings	To evaluate the effect of WRAP on personal recovery, quality of life, and self-reported psychiatric symptoms	Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983)	The Mental Health Recovery Measure (Young & Bullock, 2003)	• Positive outcomes of WRAP in domains of addictive behaviors and identity and self-esteem • No significant effect of WRAP on personal recovery, quality of life or psychiatric symptoms
5. Mak et al., 2016, China	Quasi-experimental (matched comparison group)	N = 118 (68:50) Service-users of residential and vocational rehabilitation services	To evaluate the effect of WRAP on symptom severity and recovery	Modified Colorado Symptom Index, (Conrad et al., 2001)	Recovery Markers Questionnaire (Ridgway et al., 2003)	• Significant increase in perceived social support for WRAP group • No significant change in symptom severity or recovery for WRAP group

Note. RCT = randomized control trial; TAU = treatment as usual; ABI = acquired brain injury.
[a] Demographics are described for a sample of N = 58 however outcome data were reported for 57 participants.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 2
Methodological Quality Appraisal

Study	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and dropouts	Overall quality rating
1(a). Cook, Copeland, Floyd, et al., 2012	Moderate	Strong	Strong	Moderate	Strong	Strong	Strong
1(b). Cook, Copeland, Jonikas, et al., 2012	Moderate	Strong	Strong	Moderate	Strong	Strong	Strong
2. Fukui et al., 2011	Moderate	Moderate	Strong	Moderate	Strong	Strong	Strong
3. O'Dwyer, 2014	Moderate	Moderate	Weak	Weak	Strong	Weak	Weak
4. O'Keeffe et al., 2016	Strong	Strong	Strong	Moderate	Strong	Strong	Strong
5. Mak et al., 2016	Moderate	Moderate	Strong	Weak	Strong	Strong	Moderate

Note. Selection bias—strong: selected individuals very likely representative of target population and > 80% participation; moderate: selected individuals at least somewhat likely representative of target population and 60–79% participation; weak: selected individuals not likely representative of target population; or there is < 60% participation or selection not described; and level of participation not described. Design—strong: randomized controlled trials and controlled clinical trials; Moderate: cohort analytic study, case control study, cohort design, or interrupted time-series; and weak: any other method or method unstated. Confounders—strong: controlled for ≥ 80% of relevant confounders; moderate: controlled for 60–79% of relevant confounders; and weak: <60% of relevant confounders were controlled or control not described. Blinding—good: outcome assessor unaware of intervention status of participants; and participants unaware of research question; moderate: outcome assessor unaware of intervention status of participants; or participants unaware of research question; and weak: outcome assessor aware of intervention status of participants; and participants aware of research question; or blinding not described. Data collection methods—strong: data collection tools have been shown to be valid; and shown to be reliable; moderate: tools have been shown to be valid; and have not been shown to be reliable or reliability is not described; and weak: tools have not been shown to be valid or both reliability and validity are not described. Withdrawals and drop-outs—strong: follow-up rate ≥ 80%; moderate: follow-up rate is 60–79%; and weak: follow-up rate is < 60% or withdrawals and drop-outs not described.

study used the Recovery Markers Questionnaire (Study 3). Postintervention, there was an overall significant effect favoring WRAP intervention (standardized mean difference [SMD] 0.24; 95% CI [0.06, 0.42]; $p = .01$; $I^2 = 23\%$). This effect was of small magnitude but indicated that WRAP was effective in improving recovery outcomes (relative to inactive control conditions). Limiting the meta-analysis to the randomized trials (Cook, Copeland, Floyd, et al., 2012 [Study 1a]; O'Keeffe et al., 2016 [Study 4]) reduced heterogeneity ($I^2 = 0\%$) but did not substantively alter the magnitude or significance of the overall effect (SMD: 0.24; 95% CI [0.02, 0.37]; $p = .03$). Thus, results appeared robust to (i.e., unbiased by) the inclusion of non-RCT studies.

Clinical Symptomatology

Five studies assessed clinical symptomatology as an outcome of attending WRAP training. Three investigated symptomatology as an overall construct (Studies 1b, 2, and 5), and three assessed anxiety and depression independently (Studies 1a, 3, and 4). The effects of WRAP in relation to overall clinical symptomatology, symptoms of anxiety, and symptoms of depression will each be discussed below. A final analysis will combine data from all studies, by using one marker of clinical symptomatology from each study (preferring overall symptomatology scores where available [Studies 1b, 2, and 5] and otherwise using anxiety/depression symptom scores [Studies 3 and 4]). This final analysis necessarily conflates distress-specific versus overall symptom-change but makes fuller use of available information on clinical symptomatology outcomes (affording maximal power to detect effects).

Overall Clinical Symptoms

Three studies assessed the effects of WRAP on overall clinical symptomatology (1b, 2, 5). Two studies (2, 5) used the Modified Colorado Symptom Index (MCSI) and one (1b) used the Brief Symptom Inventory (BSI). When combining study results, there was no significant effect of WRAP intervention on overall clinical symptomatology (SMD: 0.01; 95% CI [−0.30, 0.33]; $p = .94$) and heterogeneity was substantial ($I^2 = 70\%$). (see Figure 3). However, it was notable that limiting to the single RCT reporting on this outcome (1b) would suggest a small but significant effect favoring WRAP intervention over treatment-as-usual (SMD: −0.19; 95% CI [−0.38, −0.01]; $p = .04$). Restriction to RCT evidence leads to a substantive change in estimates for this outcome (albeit limited to data from a single study).

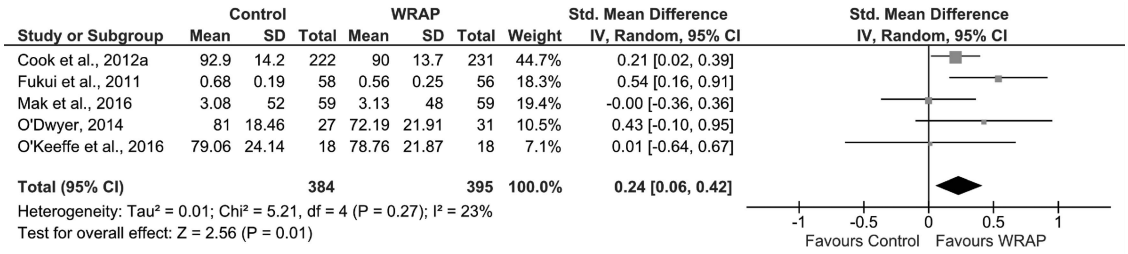

Figure 2. Effects of Wellness Recovery Action Plan on standardized recovery questionnaires.

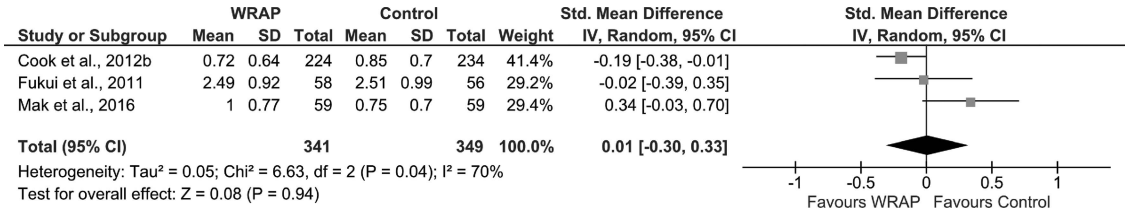

Figure 3. Effects of Wellness Recovery Action Plan on overall clinical symptomatology.

Anxiety

Three studies (Studies 1a, 3,and 4) examined the effect of WRAP on symptoms of anxiety in comparison to inactive control conditions (treatment as usual or waiting list; see Figure 4). Two studies (Studies 3 and 4) used the Hospital Anxiety and Depression Scale (HADS) and one study (Study 1a) used the anxiety scale of the BSI. No significant effect of treatment condition on symptoms of anxiety was observed (SMD: −0.21; 95% CI [−0.65, 0.23]; $p = .36$) and heterogeneity was substantial (66%). This overall finding appeared to be robust to the inclusion of non-RCT evidence: in a sensitivity analysis, removal of O'Dwyer (2014; Study 3) did not substantively alter the magnitude or nonsignificance of the overall effect (fixed effects SMD: −0.14; 95% CI [−0.32, 0.04]; $p = .12$) and heterogeneity remained substantial (53%).

Depression

Three studies examined the effects of WRAP on symptoms of depression (Studies 1a, 3, and 4; see Figure 5). Two studies used the HADS as a measure of depressive symptoms (Studies 3 and 4). No significant effect of intervention type on symptoms of depression was observed following intervention (SMD: −0.17; 95% CI [−0.50, 0.17]; $p = .34$; $I^2 = 44\%$). This overall finding appeared to be robust to the inclusion of non-RCT evidence: in a sensitivity analysis, removal of the non-RCT study (3) reduced heterogeneity ($I^2 = 0\%$) but did not substantively alter the magnitude or nonsignificance of the overall effect (fixed effects SMD: −0.09; 95% CI [−0.27, 0.09]; $p = .32$).

Finally, although not a focus of the current study and therefore not included in meta-analyses, it is important to note that two studies (Studies 1b and 2) found significant increases in hope following WRAP and another (Study 4) reported positive outcomes relating to identity, self-esteem and addictive behaviors. Findings related to quality of life were reported in two studies but mixed, with one (Study 1b) reporting significant improvement and another (Study 4) reporting no significant effect of WRAP on quality of life.

Marker of Clinical Symptomatology (Maximally Inclusive Analysis)

Collapsing (overall and distress-specific) symptom-categories, all five studies examined the effects of WRAP on at least one marker of clinical symptomatology: overall measures were used as markers where available (Studies 1b, 2, and 5) and depression/anxiety scores were used as markers in the two remaining studies (Studies 3 and 4). When combining study results, heterogeneity was substantial and there was no significant postintervention effect of WRAP intervention on markers of clinical symptomatology. This finding was consistent irrespective of whether depression-scores (SMD: −0.05; 95% CI [−0.33, 0.23]; $p = .72$; $I^2 = 62\%$) or anxiety-scores (SMD: −0.05; 95% CI [−0.36, 0.25]; $p = .73$; $I^2 = 69\%$) were used as symptom-markers for the two studies without overall symptom-indices. Limiting to RCT evidence (Studies 1b and 4) did not substantively alter these results, irrespective of whether the symptom-marker for Study 4 was depression-score (fixed effects SMD: −0.16; 95% CI [−0.34, 0.01]; $p = .07$; $I^2 = 21\%$) or anxiety-score (fixed effects SMD: −0.16; 95% CI [−0.33, 0.02]; $p = .08$; $I^2 = 56\%$).

Conclusions and Implications for Practice

Meta-analysis demonstrated a small but significant effect of WRAP training on self-perceived recovery. The evidenced beneficial effect on personal recovery, despite participants not reporting significant decreases in their clinical symptoms, supports a move to broader measurement of change—away from a reliance on clinical outcome measures (Andresen, Caputi, & Oades, 2010). A focus on symptom-based clinical outcome measures and associated caseness criteria can encourage a reductive characterization of service-users in terms of presence versus absence of illness. Results from this review indicate that there are many therapeutic benefits of managing mental illness that contribute toward personal wellbeing as a whole. Consistent with the personal recovery ethos of living a fulfilling life in the presence of ongoing illness (Deegan, 1988), results evidenced self-perceived growth (in terms

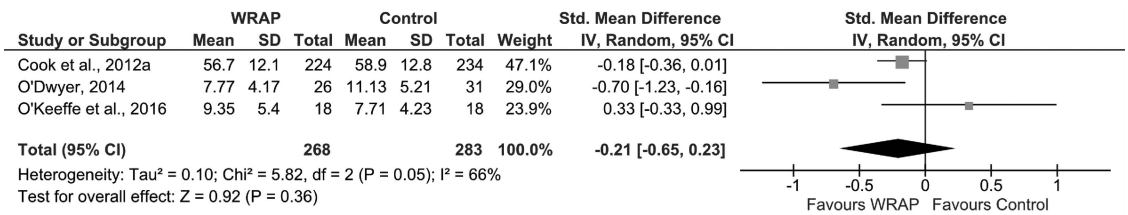

Figure 4. Effects of Wellness Recovery Action Plan on symptoms of anxiety.

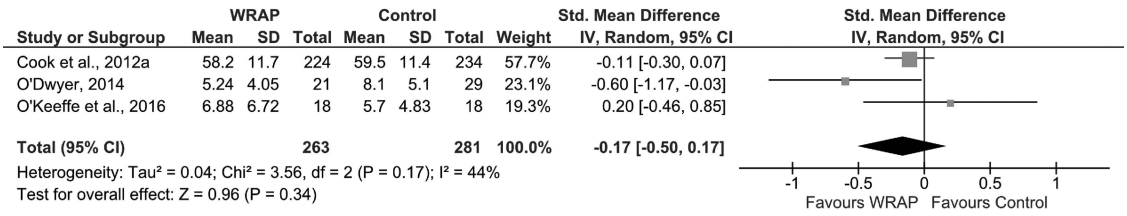

Figure 5. Effects of Wellness Recovery Action Plan on symptoms of depression.

of recovery markers, encompassing components such as self-efficacy, hope, and overall wellbeing) even in the absence of demonstrable symptom reduction. Future studies could further explore the effects of WRAP on recovery markers such as hope and quality of life.

Conversely, meta-analytic results showed no significant effect of WRAP in the reduction of clinical symptomatology, either as an overall construct or when examining anxiety and depression independently. Exceptionally, when restricting to RCT evidence alone, there was one small effect favoring WRAP for reducing overall symptoms: This estimate was effectively based on a single study, rather than a pooling of multiple estimates (i.e., meta-analysis per se), limiting confidence in its reliability and reproducibility across study-contexts—this single study was an RCT, appraised to be of strong overall quality, but the relationship between internal validity and external validity (or transferability) is complex (Jimenez-Buedo & Miller, 2010). Although WRAP does not purport to reduce the magnitude of clinical symptoms, large-scale studies conducted in the United States have reported significant reductions following WRAP training (e.g., Cook, Copeland, Floyd, et al., 2012; Cook, Copeland, Jonikas, et al., 2012; Fukui et al., 2011). Two of the three studies from outside the United States (O'Keeffe et al., 2016; Mak et al., 2016), albeit with smaller sample sizes, did not report similar findings. Whether such findings represent applicability to non-United States populations or are a result of statistical power or other variables remains unclear. The discrepancy between individual study reports and present meta-analytic results (combining data from individual study reports) is comment-worthy: Some effects that were found to be statistically significant in original studies reflected adjustments applied to outcome data (e.g., controlling for baseline values) or particular comparisons drawn (e.g., differential within-group changes vs. direct between-groups comparisons); in some cases, the significance of these individual symptomatology effects did not survive when extracted as unadjusted summary statistics and compared between groups (Cook, Copeland, Floyd, et al., 2012a; Fukui et al., 2011). Moreover, the pooling of effect estimates across studies—to provide a more precise and suitably weighted estimate of effectiveness—showed that, on balance, available evidence does not support effects of WRAP training on clinical symptomatology. This finding affirms the value of applying meta-analysis to combine (sometimes inconsistent) results from across relevant studies, to reach a more dependable conclusion about the evidence to date—which is less sensitive to the idiosyncrasies and analytic approaches of individual studies. Notwithstanding the above, confidence intervals around pooled estimates were wide, and do not exclude the possibility of beneficial effects on clinical symptoms—additional primary data are needed to increase precision of estimates.

Notably, studies provided limited (postintervention) follow-up data—and follow-up periods varied between studies, precluding meaningful synthesis. As Fukui et al. (2011) have observed, mental health recovery is a lengthy and nonlinear process: Greater insight would be afforded by studies employing longer-term follow-up periods, with repeated observations over time (and appropriate modeling of individual differences).

Reviewed studies provided little insight into variables mediating positive (short-term) outcomes for recovery: The "active ingredients" of WRAP training are difficult to discern from the available quantitative data—synthesis of qualitative data is more revealing in this regard (Canacott, Moghaddam & Tickle, in preparation). Indeed, studies did not consistently apply robust assessments of implementation fidelity (i.e., whether WRAP processes were delivered as prescribed): Two of the five reviewed studies failed to measure fidelity (O'Dwyer, 2014; Fukui et al., 2011) and one study used facilitator self-evaluation of fidelity (O'Keeffe et al., 2016), which may be particularly susceptible to bias (independent observation of session delivery would be a more defensible basis for assessment). Moreover, judgments of satisfactory fidelity in Mak et al. (2016) were arguably inconsistent with observations that some principal components (around crisis and postcrisis planning) were not afforded enough coverage—while independent of the research, it is unclear whether the evaluator in this instance (a clinical psychology graduate student) was adequately supported or trained to judge fidelity. Taken together, the reviewed evidence does not allow for strong conclusions about the particular mechanisms through which WRAP training might inculcate change—potential attribution of effects to WRAP-specific processes is limited by uncertainties around the fidelity of WRAP implementation.

Study heterogeneity limits the strength of conclusions that can be drawn from meta-analytic findings, particularly with respect to overall and anxiety symptomatology at postintervention—considerable statistical heterogeneity was evident in pooled estimates relating to these outcomes. There were various observable differences across studies (e.g., in study design, population, intervention delivery, comparison conditions, and specific outcome measures) that may have contributed to statistical heterogeneity, and which conceptually complicate pooling of effect-estimates—although the variability between studies arguably increases potential generalizability (extending to a broader range of contexts for WRAP training) and was appropriately modeled in applied meta-analyses (using random-effects models [where apt] to accommodate an assumption that different studies are estimating different-yet-related intervention effects, and sensitivity analysis to check whether results were influenced by the decision to include data from both RCT and non-RCT designs).

There was ostensibly little variation between studies in appraised quality (all but one study achieved overall ratings of moderate-to-strong quality) and most were rated as strong, with no consistent pattern in terms of areas of weakness (i.e., aspects of study conduct or reporting which should be targeted for improvement in future research) notwithstanding inherent practical constraints on blinding and selection. However, the applied quality assessment tool was broad in its appraisal criteria: This enabled application across a range of study designs but offered limited specificity for discriminating the relative quality of studies (for examining the efficacy of WRAP training). Consideration of study characteristics identified some particular sources and patterns of variability that limit the quality of available evidence: for example, as discussed above, basic checks on intervention fidelity were either omitted or of questionable adequacy in half of the reviewed studies—a clear recommendation for future research is that studies implement robust, independent assessments to check the fidelity of WRAP delivery (i.e., whether the study is predicated on implementation of WRAP processes).

To further contribute to the evidence-base, trials comparing WRAP to other recovery-oriented, peer-led interventions may produce more discriminating insight into the effectiveness of WRAP than those comparing WRAP to treatments targeting clinical symptoms. Longer-term follow-up data may provide greater understanding of the recovery journey and whether changes in personal and/or clinical recovery are sustained. As WRAP implementation continues to expand internationally, further research regarding its short- and long-term practice are essential to test and develop its position as an evidence-based intervention.

References

Allott, P., Loganathan, L., & Fulford, K. (2002). Discovering hope for recovery. *Canadian Journal of Community Mental Health, 21,* 13–33. http://dx.doi.org/10.7870/cjcmh-2002-0014

Andresen, R., Caputi, P., & Oades, L. G. (2010). Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research, 177,* 309–317. http://dx.doi.org/10.1016/j.psychres.2010.02.013

Anthony, W. A. (1993). Recovery from mental illness. The guiding vision of the mental health service system in the 1990's. *Psychosocial Rehabilitation Journal, 16,* 11–23. http://dx.doi.org/10.1037/h0095655

Buffington, E. (2003). *Wellness Recovery Action Plan: WRAP evaluation, State of Minnesota*. Minneapolis: Mental Health Consumer/Survivor Network of Minnesota.

Chen, D. G., Fang, D., & Wilson, J. R. (2017). Meta-analysis of two studies with random effects? *Journal of Minimally Invasive Gynecology, 24,* 689–690. http://dx.doi.org/10.1016/j.jmig.2017.05.008

Cochrane Collaboration. (2014). Review Manager (RevMan), Version 5.3 [Computer program]. Copenhagen, Denmark: The Nordic Cochrane Centre.

Conrad, K. J., Yagelka, J. R., Matters, M. D., Rich, A. R., Williams, V., & Buchanan, M. (2001). Reliability and validity of a modified Colorado Symptom Index in a national homeless sample. *Mental Health Services Research, 3,* 141–153. http://dx.doi.org/10.1023/A:1011571531303

Cook, J. A., Copeland, M. E., Floyd, C. B., Jonikas, J. A., Hamilton, M. M., Razzano, L., . . . Boyd, S. (2012). A randomized controlled trial of effects of Wellness Recovery Action Planning on depression, anxiety, and recovery. *Psychiatric Services, 63,* 541–547. http://dx.doi.org/10.1176/appi.ps.201100125

Cook, J. A., Copeland, M. E., Jonikas, J. A., Hamilton, M. M., Razzano, L. A., Grey, D. D., . . . Boyd, S. (2012). Results of a randomized controlled trial of mental illness self-management using Wellness Recovery Action Planning. *Schizophrenia Bulletin, 38,* 881–891. http://dx.doi.org/10.1093/schbul/sbr012

Cook, J. A., Jonikas, J. A., Hamilton, M. M., Goldrick, V., Steigman, P. J., Gray, D. D., . . . Copeland, M. E. (2013). Impact of Wellness Recovery Action Planning on service utilization and need in a randomized controlled trial. *Psychiatric Rehabilitation Journal, 36,* 250–257. http://dx.doi.org/10.1037/prj0000028

Copeland, M. (1997). *The Wellness Recovery Action Plan*. Dummerston, VT: Preach Press.

Copeland, M. (2001). Wellness Recovery Action Plan: A system for monitoring, reducing and eliminating uncomfortable or dangerous physical symptoms and emotional feelings. *Occupational Therapy in Mental Health, 17,* 127–150.

Copeland, M. (2004). *Leading a mental health recovery and WRAP facilitator training*. Brattleboro, CT: Preach Press.

Copeland Centre for Wellness and Recovery. (2014). *The way WRAP works! Strengthening core values & practices*. Retrieved from https://copelandcentre.com/resources/way-wrap-works

Davidson, L. (2005). Recovery, self-management and the expert patient: Changing the culture of mental health from a United Kingdom perspective. *Journal of Mental Health, 14,* 25–35. http://dx.doi.org/10.1080/09638230500047968

Deegan, P. E. (1988). Recovery: The lived experience of rehabilitation. *Psychosocial Rehabilitation Journal, 11,* 11–19. http://dx.doi.org/10.1037/h0099565

Deegan, P. E. (1996). Recovery as a journey of the heart. *Psychiatric Rehabilitation Journal, 19,* 91–97. http://dx.doi.org/10.1037/h0101301

Doughty, C., Tse, S., Duncan, N., & McIntyre, L. (2008). The Wellness Recovery Action Plan (WRAP): A workshop evaluation. *Australasian Psychiatry, 16,* 450–456. http://dx.doi.org/10.1080/10398560802043705

Effective Public Health Practice Project. (1998). *Quality assessment tool for quantitative studies*. Hamilton, Canada: Effective Public Health Practice Project. Retrieved from https://merst.ca/ephpp/

Fukui, S., Starnino, V. R., Susana, M., Davidson, L. J., Cook, K., Rapp, C. A., & Gowdy, E. A. (2011). Effect of Wellness Recovery Action Plan (WRAP) participation on psychiatric symptoms, sense of hope, and recovery. *Psychiatric Rehabilitation Journal, 34,* 214–222. http://dx.doi.org/10.2975/34.3.2011.214.222

Giffort, D., Schmook, A., & Wood, C. (1995). *Construction of a scale to measure consumer recovery*. Springfield, IL: Office of Mental Health.

Higgins, A., Callaghan, P., DeVries, J., Keogh, B., Morrissey, J., Nash, M., . . . Carter, T. (2012). Evaluation of mental health recovery and Wellness Recovery Action Planning education in Ireland: A mixed methods pre-postevaluation. *Journal of Advanced Nursing, 68,* 2418–2428. http://dx.doi.org/10.1111/j.1365-2648.2011.05937.x

Higgins, J. P. T., Whitehead, A., Turner, R. M., Omar, R. Z., & Thompson, S. G. (2001). Meta-analysis of continuous outcome data from individual patients. *Statistics in Medicine, 20,* 2219–2241. http://dx.doi.org/10.1002/sim.918

Jimenez-Buedo, M., & Miller, L. M. (2010). Why a trade-off? The relationship between the external and internal validity of experiments. *Theoria Revista de Teoría*. Historia y Fundamentos de la Ciencia, *25,* 301–321.

Jüni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association, 282,* 1054–1060. http://dx.doi.org/10.1001/jama.282.11.1054

Lester, H., & Glasby, J. (2010). *Mental health policy and practice*. Basingstoke, UK: Palgrave Macmillan.

Mak, W. W. S., Chan, R. C. H., Pang, I. H. Y., Chung, N. Y. L., Yau, S. S. W., & Tang, J. P. S. (2016). Effectiveness of Wellness Recovery Action Planning (WRAP) for Chinese in Hong Kong. *American Journal of Psychiatric Rehabilitation, 19,* 235–251. http://dx.doi.org/10.1080/15487768.2016.1197859

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151,* 264–269, W64. http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135

Naylor, C., Parsonage, M., McDaid, D., Knapp, M., Fossey, M., & Galea, A. (2012). *Long-term conditions of mental health: The cost of co-morbidities*. London, UK: The King's Fund, Centre for Mental Health. Retrieved from www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/long-term-conditions-mental-health-cost-comorbidities-naylor-feb12.pdf

O'Dwyer, D. (2014). *Embracing wellness, self management, and the positive interface of Eastern and Western psychology*. Retrieved from http://openaccess.city.ac.uk/14565/1/O%27Dwyer%2C%20Denise%20%28redacted%29.pdf

O'Keeffe, D., Hickey, D., Lane, A., McCormack, M., Lawlor, E., Kinsella, A., . . . Clarke, M. (2016). Mental illness self-management: A randomised controlled trial of the Wellness Recovery Action Planning intervention for inpatients and outpatients with psychiatric illness. *Irish Journal of Psychological Medicine, 33,* 81–92. http://dx.doi.org/10.1017/ipm.2015.18

Piersma, H. L., Reaume, W. M., & Boes, J. L. (1994). The Brief Symptom Inventory (BSI) as an outcome measure for adult psychiatric inpatients. *Journal of Clinical Psychology, 50,* 555–563. http://dx.doi.org/10.1002/1097-4679(199407)50:4<555::AID-JCLP2270500410>3.0.CO;2-G

Pratt, R., MacGregor, A., Reid, S., & Given, L. (2013). Experience of wellness recovery action planning in self-help and mutual support groups for people with lived experience of mental health difficulties. *The Scientific World Journal, 2013,* 180587. http://dx.doi.org/10.1155/2013/180587

Ridgway, P., Press, A., Ratzlaff, S., Davidson, L., & Rapp, C. (2003). *Report on field testing the Recovery Enhancing Environment (REE) measure*. Lawrence, KS: University of Kansas School of Social Welfare, Office of Mental Health Research and Training.

Scottish Centre for Social Research and Pratt. R. (2010). *An evaluation of wellness planning in self-help and mutual support groups.* Edinburgh, Scotland: Author. Retrieved from https://scottishrecovery.net/wp-content/uploads/2010/09/WRAP-Final-Report.pdf

Shepherd, G., Boardman, J., & Slade, M. (2008). *Making recovery a reality*. London, UK: The Sainsbury Centre for Mental Health.

Sklar, M., Groessl, E. J., O'Connell, M., Davidson, L., & Aarons, G. A. (2013). Instruments for measuring mental health recovery: A systematic review. *Clinical Psychology Review, 33,* 1082–1095. http://dx.doi.org/10.1016/j.cpr.2013.08.002

Slade, M. (2009). *Personal recovery and mental illness: A guide for mental health professionals (values-based practice).* Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511581649

Tew, J., Ramon, S., Slade, M., Bird, V., Melton, J., & Le Boutillier, C. (2012). Social factors and recovery from mental health difficulties: A review of the evidence. *British Journal of Social Work, 42,* 443–460. http://dx.doi.org/10.1093/bjsw/bcr076

Young, S., & Bullock, W. (2003). *The Mental Health Recovery Measure*. Toledo, OH: Department of Psychology, University of Toledo.

Zhang, W., Li, Y., Yeh, H., Wong, S., & Zhao, Y. (2007). *The effectiveness of the Mental Health Recovery (including Wellness Recovery Action Planning) Programme with Chinese consumers*. Retrieved from http://www.tepou.co.nz/file/Knowledge-Exchange-stories/bo-ai-she-the-effectiveness-of-the-mental-health-recovery-research-paper.pdf

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361–370. http://dx.doi.org/10.1111/j.1600-0447.1983.tb09716.x

Received May 29, 2018
Revision received March 13, 2019
Accepted March 24, 2019 ■