PEER-ADMINISTERED INTERVENTIONS FOR DEPRESSION:
A META-ANALYTIC REVIEW

by

Amanda E. B. Bryan

Copyright © Amanda E. B. Bryan 2013

A Dissertation Submitted to the Faculty of the

Department of Psychology

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2013
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Amanda Bryan entitled Peer-Administered Interventions for Depression: A Meta-Analytic Review and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 3/8/2013

Hal Arkowitz

Date: 3/8/2013

John Allen

Date: 3/8/2013

David Sbarra

Date: 3/8/2013

Richard Bootzin

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College. I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 3/8/2013

Dissertation Director: Hal Arkowitz
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Amanda E. B. Bryan
ACKNOWLEDGEMENTS

Many thanks to my advisor, Hal Arkowitz, who has been a supportive and inspiring guide and mentor. Thanks also to my committee members, John Allen, Dave Sbarra, and Dick Bootzin, each of whom has been a wonderful and influential teacher to me over the last several years. My sincerest appreciation to research assistants Karlee Frye, Sean Gallagher, and Joseph Santa Cruz for their generous and capable work on coding the studies included in this analysis. Finally, warmest thanks to my entire family for their love, patience, and support as I’ve found my circuitous way through graduate school.
DEDICATION

For Andrew, and the many years of discovery that await us.
TABLE OF CONTENTS

LIST OF FIGURES........................................................................................................8
LIST OF TABLES...........................................................................................................9
ABSTRACT....................................................................................................................10
1. BACKGROUND AND RATIONALE....................................................................12
   Effectiveness of nonprofessional mental health services.................................13
   Prior meta-analyses of peer interventions for depression.................................15
   Why peer interventions?.........................................................................................17
   The present study....................................................................................................19
2. METHOD...............................................................................................................20
   Study selection.......................................................................................................20
   Data extraction......................................................................................................21
   Effect size calculation............................................................................................22
   Weighted mean effect size calculation..................................................................25
   Moderation analysis...............................................................................................27
   Alternate metrics of outcome................................................................................28
   Follow-up data.......................................................................................................28
   Publication bias.......................................................................................................29
3. RESULTS..............................................................................................................32
   Study characteristics..............................................................................................32
   Depression measurement.......................................................................................32
   Observed effect sizes............................................................................................33
TABLE OF CONTENTS – CONTINUED

Moderation models.................................................................35

Follow-up effect sizes.........................................................37

Outcomes using alternate metrics........................................38

Impact of providing peer support..........................................39

Publication bias......................................................................39

4. DISCUSSION........................................................................41

Summary and interpretation of findings.................................41

Clinical significance of findings............................................44

Methodological limitations....................................................45

Future directions....................................................................46

FIGURES..................................................................................49

TABLES....................................................................................61

APPENDIX A: CODING FORM....................................................73

APPENDIX B: METHODOLOGICAL QUALITY SCORE (MQS)

CODING SYSTEM......................................................................77

REFERENCES...........................................................................78
LIST OF FIGURES

Figure 1: Flow of studies.................................................................49

Figure 2: Forest plot of effect sizes comparing
PAI to active therapy conditions..............................................50

Figure 3: Forest plot of effect sizes comparing
PAI to TAU/WL conditions.....................................................51

Figure 4: Forest plot of effect sizes of PAI conditions
(change from pre- to post-treatment)...........................................52

Figure 5: Moderators of standardized mean gain (SMG) for
PAI conditions (change from pre- to post-treatment)..............53

Figure 6: Moderators of standardized mean difference (SMD)
comparing PAI to active therapy conditions............................54

Figure 7: Moderators of standardized mean difference (SMD)
comparing PAI to TAU/WL conditions.......................................55

Figure 8: Methodological quality is not significantly
associated with pre-post effect size.........................................56

Figure 9: Methodological quality is not significantly
associated with between-groups effect size

(a) PAI compared with active therapy conditions.............57

(b) PAI compared with TAU/WL..............................................58

Figure 10: Funnel plot of effect sizes comparing
PAI to active therapy conditions............................................59

Figure 11: Funnel plot of effect sizes comparing
PAI to TAU/WL conditions..................................................60
**LIST OF TABLES**

Table 1: Descriptive characteristics of clinical trials involving PAIs...............................................................61

Table 2: Characteristics of depression symptom measurement instruments................................................................67

Table 3: Results of included studies.................................................................................................................68

Table 4: Mean SMD estimates (PAI compared with active therapy conditions) adjusted for four models of selection (publication) bias........................................71

Table 5: Mean SMD estimates (PAI compared with TAU/WL conditions) adjusted for four models of selection (publication) bias........................................72
ABSTRACT

A variety of psychotherapies have been demonstrated to be efficacious and effective treatments for depression. The cost of psychotherapy, however, and its low availability in some contexts pose significant treatment barriers for many depressed individuals. Based on the idea that peers (i.e., individuals who have successfully recovered from similar problems) may be uniquely able to provide empathy and support to those currently receiving treatment, some community mental health centers have implemented peer treatment models that employ recovered former clients as cost-effective adjunct providers. The effectiveness of these and other peer-administered interventions (PAIs) has not been well-established. The current study is a meta-analysis of the existing outcome research on PAIs for depression. Twenty-six studies were identified as eligible for inclusion and yielded 30 between-groups effect sizes and 29 pre-post PAI effect sizes. Study characteristics and methodological quality were coded and random-effects models were used to calculate and compare mean effect sizes. PAIs produced significant pre-to-post treatment reductions in depression symptoms that were comparable to those found in well-established professionally-administered interventions (.4554). In direct comparisons, PAIs performed as well as professionally-administered treatments (.0848). but not significantly better than treatment-as-usual (e.g., periodic physician check-ins or availability of community mental health services) and wait-list control conditions (.0978). These findings did not change after adjusting for the moderate degree of publication bias in the data. Moderation models revealed that professionally-co-administered PAIs produced significantly worse outcomes than those that were purely peer-administered,
and that educational/skills-based PAIs (but not supportive PAIs) produced better outcomes compared with professional treatments. Limitations of this analysis included the heterogeneity of the included interventions and the lack of data on mediators and moderators. Still, these findings suggest that PAIs have promise as effective depression treatments and are worthy of further study.
1. BACKGROUND AND RATIONALE

“You never really understand a person until you consider things from his point of view... Until you climb inside of his skin and walk around in it.”

— Harper Lee, To Kill a Mockingbird (1960)

Approximately 16.6% of Americans will meet the diagnostic criteria for Major Depressive Disorder (MDD) during their lives (Kessler et al., 2005). The costs of MDD both to individuals and to society are substantial, with suicide attempt rates 11 times higher in individuals with lifetime diagnoses of depression than in individuals without psychological diagnoses (Kessler, Borges, & Walters, 1999), and an estimated annual economic burden of $83.1 billion incurred through direct, suicide-related, and lost-productivity costs of depression in the United States (Greenberg et al., 2003). Yet just over half of individuals who have met DSM criteria for major depression within the last year have received any treatment, and only about one-fifth have received adequate treatment, defined as receiving pharmacotherapy for at least 30 days with at least four physician visits, or at least eight psychotherapy sessions (Kessler et al., 2003).

Psychotherapeutic treatments for depression have demonstrated comparable efficacy to antidepressant medication and possibly superior outcomes in preventing relapse (Hollon et al., 2002), but the proportion of depressed individuals who receive psychotherapy has declined in recent decades, as has the mean number of psychotherapy sessions for depressed people who do receive it (Olfson et al., 2002). Contributing to this
discrepancy is the high cost of psychotherapy, which decreases the rate at which psychotherapy is utilized in outpatient settings (e.g., Simon et al., 1996), especially among lower-SES individuals (e.g., Steele, Dewa, & Lee, 2007). The expenses of training and reimbursement for master’s- and doctoral-level providers contribute to the costs of therapy, yet it is far from clear that level of training is correlated with treatment quality or outcome (Stein & Lambert, 1984; Bickman, 1999); thus, exploring areas of treatment need that could be fulfilled at a lower cost by providers with less training is worthwhile. Along these lines, it has been suggested (e.g., Pfeiffer et al., 2011) that peer support services could substantially reduce the barriers to psychosocial treatment of depression by making services available at low costs, and by circumventing some of the stigma associated with traditional psychotherapeutic and pharmacological treatments.

Effectiveness of nonprofessional mental health services

Similar to but not synonymous with peer support services, services provided by paraprofessionals or “lay therapists” (i.e., individuals without professional credentials as mental health providers, but not necessarily with personal experiences of the problems they are treating) have shown promise in numerous research studies, suggesting that individuals with little or no prior training as mental health providers may be able to contribute effectively to treatments for depression. Paraprofessional therapists may receive brief, role-specific preparatory training, but do not hold professional credentials as health or mental health professionals, and generally have not received broad or extensive training in mental health treatment. Den Boer, Wiersma, Russo, and van den
Bosch (2005) meta-analyzed controlled studies of paraprofessional treatment of anxiety and depression, and found that the outcomes achieved by paraprofessionals delivering a variety of cognitive-behavioral and supportive interventions were not statistically different from those achieved by professionals, and were significantly better than waitlist and placebo control conditions. In a recent review of cognitive-behavioral treatments for anxiety and depression (Montgomery, Kunik, Wilson, Stanley, & Weiss, 2010), the authors identified four studies that had compared professional and paraprofessional outcomes and concluded that “paraprofessionals generally achieved comparable overall outcomes” (p. 57). These and other summaries (e.g., Durlak, 1979; Hattie, Sharpley, & Rogers, 1984; Berman & Norton, 1985; Christensen & Jacobson, 1994) have concluded it is possible that high-quality treatment can be delivered by individuals who are not trained as therapists.

Taking nonprofessional treatment services a step further, some community mental health organizations have begun to integrate peer “consumer-provided” care – that is, services delivered by current or former clients – into their treatment programs with promising results. Like paraprofessionals, consumer providers typically have neither professional mental health credentials nor a broader training background in a mental health field; they, too, may receive brief, role-specific preparatory training, or they may simply rely on their own treatment experiences for preparation. Chinman, Rosenheck, Lam, and Davidson (2000) evaluated a case management program for seriously mentally ill individuals in which some case managers were consumer providers and others were not mental health care consumers; they found no significant interaction between time and
case manager type on clinical outcomes over 12 months. In a broader investigation of consumer-provided services nationwide, Goldstrom and colleagues (2006) reported on a survey conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA) of mutual support groups, self-help organizations “run by and for mental health consumers,” and consumer-operated mental health services. They found that these groups exceeded in number traditional mental health organizations, and that they served upwards of half a million people per year. Regarding the effectiveness of these services, Doughty and Tse (2011) reviewed 29 studies of consumer-led mental health services and concluded that they produced outcomes similar to the outcomes of traditional mental health services. Taken together, these findings suggest that peer-administered mental health interventions are widely used and possibly effective in improving outcomes.

Prior meta-analyses of peer interventions for depression

In the more specific realm of depression, two recent meta-analyses have evaluated the literature on interventions for depression that focus on the development of relationships with peers (as opposed to with mental health professionals). Pfeiffer and colleagues (2011) analyzed “peer support interventions” for depression, which were defined as interventions that “placed individuals with current depression in regular contact with at least one other person with either current or prior depression” (p. 30). They included 10 controlled studies, and concluded that outcomes of peer support interventions did not differ significantly from outcomes of group CBT, and that the effects of peer support interventions compared with usual care were similar in size to the
effects of CBT compared with usual care in other published analyses. Mead and
colleagues (2010) evaluated the effect on depression symptoms of “befriending,” defined
as “an intervention that introduces the client to one or more individuals whose main aim
is to provide the client with additional social support through the development of an
affirming, emotion-focused relationship over time” (p. 96). They identified 24 studies to
include in their analyses, and found that befriending had a modest effect on depression
symptoms compared with usual care or no treatment, and was less effective than alternate
interventions, including cognitive-behavioral and family-based interventions.

In addition to their somewhat different results, each of these analyses has features
that complicate interpretation. First, in both analyses, treatments that were administered
or facilitated by a mental health professional were included as long as the treatment also
involved some degree of interaction with nonprofessional peers. This makes it difficult
to interpret the extent to which peer support actually had an effect on symptoms,
especially because neither analysis examined professional facilitation as a potential
moderator of treatment outcome. Research has shown (e.g., Toro et al., 1988) that the
presence of a mental health professional in a mutual help group alters the behaviors of
group members and their perceptions of the group.

Second, the inclusion of studies was not comprehensive in either analysis: some
studies were either missed or excluded for reasons that are unclear, and internet-based
interventions were not included in either analysis in spite of other evidence that
computer-based treatments for depression produce effect sizes similar to face-to-face
treatments (e.g., Andersson & Cuijpers, 2009). Additionally, each of these analyses
included only controlled studies, yet a substantial minority of the published studies in this area have been uncontrolled; therefore, many relevant findings were excluded. Although controlled studies produce effect sizes that are easier to interpret than pre-post effect sizes (see p. 23), pre-post comparisons may also be informative when uncontrolled studies comprise much of a body of literature, because there may be systematic differences between controlled and uncontrolled studies that could affect outcome. Researcher allegiance is a hypothetical example: community organizations with a strong allegiance to peer support as a treatment approach may also be less likely than university researchers to have the resources to conduct controlled studies.

Why peer interventions?

Why might one expect interventions administered by non-therapist peers to effectively reduce depression symptoms? One reason is that peer-administered interventions may provide a relationship characterized by an especially high degree of empathy. Although there is a dearth of evidence regarding the effect of empathy on depression symptoms in everyday interactions, it has been demonstrated several times that therapist empathy is associated with better therapy outcome (see Greenberg, Elliot, Watson, & Bohart, 2001, for a brief meta-analysis), and there is evidence that therapist empathy causally affects outcome in cognitive therapy for depression (Burns & Nolen-Hoeksema, 1992). This suggests that receiving empathy improves symptoms, at least in the context of therapy. Yet it is not clear that therapists are better equipped than peers to provide a high degree of empathy. Hassenstab and colleagues (2007) showed that
therapists scored higher than non-therapist control subjects (matched on age, gender, intelligence, and education level) on self-report measures of cognitive empathy (i.e., the ability to understand others’ emotions, thoughts, or intentions), but lower than controls on emotional empathy (i.e., one’s own emotional responses to others’ affective states), suggesting that therapists may actually sacrifice some empathy in favor of emotional control.

On the other hand, peers who are identified specifically because of their similar experiences may be uniquely equipped to provide empathy. Several studies have demonstrated that individuals are more empathetic in response to others’ distress if they themselves have experienced a similar distressing situation. For example, Batson and colleagues (1996) presented subjects with vignettes that described an adolescent experiencing distress about either severe acne or interpersonal rejection. They found that women, but not men, reported experiencing more empathy if they had personally experienced the same type of distress as the adolescent in the vignette they had read. Others have reported the similar effects in women only across different targets of empathy, including fear of darkness, fear of abandonment, and loss of a pet (Eklund, Andersson-Straberg, & Hansen, 2009), and pregnancy (Hodges, Kiel, Kramer, Veach, & Villanueva, 2010). Furthermore, it appears that therapists are no exception: they reported experiencing greater empathy in response to fictional clients if they were able to generate “reference points” from their own experiences that were in some way similar to the clients’ (Hatcher et al., 2005).
The present study

The research summarized above suggests that (1) peers may be especially well-suited to provide social support characterized by a high degree of empathy; (2) receiving empathy may reduce depression symptoms; and (3) previous evaluations of research on peer-support interventions for depression have not been comprehensive and have included studies that were administered primarily by professionals. These findings jointly support the need for the present analysis, which meta-analytically aggregates the published effects of peer-administered interventions on depression symptoms using (a) a highly specific definition of “peer-administered” that explicitly qualifies any involvement by health or mental health professionals, and (b) a comprehensive search strategy to identify studies that have been excluded from previous analyses. Additionally, the analysis reported here includes examinations of several possible moderators of treatment outcome, and evaluates the clinical significance of the findings. Finally, the effects (if any) of administering treatment to peers on well-being will be examined if a sufficient number of studies report data on this.
2. METHOD

Study selection

For this analysis, a peer-administered intervention (PAI) was defined as an intervention administered by one or more lay individuals who have experienced problems similar to those being treated, with minimal or no assistance from a health or mental health professional. Interventions meeting this definition might include a support group for bereaved spouses facilitated by someone who is widowed; an internet support forum for depression in which depressed individuals provide support to each other; and a program in which mothers with prior perinatal depression have regular contact with mothers who are currently experiencing perinatal depression. Under this definition, it is acceptable for peer facilitators to have received supervision or training from a professional in preparation to administer the intervention, and/or ongoing supervision from a professional during the course of administering the intervention. Interventions that were primarily administered by a peer, but had a professional present in a supportive role, were also included in this analysis, and the effect of having a professional present was examined statistically (see Results section).

To be included in this analysis, studies were required to report outcomes in which: (a) at least one condition involved a PAI as defined above; (b) depression symptoms were measured as an outcome; (c) participants were adult outpatients with diagnosed depression or elevated depression scores on a well-validated symptom measure; and (d) the report was published in a peer-reviewed journal and was available in English. Samples were not required to be defined by a diagnosis with depression; studies of
interventions for chronic illness populations or other stressed populations could be included, but only if those samples met minimum clinical cutoffs for probable depression on well-established measures. In order to allow for the most comprehensive review possible, uncontrolled and quasi-experimental studies were included if they reported, at a minimum, pre- and post-treatment outcomes for a PAI condition.

Figure 1 shows the progression of the study selection process. Searches of the PsycINFO and PubMed online databases were conducted in November 2011 using the search terms “(Peer OR mutual OR paraprofessional) AND depress*”. The abstract of each search return was reviewed, and articles that were clearly ineligible (e.g., because they did not report on an intervention study) were discarded, leaving a list of “potentially-eligible” studies. Reference sections of these studies were then hand-searched to identify any additional potentially-eligible studies that the database searches may have missed. Finally, the complete text of each potentially-eligible study was reviewed and its status as eligible or ineligible was determined. For articles that reported on an eligible study but did not include sufficient statistical information to calculate effect size, the needed data were requested from corresponding authors.

Data extraction

A sample coding form, which lists all coded variables, and the Methodological Quality Scale, which raters also completed for all studies, are included as Appendices A and B, respectively.
Each study was independently coded by two raters: the author (Bryan) and a trained undergraduate research assistant. Interrater discrepancies were resolved by discussion to decide values for the final dataset. Interrater reliability for the continuous Methodological Quality Score (MQS) was calculated using Spearman’s rank correlation coefficient (Spearman, 1904), calculated as

$$\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

where $x_i$ and $y_i$ are each rater’s raw ratings converted to ranks. The MQS showed good interrater reliability ($\rho = .82$).

Interrater reliability for categorical variables was calculated using Cohen’s kappa (Cohen, 1960), calculated as

$$\kappa = \frac{Pr(a) - \gamma r(e)}{1 - \gamma r(e)}$$

where $Pr(a)$ is the relative observed agreement among raters, and $Pr(e)$ is the hypothetical probability of chance agreement. Kappas for categorical variables were all at least moderately high, ranging from $.71$-.89.

Effect size calculation

Standardized mean gains. Pre-post effect sizes (standardized mean gains, or SMGs) were calculated for all PAI conditions in controlled and uncontrolled studies, defined as
where $\bar{X}_{T1}$ is the mean at Time 1, $\bar{X}_{T2}$ is the mean at Time 2, $n$ is the common sample size at Time 1 and Time 2, $r$ is the correlation between Time 1 and Time 2 scores, and $s_p$ is the pooled standard deviations of Time 1 and Time 2 means, calculated as

$$s_p = \sqrt{s^2_{T1} + s^2_{T2} / 2}$$

As Westen and Morrison (2001) note, interpretation of the SMG is limited for many of the same reasons as its components: effects of the treatment are likely to be correlated with placebo effects, the effects of time’s passage, and the effects of common factors. At the same time, it may be useful to compare SMGs of PAI conditions with pre-post effect sizes from other published meta-analyses of empirically supported therapies for depression – interpreting these comparisons judiciously, of course.

**Standardized mean differences.** For studies that included randomized control groups, between-groups effect sizes were calculated as *standardized mean differences*, defined as (Lipsey & Wilson, 2001)

$$ES_{sm} = \frac{\bar{X}_{G1} - \bar{X}_{G2}}{s_p}.$$
where $\bar{X}_{G1}$ the posttest mean for Group 1 (comparison group in this analysis), $\bar{X}_{G2}$ is the posttest mean for Group 2 (PAI group in this analysis), and $s_p$ is the pooled posttest standard deviations of the two group means, defined as

$$s_p = \sqrt{\frac{(n_{tx} - )s^2_{tx} + n_{ctrl} - )s^2_{ctrl}}{(n_{tx} - ) + n_{ctrl} - )}.$$  

Because standardized mean differences tend to be upwardly biased in samples of fewer than 20 participants (Hedges, 1981), all standardized mean differences were corrected using Hedges’ correction factor, and these corrected (unbiased) effect size estimates were used in all analyses:

$$ES'_{sm} = \left[1 - \frac{3}{4N - 9}\right]ES_{sm},$$

$$SE_{sm} = \sqrt{\frac{n_{G1} + t_{G2}}{n_{G1}n_{G2}} + \frac{(ES'_{sm})^2}{2(n_{G1} + t_{G2})}},$$

$$w_{sm} = \frac{1}{SE_{sm}^2}.$$  

For this analysis, two sets of SMDs were calculated: those comparing a PAI condition to another active intervention (e.g., a professionally-administered psychotherapy group), and those comparing a PAI condition to a treatment-as-usual (TAU) or wait-list (WL) condition. The decision to combine TAU and WL control conditions into a single meta-analytic comparator was based on the observation that the distinction between the two categories was often blurry or impossible to make in a
systematic way. For example, participants in chronic illness samples almost invariably continued to receive standard medical treatment during the course of the study, even if they were in a wait-list condition for the PAI or other psychosocial interventions; these standard medical treatments may or may not have involved “therapeutic” interactions with health care professionals. Similarly, across sample types, conditions described as “wait-list control” often included periodic check-ins or “case management” visits that may have had therapeutic value. Conversely, in conditions described as TAU or “standard care,” often this simply meant that ongoing care of some kind was available to participants, not necessarily that it was utilized during the study period. The difficulty of distinguishing between TAU and WL conditions in this sample was corroborated by low interrater reliability between the two (κ = .25). Thus, rather than trying to categorize based on incomplete data or arbitrary differences, the two were combined as a single category. This had the added benefit of increasing statistical power for comparisons between PAIs and TAU/WL conditions.

Weighted mean effect size calculation

In order to summarize effect sizes across studies, weighted mean effect size estimates were calculated separately for between-groups ESs (SMDs) and within-groups ESs (SMGs) using the formula

$$\overline{ES} = \frac{\sum (w_i ES_i)}{\sum w_i}$$
where $w_i$ is the inverse variance weight for $ES_i$. Upper and lower limits of the 95% confidence intervals around mean effect sizes were calculated as:

$$\bar{ES} \pm t_{\alpha/2} \cdot (SE_{\bar{ES}})$$

where $SE_{\bar{ES}}$ is the standard error of the mean effect size, calculated as

$$SE_{\bar{ES}} = \sqrt{\frac{1}{\sum w_i}}.$$

To determine whether the effect sizes included in these means all estimate the same population effect size (e.g., Hedges, 1982, Rosenthal & Rubin, 1982), the heterogeneity statistic $Q$ was computed for each mean effect size:

$$Q = \frac{\left(\sum w_i ES_i^2\right) - \left(\sum w_i ES\right)^2}{\sum w_i}.$$

$Q$ falls in a chi-square distribution with $k - 1$ degrees of freedom, where $k$ is the number of effect sizes included in its calculation (Hedges & Olkin, 1985). If $Q$ is significant for a given mean effect size (i.e., the null hypothesis of homogeneity is rejected), this indicates that the dispersion of effect sizes around that mean is greater than would be expected from sampling error alone (e.g., that differences in particular study characteristics account for some of the variance in effect size). In these cases, it may be appropriate to use random effects models, which use an adjusted inverse variance weight based on the assumption that in addition to subject-level sampling error, there is an additional randomly-distributed study-level source of error.
Moderation analysis

First, fixed-effects (ordinary least squares) models were calculated for all hypothesized moderators using David Wilson’s MetaReg and MetaF macros for SPSS (Wilson, 2010). Fixed-effects models partition effect-size variance into two categories: between-groups variance (i.e., variance explained by the moderator), and within-group variance (i.e., residual unexplained variance, which is assumed to be subject-level sampling error). If $Q_w$, the heterogeneity statistic for the residual variance, is significant, this indicates that modeling only systematic variance and subject-level sampling error (as in a fixed-effects model) is insufficient (i.e., fails to meet the model’s assumptions that these are the only sources of variability), and a mixed-effects model may be appropriate (Lipsey & Wilson, 2001).

Because fixed-effects models produced significant $Q_w$ values for all of the mean effect sizes (see Results section), we turned to mixed-effects models using method-of-moments estimation procedures, which Wilson’s (2010) macros can also calculate. Mixed-effects models assume “that there is a remaining unmeasured (and possibly unmeasurable) random effect in the effect size distribution in addition to sampling error. That is, variability in the effect size distribution is attributed to systematic (modeled) between-study differences, subject-level sampling error, and an additional random component” (Lipsey & Wilson, 2001, p. 124). Mixed-effects models have lower statistical power to detect moderation effects and result in wider confidence intervals than fixed-effects models; they are, however, less prone to Type I errors, and their assumptions are not violated by heterogeneity as the assumptions of fixed-effects models
are (Lipsey & Wilson, 2001). Thus, they were deemed more appropriate for testing moderation in this set of data.

**Alternate metrics of outcome**

In addition to conventional effect size estimates, two alternate metrics of outcome were examined to the extent that they were reported or available in order to determine whether any treatment effects were clinically (not just statistically) significant. Clinical significance has been defined as “the extent to which therapy moves someone outside the range of the dysfunctional population or within the range of the functional population” (Jacobson & Truax, 1991, p. 12), and is an important consideration when determining whether detected effects represent meaningful changes in symptomology. First, percentage of participants considered “improved” at treatment termination was coded when reported. Second, posttreatment symptomology (i.e., mean scores on outcome measures) was compared to published or widely-used guidelines for what constitutes clinical elevation on the respective measures of depression symptoms.

**Follow-up data**

When available, follow-up data were examined to evaluate the durability of treatment effects. SMGs and SMDs were calculated in the same manner as for pre-post data (with SMGs indicating effect size between post-test and follow-up) When a study
reported on multiple follow-up time points (e.g., three- and six-month follow-ups), data from the last available (e.g., the six-month follow-up) were used here.

For each set of follow-up effect sizes, two variables were tested as possible moderators using the same procedures as described previously. First, the length of the follow-up period was tested, as it was possible that shorter follow-ups would indicate better maintenance simply because there was less time for symptoms to recur (and not because interventions had a particularly enduring effect). Second, effect size at posttreatment was tested as a moderator to determine whether PAIs that were more effective in the short-term also produced more enduring benefits.

Publication bias

Several methods were considered for examining publication bias. The two most commonly-reported indicators of publication bias, the funnel plot and the fail-safe N, each have features that make reliable and valid interpretation difficult (Sutton, 2009), rendering them less-than-ideal as standalone indicators. Funnel plots (Light & Pillemer, 1984) are scatterplots of effect size on the x-axis and standard error on the y-axis; in data without publication bias, these plots are generally pyramid-shaped and symmetrical. Interpretation of funnel plots typically relies on visual inspection for asymmetry. Although funnel plots can be useful for developing hypotheses about publication bias, interpreting them objectively and/or quantitatively presents challenges; furthermore, when study characteristics moderate effect sizes, funnel plots can be misleading because they effectively superimpose effect sizes drawn from multiple populations (Vevea &
Woods, 2005). The fail-safe N (Rosenthal, 1979; Orwin, 1983) is “the number of studies with an effect size of zero needed to reduce the mean effect size to a specified or criterion level” (Lipsey & Wilson, 2001, p. 166). This method has been criticized for its assumption that all missing studies have effect sizes of zero, which is unlikely to be reliably true (Vevea & Woods, 2005). Additionally, there is no theoretically-based criterion for determining a threshold number of missing studies as indicative of publication bias, and such a number may not be particularly informative anyway; as Field and Gillett (2010) write, “…any fail-safe N method addresses the wrong question: it is usually more interesting to know the bias in the data one has and to correct for it than to know how many studies would be needed to reverse a conclusion” (p. 686). One commonly-used method for correcting for publication bias, the iterative “trim-and-fill” technique (Duval & Tweedie, 2000), produces a mean effect size corrected for the number of studies estimated to be missing; this method, however, can lead to overcorrection, particularly in datasets that contain heterogeneity, because it assumes that all of the missing studies are those with the smallest effect sizes (Vevea & Woods, 2005).

Following the recommendation of Field and Gillett (2010), the analysis reported here utilized Vevea and Woods’s (2005) weight-function sensitivity analysis to estimate and correct for publication bias. Generally speaking, weight-function models estimate the influence of different study characteristics on the likelihood of publication and make corrections that are weighted according to these influences, thus avoiding the assumption that missing studies are simply those with the smallest effect sizes. Typically, weight-function models require very large samples sizes ($k > 100$; e.g., Vevea & Hedges, 1995),
but Vevea and Woods (2005) offer a modification for use with smaller-scale meta-
analyses that “impose[s] a set of fixed weights determined a priori and chosen to
represent a specific form and severity of biased selection” (p. 432). This procedure
provides corrected effect size estimates for four different models of selection bias,
allowing the user to judge the most likely conditions of publication bias in his data (e.g.,
based on visual inspection of funnel plots) and select the model that most closely
represents those conditions. The models reported in this analysis were calculated using
the macros provided by Field and Gillett (2010).
3. RESULTS

*Study characteristics*

Search procedures identified 26 studies meeting inclusion criteria for which sufficient data to calculate at least one effect size were either reported or obtainable from authors. Included studies were published between 1988 and 2011 and reported on a total of 29 PAI conditions. Twenty studies included at least one control group; in two of these studies, however, participants were not randomly assigned to conditions, and only pre-post data from the PAI conditions from those studies were included in this analysis. In all, this yielded 30 effect sizes of PAI compared with another condition, and 29 pre-post PAI effect sizes.

In all included studies, participants’ average pre-treatment depression scores met minimum clinical cutoffs for likely diagnosable depression. In nine studies, elevation in depression symptoms was the primary defining feature of the sample. Eleven studies’ samples were defined primarily as having a chronic illness, and the remaining six were identified primarily by the experience of a severe psychosocial stressor. Additional study characteristics are summarized in Table 1.

*Depression measurement*

Table 2 describes the measures of depression symptoms that were used in included studies, including frequency of use and threshold scores for probable depression.
**Observed effect sizes**

Results of included studies are summarized in Table 3. Effect sizes are shown in Figures 2-4.

**Standardized mean gains (pre-post effect sizes).** It should be noted that, although improvement in depression symptoms is indicated by a decreased score on all included depression measures, for the sake of consistency all signs were reversed so that a positive SMG indicates that symptoms improved from pre- to post-treatment.

The mean SMG for all PAI conditions using a fixed-effects model was .2194 (N = 29, \( p < .0001 \), 95% CI: .2003 - .2385), indicating that on average, depression scores decreased between pre- and post-treatment in PAI conditions. However, the heterogeneity statistic was highly significant (\( Q = 890.84, df = 28, p < .0001 \)), suggesting that the effect sizes comprising this mean do not all estimate the same population mean. Using a random-effects model to account for the heterogeneity increased the estimated mean SMG to .4554 (\( p < .0001 \), 95% CI: .3391 - .5717). As a standalone effect size indicator, this estimate must be interpreted cautiously.

**Standardized mean differences (between-groups effect sizes).** It should be noted that the PAI condition was considered “Group 2” for all SMD calculations; therefore, a positive SMD indicates that PAI conditions produced fewer symptoms (i.e., better outcomes) at post-treatment than control conditions.

Using a fixed-effects model, the mean SMD comparing PAI conditions to active control conditions (i.e., other, non-PAI therapies) was .0306 (N = 13, \( p = .6664 \), 95% CI: -.1085 - .1696), indicating that posttreatment depression scores were not significantly
different between PAI and active-control conditions. The heterogeneity statistic was significant ($Q = 30.15$, $df = 12$, $p = .0027$), suggesting greater variability than would be expected from sampling error if all effect sizes were estimates of the same population mean; thus, a random-effects model may be more accurate. The mean PAI-vs-active control SMD using a random-effects model was $0.0848$ ($p = .4706$, 95% CI: $-0.1455$ - $0.3151$), indicating that the effect size did not achieve statistical significance even after accounting for the heterogeneity in the data.

Using a fixed-effects model, the mean SMD comparing PAI conditions to TAU/WL conditions was $0.0913$ ($N = 17$, $p = .0235$, 95% CI: $0.0123$ - $0.1703$), indicating that posttreatment depression scores were significantly lower in PAI conditions than in TAU/WL conditions. Again, the heterogeneity statistic was significant ($Q = 72.88$, $df = 16$, $p < .0001$), suggesting greater variability than would be expected from sampling error if all effect sizes were estimates of the same population mean; thus, a random-effects model may be more accurate. The mean PAI-vs-TAU/WL SMD using a random-effects model was $0.0978$ ($p = .2966$, 95% CI: $-0.0859$ - $0.2815$), indicating that the effect size did not maintain statistical significance after accounting for heterogeneity. Given the direction and small magnitude of the difference in effect size between the fixed- and random-effects models, this discrepancy in statistical significance is more likely an artifact of low statistical power than a meaningful indicator about the effect itself; that is, although the heterogeneity in the data was statistically significant, its impact on effect size was insufficient to support the statistical power demands of the random-effects procedure.
In summary, these mean effect size estimates suggest that PAI interventions produced significant reductions in depression symptoms from pre- to post-test, performed similarly to the other active therapies to which they were compared, and did not significantly outperform TAU/WL conditions. The significant heterogeneity in all three estimates, however, suggests the presence of other significant sources of variability (e.g., moderators) that need to be considered in a valid interpretation.

**Moderation models**

Hypothesized moderators were PAI format (group or one-on-one), professional coadministration of PAI (yes or no), sample type (depression, stressor, or chronic illness), mutuality (mutual support or one person supporting another), training of peer administrators (yes or no), professional supervision of peer administrators (yes or no), PAI content (supportive or psychoeducational), PAI delivery mode (face-to-face, telephone, or internet), PAI meeting frequency (weekly, more than weekly, or less than weekly), blindness of assessments (yes or no), researcher allegiance (for or against PAI), and total methodological quality score (MQS). Each of these constructs was tested as a possible moderator of SMGs and SMDs, with two exceptions: researcher allegiance, for which there was insufficient variability to test for moderation (i.e., all but two studies expressed allegiance to PAI); and blindness of assessments, for which there were too many missing data to test for moderation (only 56% of controlled studies reported on blinding).
Results of mixed-effects moderation models are displayed in Figures 5-9. Pre-post (SMG) models revealed a significant moderation effect of professional coadministration, such that PAIs that were coadministered by health or mental health professionals were less effective in reducing depression symptoms between pre- and post-test than PAIs that were entirely peer-administered. There was also a significant omnibus test for supervision, but this is unlikely to be meaningful since post-hoc tests indicated that the significant difference was between non-supervised groups and groups in which supervision status was unspecified. Finally, there was an almost-significant ($p = .0546$) omnibus test for sample type, such that pre-post improvements were greater for depression samples than for stressor samples (with neither differing significantly from chronic illness samples). There were no other variables that significantly moderated SMG.

Mean SMD models comparing PAIs with active comparison conditions indicated a significant moderation effect of PAI content, such that PAIs whose content was primarily educational performed better than their active comparison groups, whereas PAIs whose content was primarily supportive did not perform significantly differently than their active comparison groups. There was also a moderating effect approaching significance of meeting frequency ($p = .0981$), such that PAIs meeting weekly performed better than their active comparison conditions, whereas PAIs with variable or unspecified meeting frequencies performed no differently than their active comparison conditions. There were no other significant moderators of SMDs between PAI and active control conditions.
There were no significant moderators of SMDs between PAI and TAU/WL comparison conditions.

*Follow-up effect sizes*

Eight studies reported sufficient follow-up data to calculate at least one effect size. (A ninth study [Tudiver et al., 1992] reported follow-up depression scores, but no standard deviations or other estimate of variability was available, precluding the calculation of an effect size.) This resulted in eight follow-up SMGs comparing outcomes at post-treatment to outcomes at follow-up, and nine follow-up SMDs (five comparing PAI to active comparison, four comparing PAI to TAU/WL). Follow-up periods ranged from 3-12 months (M = 6.89, SD = 3.62).

The mean SMG from post-treatment to follow-up was .0016 (p = .99), suggesting that improvements in depression symptoms were maintained (but did not significantly increase) after treatment ended. This effect was not moderated either by follow-up length (p = .9214) or post-treatment SMG (p = .9501).

The mean SMD between PAI and active comparison conditions was .0477 (p = .6322), indicating that PAI and active comparisons remained equally effective at follow-up. This effect was not significantly moderated either by follow-up length (p = .3922) or by post-treatment SMD (p = .9777).

The mean SMD between PAI and TAU/WL conditions was .1226 (p = .0744), which approached significance and, if significant, would indicate that PAI conditions had maintained gains better (or deteriorated less) at follow-up compared with TAU/WL
conditions. This effect was not significantly moderated by follow-up length \((p = .6505)\), but it was significantly moderated by SMD at posttest \((p = .0012)\), such that PAIs that had performed better than TAU/WL comparisons at posttest also performed better at follow-up.

**Outcomes using alternate metrics**

Only one of the 26 studies included in this meta-analysis (Bright et al., 1999) reported an estimate of the percentage of participants who were considered improved after treatment, precluding the aggregated use of this metric as an indicator of clinical significance across studies.

Another indicator of clinical significance is whether post-treatment depression scores fell below the thresholds for clinical elevation for the measures utilized. (Table 2 provides the elevation threshold for each measure included in this analysis.) In theory, at least, falling below such a threshold would mean that participants improved enough that their symptom levels at post-treatment were more similar to non-clinical than clinical population means. Of the 29 PAI conditions for which they were available, post-treatment depression means fell below the elevation threshold for 10 conditions, indicating that in almost two-thirds of the PAI conditions studied, participants were still experiencing significant symptoms after receiving treatment. It should be noted that most of these post-treatment means were well within one standard deviation of their thresholds (and in many cases, fell mere fractions of points above or below); furthermore, this indicator is only informative to the extent that variability is low (i.e., that a condition’s
mean outcome is a good representation of a typical participant’s outcome in that condition). Therefore, it should be interpreted cautiously. To the extent that it is accurate, it suggests that PAIs alone may not have been sufficient – at least in the doses provided – to reduce symptoms to below a clinical level.

*Impact of providing peer support*

Only one of the studies included in this meta-analysis (Giese-Davis et al., 2006) reported on any indicators of well-being for peers providing intervention. In that study, peer “Navigators” (breast cancer survivors) who provided support to women newly diagnosed with breast cancer did not show changes in depression symptoms or well-being over the course of the study (though they did indicate increased dissatisfaction with their medical teams). The lack of other outcome measures for peer supporters ruled out aggregate analysis of the impact on peers of providing support.

*Publication bias*

Visual inspection of the funnel plot for PAI-vs-active therapy effect sizes (Figure 10) suggests a lack, but not a complete absence, of effect sizes on the left side of the funnel (i.e., negative effect sizes, those that would indicate superiority of other therapy over PAI). Therefore, the moderate one-tailed selection bias model (Table 4) was the most likely to fit the data, and indicated that adjusting for publication bias would produce a mean SMD of -.017 using a fixed-effects model (compared with an unadjusted SMD of .031), or a mean SMD of -.014 using a random-effects model (compared with an unadjusted SMD of .083). These adjustments suggest that the true mean SMD may be
closer to zero than the observed mean SMD. The observed mean SMD, however, was not statistically different from zero; thus, the finding does not change when publication bias is accounted for.

Visual inspection of the funnel plot for PAI vs TAU/WL conditions (Figure 11) suggests a possible, but not dramatic, lack of effect sizes on the far left side of the funnel (i.e., negative effect sizes, those that would indicate superiority of TAU/WL over PAI). Therefore, here too the moderate one-tailed selection bias model (Table 5) was the most likely to fit the data, and indicated that adjusting for publication bias would produce a mean SMD of .061 using a fixed-effects model (compared with an unadjusted SMD of .091), or a mean SMD of .006 using a random-effects model (compared with an unadjusted SMD of .102). These adjustments suggest that the true mean SMD is closer to zero than the observed mean SMD (possibly enough closer that it would not be statistically different from zero).

In summary, it appears that there was a modest degree of publication bias in this dataset, such that studies showing poorer outcomes for PAIs were less likely to be published. Adjusting for this bias did not change the interpretation of comparisons between PAIs and active therapy conditions; the outcomes remained statistically equivalent. Comparisons between PAI and TAU/WL conditions also remained equivalent after this adjustment. Therefore, the practical impact of this adjustment is minimal (see Discussion, pp. 43-44).
4. DISCUSSION

**Summary and interpretation of findings**

The meta-analysis reported here examined the effects of peer-administered interventions (PAIs) on symptoms of depression, defined as interventions administered by one or more lay individuals who have experienced problems similar to those being treated, with minimal or no assistance from a health or mental health professional. This is the first analysis to comprehensively aggregate outcomes from studies of PAIs for depression; prior meta-analyses of peer interventions for depression have suffered from incomplete identification/inclusion of eligible studies and from overly-broad inclusion criteria (e.g., including interventions that were primarily administered by professionals). Effect size comparisons indicated that PAIs reduced depression symptoms comparably to other active interventions (e.g., professionally-administered psychotherapies), but did not significantly outperform TAU/WL conditions. Overall, these findings suggest that PAIs may be effective, cost-efficient treatments for depression; their effectiveness, however, is not unambiguous.

This is also the first meta-analysis to examine potential moderators of PAI outcomes. Moderation models revealed that PAIs that were co-administered by health or mental health professionals were less effective in reducing depression symptoms between pre- and post-treatment than PAIs that were entirely peer-administered. Contrary to the intuitive perspective that additional expertise should carry additional benefit, this finding suggests that a “pure” approach to peer intervention may be preferable. It is possible that the involvement of a professional undermined an important mechanism of peer change;
for example, perhaps peers were seen as less knowledgeable when a professional co-administrator was also present, reducing their impact as strong, capable role models of successful recovery. Given that this moderation effect pertained to pre-post comparisons, it is also possible that it was an artifact of a third variable unrelated to mechanisms of therapeutic change, such as researcher allegiance. For example, researchers who included professionals in their “peer” interventions may have started out with low confidence in the abilities of peers to be independently therapeutic. Alternately, researchers who knew that professionals would be involved may have seen less value in dedicating time and resources toward preparing peers to intervene effectively.

There was also a moderating effect of intervention content, such that PAIs that were psychoeducational or skills-based (e.g., teaching basic cognitive-behavioral skills) performed better than other active therapies to which they were compared, whereas PAIs that were primarily supportive did not perform differently than other active therapies. There are several plausible interpretations of this finding, including: (1) that gaining education and skills directly impacted depression symptoms more than receiving support; (2) that peers administering educational/skills-based interventions were equally supportive and empathic as those delivering primarily-supportive PAIs, resulting in additive effects of skills and support; or (3) that peers delivering educational/skills-based interventions were more likely to have undergone a brief preparatory training program to prepare them for their roles, which might have resulted in a variety of advantages, including higher confidence and better interpersonal skills.
It is also notable that PAIs performed similarly to other active therapies, yet did not outperform TAU/WL conditions. Initially, this might seem to suggest that neither PAIs nor active therapy comparisons were especially helpful. However, examination of standardized mean gains (pre-post ESs) indicates otherwise: it appears that symptoms tended to improve substantially between pre- and post-treatment not only for PAIs and other active therapies, but also for participants in TAU/WL conditions (SMG\textsubscript{PAI} = .46; SMG\textsubscript{ACTCTRL} = .55; SMG\textsubscript{TAU/WL} = .32). This suggests that the failure of PAIs to perform much better than TAU/WL conditions may be because participants improved regardless of condition. One possible explanation for the high degree of improvement among TAU/WL participants is that a substantial minority of included studies involved depressed chronically ill populations, whose “usual” or “standard” care was likely to involve regular contact with medical professionals. These contacts may, in some cases, have been psychologically therapeutic, even if that was not their intended function. A second explanation pertains to the fact that these analyses are based on study completers; participants who remained in a study despite being assigned to a control condition may have been highly motivated to change and/or less distressed to begin with, and thus likely to improve whether or not they received treatment.

Publication bias appeared to be modestly present in the data, such that between-groups comparisons that did not favor PAIs were somewhat less likely to appear in the published literature. When effect sizes were adjusted for the likely degree of publication bias, PAIs and other active therapies maintained their equivalent effectiveness, and the difference in outcome between PAI and TAU/WL conditions remained non-significant.
Therefore, although it is notable that there was bias in the data, it posed little complication to the interpretation of the findings.

**Clinical significance of findings**

The mean pre-post effect size (standardized mean gain) observed across PAI conditions was statistically greater than zero, but statistical significance does not necessarily imply clinical significance. What, if anything, does that number say about the extent to which PAI participants experienced meaningful symptom reduction? There are multiple ways to address this question.

First, we can calculate how an effect size translates to score reductions on the symptom measures that were employed. Using the pooled standard deviation observed in this sample of studies, the .4554 SMG is equivalent to a reduction of 3.54 points on the BDI-II, or 4.41 points on the CES-D (by far the two most commonly-used measures in this analysis). To put those numbers in context, we can consider them as minimal proportions of the threshold for clinical elevation. For the BDI-II, scores ≥14 were considered elevated; thus, a 3.54-point reduction would constitute a 25.3% minimum reduction in symptoms. For the CES-D, scores >15 were considered elevated; thus, a 4.41-point reduction would constitute a minimum 29.4% reduction in symptoms. This provides a conservative estimate of proportional score reductions for PAI participants (since symptom means often exceeded the threshold considerably). Still, these average raw score reductions are small in absolute terms – that is, they represent relatively small changes in actual symptom presentation.
Second, we can use as benchmarks the pre-post effect sizes observed in recently-published meta-analyses of well-established, empirically-supported therapies for depression. Large standardized mean gains have been observed in meta-analyses of a variety of efficacious and possibly-efficacious (Chambless & Hollon, 1998) therapies for adult unipolar depression, including but not limited to cognitive-behavioral therapy ($d = 1.06$; Hans & Hiller, 2013), short-term psychodynamic psychotherapy ($d = 1.34$; Driessen, Cuijpers, Maat, Abbass, de Jonghe, & Dekker, 2010), and acceptance and commitment therapy (Hedges’ $g = .59$; Hofmann, Sawyer, Witt, & Oh, 2010). These previously-observed standardized mean gains are more than twice those observed in this sample of PAIs. These comparisons suggest that although PAI participants experience reduced symptoms, these symptom reductions are inferior to those seen in other, well-established psychotherapies for depression.

**Methodological limitations**

A significant limitation of this analysis was the heterogeneity of the interventions falling under the PAI umbrella. All had the common feature of being peer-administered, but they differed substantially in structure, duration, and content. Although every attempt was made to code these differences and examine their moderating effects, it is possible that some important moderators were missed or unavailable, or that significant moderation effects did not achieve significance because of low statistical power. Furthermore, the lack of clear distinction between TAU and WL conditions, and the
resulting need to combine them as a single comparison category are less than ideal, as there was substantial variety in these conditions.

Another limitation was the lack of available data on potential moderators and mediators of PAI outcomes. Although methodological features were coded and tested as moderators, participant-level variables beyond basic demographics were largely unavailable, let alone measured consistently enough across studies to test meta-analytically. For example, race and ethnicity were too inconsistently reported in these studies to be tested as moderators, yet it is probable that racial and cultural factors impact clients’ views on help-seeking generally, and their attitudes toward utilizing peers versus professionals to manage distress (e.g., Diala, Muntaner, Walrath, Nickerson, LaVeist, & Leaf, 2001; Sheu & Sedlacek, 2004; Townes, Chavez-Korell, & Cunningham, 2009; Chen & Mak, 2008). Mechanisms of action underlying PAIs’ effectiveness were, for the most part, neither tested nor speculated upon in this literature. One obvious example is empathy: the capacity of peers to empathize with similar others’ distress is a hypothesized process by which PAIs might be effective, yet too few studies measured empathy or any other relationship or process-level variables to examine this hypothesis. Thus, the currently-available literature largely leaves us to speculate about how, why, and for whom PAIs are beneficial.

Future directions

Given the limitations details above, the literature on PAIs for depression would benefit from future research examining for whom, under what circumstances, and via
what processes PAIs are helpful for depression. For example, in this analysis, there were not statistically significant pre-post or between-groups differences in outcome for the different sample types (depressed, stressed, or chronically ill), but there was a consistent pattern of smaller effect sizes for stressor samples that would be worth investigating further with a larger sample size. Furthermore, from a practical standpoint, it would be useful to know whether PAIs function best as standalone interventions or as adjuncts to traditional treatments (and, if as adjuncts, whether best delivered concomitantly or as relapse prevention). Finally, in addition to empathy, other likely mediating processes might include recovery modeling, the provision of hope, and specific techniques or coping strategies.

Another distinction remains to be addressed. PAIs may be (1) treatments that are already supported as professionally-administered interventions, but now administered by peers; or (2) interventions that are specifically designed for peers to administer, without prior empirical support as professionally-administered interventions. The former may be more justifiable in terms of research funding; at the same time, interventions designed specifically for peers may have therapeutic qualities that traditional interventions miss. This distinction is worthy of attention in future research.

Another area of interest for future research is the effect on peers of delivering a PAI. Ideally, helping others with problems one has personally overcome would reinforce recovery or at least provide a subjectively positive experience. There is some evidence that peer counselors have experienced increased well-being, increased confidence, and decreases in mild depression symptoms (e.g., Garcia, Metha, Perfect, & McWhirter,
1997; de Vries & Petty, 1992; Byrd, 1984), but this evidence is meager and methodologically limited. Moreover, we know little about how individual differences among peer supporters might influence these outcomes. For example, for peers who are currently depressed or at higher risk for relapse, spending time with other depressed individuals and exerting energy in a helping capacity could be emotionally draining, increase the probability of relapse, or prolong symptoms. On the other hand, for those with solid, durable recovery experiences, supporting peers may provide gratification, connectedness, a sense of purpose, or other positive emotional states that could buffer against relapse and improve quality of life.
FIGURES

Figure 1: Flow of studies

Search terms:
“(Peer OR mutual OR paraprofessional) AND depress*”

PsycINFO: 3344 results
Medline: 1491 results

4723 results excluded because non-eligibility was clear from abstract (e.g., not an empirical study or not an intervention study)

112 potentially-eligible studies for review from database searches

33 potentially-eligible studies for review from hand-search of reference sections

145 potentially-eligible studies

26 studies met inclusion criteria

20* controlled
6 uncontrolled

*Includes two non-randomized controlled studies, which were not used to calculate between-groups differences.
Figure 2: Forest plot of effect sizes comparing PAI to active therapy conditions

<table>
<thead>
<tr>
<th>First Author/Year</th>
<th>N</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ludman, 2007</td>
<td>93</td>
<td>Depression</td>
</tr>
<tr>
<td>Bright, 1999</td>
<td>134</td>
<td>Depression</td>
</tr>
<tr>
<td>McKay, 2002</td>
<td>200</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Onrust, 2010</td>
<td>188</td>
<td>Bereavement</td>
</tr>
<tr>
<td>Marnar, 1988</td>
<td>55</td>
<td>Bereavement</td>
</tr>
<tr>
<td>Anderson, 2005</td>
<td>85</td>
<td>Depression</td>
</tr>
</tbody>
</table>

Effect Sizes (Cohen's d)
Figure 3: Forest plot of effect sizes comparing PAI to TAU/WL conditions

<table>
<thead>
<tr>
<th>First Author/Year</th>
<th>N</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nagal, 1999</td>
<td>90.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Robinson-Whelan, 2007</td>
<td>65.0</td>
<td>Dep + Chronic Illness</td>
</tr>
<tr>
<td>Long, 2009</td>
<td>204.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Dennis, 2009</td>
<td>613.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Gifford, 1998</td>
<td>69.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Hill, 2008</td>
<td>100.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Ludman, 2007</td>
<td>45.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Helfer, 1991</td>
<td>384.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Simoni, 2007</td>
<td>105.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Gotay, 2007</td>
<td>252.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Silver, 1997</td>
<td>343.0</td>
<td>Child's Illness</td>
</tr>
<tr>
<td>Letourneau, 2011</td>
<td>51.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Tudor, 1992</td>
<td>94.0</td>
<td>Bereavement</td>
</tr>
</tbody>
</table>
Figure 4: Forest plot of effect sizes of PAI conditions (change from pre- to post-treatment)

Effect sizes of PAI conditions (change from pre- to post-treatment)

<table>
<thead>
<tr>
<th>First Author/Year</th>
<th>N</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lieberman, 2005</td>
<td>91.0</td>
<td>Chronic illness</td>
</tr>
<tr>
<td>Bright, 1999a</td>
<td>14.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Bright, 1999b</td>
<td>12.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Letourneau, 2011</td>
<td>23.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Dennis, 2009</td>
<td>297.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Nagel, 1968a</td>
<td>20.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Nagel, 1968b</td>
<td>20.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Ludman, 2007</td>
<td>22.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Freyde, 2003</td>
<td>24.0</td>
<td>Premature birth</td>
</tr>
<tr>
<td>Lieberman, 2006</td>
<td>52.0</td>
<td>Chronic illness</td>
</tr>
<tr>
<td>Robinson-Whealan, 1997</td>
<td>42.0</td>
<td>Dep + Chronic Illness</td>
</tr>
<tr>
<td>Travis, 2010</td>
<td>32.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Giese-Davis, 2006</td>
<td>19.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Long, 2000</td>
<td>181.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Hill, 2006</td>
<td>43.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Roman, 2005</td>
<td>22.0</td>
<td>Premature birth</td>
</tr>
<tr>
<td>Gotay, 2007</td>
<td>128.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Christ, 2010</td>
<td>93.0</td>
<td>Bereavement</td>
</tr>
<tr>
<td>Simoni, 2007</td>
<td>54.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Anderson, 2005</td>
<td>46.0</td>
<td>Depression</td>
</tr>
<tr>
<td>McKay, 2002a</td>
<td>30.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>McKay, 2002b</td>
<td>33.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Heller, 1991b</td>
<td>46.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Uccelli, 2004</td>
<td>42.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Silver, 1997</td>
<td>174.0</td>
<td>Child's Illness</td>
</tr>
<tr>
<td>Marmar, 1990</td>
<td>26.0</td>
<td>Bereavement</td>
</tr>
<tr>
<td>Heller, 1991a</td>
<td>46.0</td>
<td>Depression</td>
</tr>
<tr>
<td>Gifford, 1998</td>
<td>25.0</td>
<td>Chronic Illness</td>
</tr>
<tr>
<td>Cheung, 2000</td>
<td>65.0</td>
<td>Depression</td>
</tr>
</tbody>
</table>
Figure 5: Moderators of standardized mean gain (SMG) for PAI conditions (change from pre- to post-treatment)

Darker bars denote effect sizes that are significantly different from zero.

** p<.01
* p<.05
† p<.1
Figure 6: Moderators of standardized mean difference (SMD) comparing PAI to active therapy conditions

Darker bars denote effect sizes that are significantly different from zero.

** p < .01
* p < .05
† p < .1
Figure 7: Moderators of standardized mean difference (SMD) comparing PAI to TAU/WL conditions

Darker bars denote effect sizes that are significantly different from zero.

** p<.01
* p<.05
† p<.1
Figure 8: Methodological quality is not significantly associated with pre-post effect size.
Figure 9a: Methodological quality is not significantly associated with between-groups effect size (PAI compared with active therapy conditions)
Figure 9b: Methodological quality is not significantly associated with between-groups effect size (PAI compared with TAU/WL)
Figure 10: Funnel plot of effect sizes comparing PAI to active therapy conditions.
Figure 11: Funnel plot of effect sizes comparing PAI to TAU/WL conditions
Table 1: Descriptive characteristics of clinical trials involving PAIs

<table>
<thead>
<tr>
<th>1st Author, Year</th>
<th>Setting/ format</th>
<th>Inclusion criteria</th>
<th>Depression measure*</th>
<th>PAI(s)</th>
<th>Control condition(s)</th>
<th>Duration of intervention</th>
<th>Follow up? (months post tx end)</th>
<th>MQS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson, 2005</td>
<td>Internet/ group</td>
<td>Swedish adults with likely mild-to-moderate unipolar depression (per CIDI, and MADRS-S score 15-30)</td>
<td>BDI</td>
<td>Moderated online discussion group; therapist-monitored</td>
<td>Online self-help CBT program + moderated online discussion group</td>
<td>M = 10 weeks</td>
<td>6 months</td>
<td>8</td>
</tr>
<tr>
<td>Bright, 1999</td>
<td>Outpatient psychology dept clinic/ group facilitated by prior participants in a community-based self-help org</td>
<td>Memphis area adults scoring ≥10 on HRSD and meeting SCID-NP criteria for MDE, dysthymia, or dep NOS</td>
<td>HRSD [BDI]</td>
<td>(1) Mutual support group led by paraprofessional (2) CBT group led by paraprofessional</td>
<td>(1) CBT group led by professional (2) Mutual support group led by paraprofessional</td>
<td>10 weekly 90-minute sessions</td>
<td>None reported</td>
<td>5</td>
</tr>
<tr>
<td>Cheung, 2000</td>
<td>Community mutual aid organization/ group</td>
<td>Hong Kong adults who were still symptomatic after a 12-week group CBT program</td>
<td>CES-D: Chinese version</td>
<td>Mutual-aid groups of about 8 members each</td>
<td>Monthly meetings over 12 months (mean # attended not reported)</td>
<td>None reported</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Dennis, 2009</td>
<td>Ontario public health departments (hospitals)/ 1:1 telephone with</td>
<td>Immediately postpartum Ontario women scoring &gt;9 on EPDS</td>
<td>EPDS [SCID]</td>
<td>Usual postpartum care plus peer telephone support</td>
<td>Usual postpartum care only</td>
<td>Minimum of 4 contacts; M = 8.8 contacts over 12 weeks</td>
<td>3 months (note that 29% had continued contact</td>
<td>10</td>
</tr>
<tr>
<td>Study</td>
<td>Design/Patient Characteristics</td>
<td>Contact-Risk Score</td>
<td>Protocol/Format</td>
<td>Intervention Details</td>
<td>Timeframe for Follow-up</td>
<td>Contact Frequency</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>---------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Giese-Davis, 2006</td>
<td>California adult women with recent b.c. dx (M = 2.23 mo post-dx)</td>
<td>CES-D</td>
<td>Supportive and educational peer counseling</td>
<td>3-6 months of contact 1-4x/week (M = 4 months)</td>
<td>None reported</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gifford, 1998</td>
<td>San Francisco area adults receiving care for symptomatic HIV/AIDS</td>
<td>CES-D</td>
<td>Group (10-15 participants) educational intervention taught by teams of pair leaders (1 with HIV/AIDS)</td>
<td>Usual medical care 7 weekly sessions / 16 total hours</td>
<td>None reported</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gotay, 2007</td>
<td>Adults experiencing a first recurrence of b.c. (dx'ed in last 56 days)</td>
<td>CES-D</td>
<td>Peer-delivered telephone support</td>
<td>Usual medical care 3 months/4-8 sessions</td>
<td>None reported</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heller, 1991</td>
<td>Indiana area, low-support, community living, elderly women</td>
<td>CES-D</td>
<td>(1) Peer contact (dyad initiator)</td>
<td>(1) Supportive staff contact</td>
<td>20-30 weeks/1-2 contacts per week</td>
<td>None reported</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Hill, 2006</td>
<td>Western US adult rural-living women with chronic illness</td>
<td>CES-D</td>
<td>(2) Peer contact (dyad receiver)</td>
<td>(2) Assessment only</td>
<td>22 weeks</td>
<td>None reported</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Letourneau, 2011</td>
<td>Canadian adult postpartum women caring for</td>
<td>EPDS</td>
<td>Peer support and maternal infant interaction program</td>
<td>Waiting list (received 2 weeks of peer support)</td>
<td>12 weeks/ M = 9 visits and/or phone</td>
<td>None reported</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Setting/Cohort</td>
<td>Intervention Characteristics</td>
<td>Effectiveness</td>
<td>Follow-up Duration</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lieberman, 2005</td>
<td>Internet/group led by peers</td>
<td>Participation in active peer-support breast cancer “bulletin board” forum</td>
<td>6 months/variable (self-directed) participation frequency</td>
<td>None reported</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lieberman, 2006</td>
<td>Internet/group led by peers</td>
<td>Participation in active peer-support breast cancer “bulletin board” forum</td>
<td>6 months/variable (self-directed) participation frequency</td>
<td>None reported</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorig, 2009</td>
<td>Community settings/group led by peers</td>
<td>Educational peer-led diabetes self-management program</td>
<td>Usual medical care (waitlist for program)</td>
<td>6 weekly 2.5-hour sessions</td>
<td>6 months</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ludman, 2007</td>
<td>Outpatient medical clinic/group led by peers</td>
<td>Telephone care management + peer-led group</td>
<td>(1) Telephone care management</td>
<td>6 weekly workshops followed by ongoing bi-monthly groups</td>
<td>9 months</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marmar, 1988</td>
<td>Not specified/group led by peer women who had lost husbands several years</td>
<td>Supportive mutual-help group treatment</td>
<td>Professionally-administered brief dynamic psychotherapy</td>
<td>12 weekly 1.5-hour sessions</td>
<td>12 months</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Setting</td>
<td>Description</td>
<td>Intervention</td>
<td>Comparison</td>
<td>Participation frequency</td>
<td>Notes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>McKay, 2002</td>
<td>Internet</td>
<td>Adults diagnosed with type-2 diabetes</td>
<td>CES-D</td>
<td>(1) Educational material + online peer support forum + professional self-management coach</td>
<td>10 months/ variable (self-directed) participation frequency</td>
<td>None reported</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Nagel, 1988</td>
<td>Nursing home/ 1:1 in person</td>
<td>Elderly nursing home residents with mild to moderate depression symptoms</td>
<td>Zung SDS</td>
<td>(1) Empathy-focused peer counseling + (2) Education-focused peer counseling</td>
<td>No treatment</td>
<td>5 weeks/ two 1-hour meetings per week</td>
<td>None reported</td>
<td>6</td>
</tr>
<tr>
<td>Onrust, 2010</td>
<td>Home-based; 1:1 in person with</td>
<td>Netherlands adults age ≥55 whose spouses died during the past year</td>
<td>CES-D</td>
<td>Peer support home visiting service + Brief informational brochure on depression symptoms</td>
<td>M = 8.3 home visits</td>
<td>6-9 months</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Preyde, 2003</td>
<td>Hospital NICU/1:1 in person</td>
<td>Adult women immediately postpartum with very preterm infants less than 10 days old</td>
<td>BDI short form</td>
<td>“Parent buddy” peer support program + Standard medical care</td>
<td>16 weeks/ M = 9 contacts</td>
<td>None reported</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Robinson-Whelan, 2007</td>
<td>Centers for independent living/ group led by peer women with disabilities</td>
<td>Adult rural-living women with a physical disability and mild to moderate depression</td>
<td>BDI-II [CES-D]</td>
<td>Peer-led depression self-management program + Usual medical care</td>
<td>8 weekly 2.5-hour sessions</td>
<td>3 months</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Setting</td>
<td>Participants</td>
<td>Symptoms Measured</td>
<td>Support Intervention</td>
<td>Treatment Duration</td>
<td>Contact Times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>--------------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roman, 1995</td>
<td>Regional perinatal center/ 1:1 in person and phone with peers who had a preterm infant ≥1 year before</td>
<td>Adult women immediately postpartum, no prior NICU experience, with preterm infants less than 6 days old</td>
<td>POMS: depression subscale</td>
<td>Veteran parent peer support program</td>
<td>No treatment</td>
<td>12 months/ variable contact (M = 34 hrs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silver, 1997</td>
<td>Urban medical center/ 1:1 in person and phone with peer women who had raised children with health problems</td>
<td>Adult mothers of 5-8 y.o. children with ongoing health conditions</td>
<td>PSI: depression subscale</td>
<td>Supportive peer contact program</td>
<td>No treatment</td>
<td>12 months/ 6 face-to-face meetings and at least biweekly phone contact</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simoni, 2007</td>
<td>HIV primary care outpatient clinic/ group in person and 1:1 phone calls with peer supporters (HIV+ and on HAART)</td>
<td>New York area adult HIV-positive patients on HAART regimen</td>
<td>CES-D</td>
<td>Peer support intervention targeting medication adherence and depression symptomology</td>
<td>Standard medical care</td>
<td>3 months/ 6 bimonthly group meetings, and phone calls 1-3x/week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travis, 2010</td>
<td>Outpatient mental health clinics/ 1:1 phone with another participant</td>
<td>Michigan adults with depression symptoms (BDI-II &gt;12) and hx of at least 2 antidep med trials</td>
<td>BDI-II</td>
<td>Telephone-based mutual peer support program</td>
<td>12 weeks/ at least 1 weekly call (M = 10.3 completed calls)</td>
<td>None reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tudiver, 1992</td>
<td>Community facility/ group led by peer (widowed) facilitator</td>
<td>Adult men recently widowed (3-12 months prior)</td>
<td>BDI short form [GHQ]</td>
<td>Peer-facilitated mutual help groups</td>
<td>Waiting list (offered tx after 8 months)</td>
<td>9 weekly 1.5-hour sessions</td>
<td>8 months (not included in analyses b/c of</td>
<td></td>
</tr>
<tr>
<td>Uccelli, 2004</td>
<td>Not specified/group led by peer (MS patient nominated by neurologist)</td>
<td>Italian adults diagnosed with MS</td>
<td>BDI: MS adaptation</td>
<td>Peer support group intervention</td>
<td>8 weekly sessions</td>
<td>None reported</td>
<td>4</td>
<td>insufficient data reported</td>
</tr>
</tbody>
</table>
### Table 2: Characteristics of depression symptom measurement instruments

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Author</th>
<th>Description</th>
<th>(N)</th>
<th>Score threshold for elevation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beck Depression Inventory (BDI)</td>
<td>Beck et al., 1961</td>
<td>21-item self-report; scores range 0-63</td>
<td>2</td>
<td>(\geq 10)</td>
</tr>
<tr>
<td>Beck Depression Inventory-II (BDI-II)</td>
<td>Beck et al., 1996</td>
<td>21-item self-report; scores range 0-63</td>
<td>2</td>
<td>(\geq 14)</td>
</tr>
<tr>
<td>Hamilton Rating Scale for Depression (HRSD)</td>
<td>Hamilton, 1960</td>
<td>Interviewer-administered; scores range 0-52</td>
<td>1</td>
<td>(\geq 8)</td>
</tr>
<tr>
<td>Center for Epidemiological Studies-Depression (CES-D)</td>
<td>Radloff, 1977</td>
<td>20-item self-report; scores range 0-60</td>
<td>11</td>
<td>(\geq 15)</td>
</tr>
<tr>
<td>Edinburgh Postnatal Depression Scale (EPDS)</td>
<td>Cox et al., 1987</td>
<td>10-item self-report; scores range 0-30</td>
<td>2</td>
<td>(\geq 9)</td>
</tr>
<tr>
<td>Patient Health Questionnaire-9 (PHQ-9)</td>
<td>Kroenke et al., 2001</td>
<td>9-item self-report; scores range 0-27</td>
<td>1</td>
<td>(\geq 5)</td>
</tr>
<tr>
<td>Symptom Check List-90 (SCL-90)(^2)</td>
<td>Derogatis et al., 1977</td>
<td>20-item self-report; scores range 0-60</td>
<td>1</td>
<td>(\geq 25)</td>
</tr>
<tr>
<td>Zung Self-Rating Depression Scale</td>
<td>Zung, 1965</td>
<td>20-item self-report; scores range 20-80</td>
<td>1</td>
<td>(\geq 40)</td>
</tr>
<tr>
<td>Beck Depression Inventory-Short Form</td>
<td>Beck et al., 1993</td>
<td>13-item self-report; scores range 0-39</td>
<td>3</td>
<td>(\geq 5)</td>
</tr>
<tr>
<td>Profile of Mood States (POMS)(^2)</td>
<td>McNair et al., 1981</td>
<td>15-item self-report; scores range 0-60</td>
<td>1</td>
<td>(\geq 7)</td>
</tr>
<tr>
<td>Psychiatric Symptom Index (PSI)(^2)</td>
<td>Ilfeld, 1976</td>
<td>10-item self-report; scores range 0-100 (standardized)</td>
<td>1</td>
<td>(\geq 20)</td>
</tr>
</tbody>
</table>

\(^1\)Number of studies included in meta-analysis for which this was considered the primary depression measure  
\(^2\)Depression subscale score  
\(^3\)Represents summed item scores; sometimes reported as mean score per item (as in Ludman et al., 2007)
Table 3: Results of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Defining condition of sample</th>
<th>% female</th>
<th>Mean age</th>
<th>Group</th>
<th>N Enrolled</th>
<th>N Completed</th>
<th>N Follow-up</th>
<th>Depression mean (SD) Pre-treatment</th>
<th>Post-treatment</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson, 2005</td>
<td>Depression</td>
<td>74</td>
<td>36.1</td>
<td>PAI</td>
<td>60</td>
<td>49</td>
<td>35</td>
<td>20.9 (8.5) 20.5 (6.7)</td>
<td>19.5 (8.1)</td>
<td>13.1 (7.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG</td>
<td>57</td>
<td>36</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bright, 1999</td>
<td>Depression</td>
<td>71</td>
<td>45.8</td>
<td>PAI(1)</td>
<td>22</td>
<td>14</td>
<td>18</td>
<td>18.01 (4.7) 17.86 (5.18)</td>
<td>17.11 (3.59)</td>
<td>17.67 (4.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PAI(2)</td>
<td>21</td>
<td>13</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG(1)</td>
<td>27</td>
<td>18</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG(2)</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheung, 2000</td>
<td>Depression</td>
<td>83</td>
<td>38</td>
<td>PAI</td>
<td>83</td>
<td>65</td>
<td></td>
<td>21.54 (12.28) 23.57 (13.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dennis, 2009</td>
<td>Depression (postnatal)</td>
<td>100</td>
<td>78% ages 20-34</td>
<td>PAI</td>
<td>349</td>
<td>297</td>
<td>289</td>
<td>12.5 (2.8) 12.62 (2.76)</td>
<td>7.93 (4.68)</td>
<td>7.00 (4.66)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG</td>
<td>352</td>
<td>316</td>
<td>311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giese-Davis, 2006</td>
<td>Chronic illness (breast cancer)</td>
<td>100</td>
<td>51.31</td>
<td>PAI</td>
<td>29</td>
<td>19</td>
<td></td>
<td>21.75 (12.51) Not given</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gifford, 1998</td>
<td>Chronic illness (HIV/AIDS)</td>
<td>0</td>
<td>45.26</td>
<td>PAI</td>
<td>34</td>
<td>25</td>
<td></td>
<td>17.3 (8.8) 19.6 (11.3)</td>
<td>18.7 (10.5)</td>
<td>20.0 (11.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG</td>
<td>37</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gotay, 2007</td>
<td>Chronic illness (breast cancer)</td>
<td>100</td>
<td>54</td>
<td>PAI</td>
<td>152</td>
<td>128</td>
<td>124</td>
<td>17.72 (10.71) 15.62 (10.12)</td>
<td>15.39 (9.69)</td>
<td>14.14 (9.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG</td>
<td>152</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heller, 1991</td>
<td>Depression</td>
<td>100</td>
<td>74</td>
<td>PAI(1)</td>
<td>49</td>
<td>49</td>
<td></td>
<td>31.1 (8.4) 31.6 (6.8)</td>
<td>31.3 (10.1)</td>
<td>30.5 (7.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PAI(2)</td>
<td>49</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG(1)</td>
<td>44</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CG(2)</td>
<td>53</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hill, 2006</td>
<td>Chronic illness</td>
<td>100</td>
<td>92% older</td>
<td>PAI</td>
<td>61</td>
<td>43</td>
<td></td>
<td>18.52 (11.64) 15.50 (11.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Diagnosis</td>
<td>Sample Size</td>
<td>Gender</td>
<td>Mean Age</td>
<td>PAI (SD)</td>
<td>CG (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>-------------</td>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letourneau, 2011</td>
<td>Depression (postnatal)</td>
<td>100</td>
<td>Majority 26-35</td>
<td>42</td>
<td>18.20 (4.76)</td>
<td>16.1 (4.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lieberman, 2005</td>
<td>Chronic illness (breast cancer)</td>
<td>100</td>
<td>46.2</td>
<td>91</td>
<td>20.0 (.85)</td>
<td>15.3 (1.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lieberman, 2006</td>
<td>Chronic illness (breast cancer)</td>
<td>100</td>
<td>45.5</td>
<td>52</td>
<td>19.90 (10.01)</td>
<td>14.60 (10.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorig, 2009</td>
<td>Chronic illness (type 2 diabetes)</td>
<td>66</td>
<td>66.7</td>
<td>133</td>
<td>6.35 (5.85)</td>
<td>4.84 (4.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ludman, 2007</td>
<td>Depression (spousal bereavement)</td>
<td>71</td>
<td>50.2</td>
<td>26</td>
<td>1.63 (.68)</td>
<td>1.22 (.54)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marmar, 1988</td>
<td>Stressor (spousal bereavement)</td>
<td>100</td>
<td>58</td>
<td>30</td>
<td>8.80 (4.66)</td>
<td>8.77 (5.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McKay, 2002</td>
<td>Chronic illness (type 2 diabetes)</td>
<td>47</td>
<td>59</td>
<td>40</td>
<td>15.83 (12.79)</td>
<td>13.86 (12.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagel, 1988</td>
<td>Depression</td>
<td>Not available</td>
<td>74</td>
<td>20</td>
<td>45.60 (4.77)</td>
<td>39.80 (6.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onrust, 2011</td>
<td>Stressor</td>
<td>64</td>
<td>68.9</td>
<td>20</td>
<td>49.60 (6.97)</td>
<td>48.50 (6.20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Condition</td>
<td>Sample</td>
<td>PAI</td>
<td>CG</td>
<td>PAI</td>
<td>CG</td>
<td>PAI</td>
<td>CG</td>
<td>PAI</td>
<td>CG</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>2010</td>
<td>(spousal bereavement)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preyde, 2003</td>
<td>Stressor (preterm infant)</td>
<td>100</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.53</td>
<td>5.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.20</td>
<td>4.88</td>
</tr>
<tr>
<td>Robinson-Whelan, 2007</td>
<td>Depression + Chronic Illness</td>
<td>100</td>
<td>51.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.2</td>
<td>29.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.93</td>
<td>27.50</td>
</tr>
<tr>
<td>Roman, 1995</td>
<td>Stressor (preterm infant)</td>
<td>100</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.7</td>
<td>14.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.00</td>
<td>7.78</td>
</tr>
<tr>
<td>Silver, 1997</td>
<td>Stressor (child’s ongoing health problem)</td>
<td>100</td>
<td>34.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.8</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.1</td>
<td>19.6</td>
</tr>
<tr>
<td>Simoni, 2007</td>
<td>Chronic illness (HIV)</td>
<td>44</td>
<td>42.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.9</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17.6</td>
<td>17.9</td>
</tr>
<tr>
<td>Travis, 2010</td>
<td>Depression</td>
<td>37</td>
<td>52.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.0</td>
<td>22.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.4</td>
<td>17.9</td>
</tr>
<tr>
<td>Tudiver, 1992</td>
<td>Stressor (spousal bereavement)</td>
<td>0</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7.3</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.2</td>
<td>5.4</td>
</tr>
<tr>
<td>Uccelli, 2004</td>
<td>Chronic illness (multiple sclerosis)</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.19</td>
<td>26.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.4</td>
<td>6.80</td>
</tr>
</tbody>
</table>
Table 4: Mean SMD estimates (PAI compared with active therapy conditions) adjusted for four models of selection (publication) bias

<table>
<thead>
<tr>
<th></th>
<th>FIXED EFFECTS</th>
<th>RANDOM EFFECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
<td>Parameter</td>
</tr>
<tr>
<td></td>
<td>estimate</td>
<td>estimate</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>( V ) (estimate of)</td>
</tr>
<tr>
<td>Unadjusted ES (SMD)</td>
<td>0.03141657</td>
<td>0.08347356</td>
</tr>
<tr>
<td></td>
<td>0.07084147</td>
<td>(SE = 0.11564934)</td>
</tr>
<tr>
<td>Moderate one-tailed selection</td>
<td>-0.01738422</td>
<td>-0.01431520</td>
</tr>
<tr>
<td>Severe one-tailed selection</td>
<td>-0.09302688</td>
<td>-0.83928611</td>
</tr>
<tr>
<td>Moderate two-tailed selection</td>
<td>0.02677201</td>
<td>0.07177643</td>
</tr>
<tr>
<td>Severe two-tailed selection</td>
<td>0.02142850</td>
<td>0.05942731</td>
</tr>
</tbody>
</table>
Table 5: Mean SMD estimates (PAI compared with TAU/WL conditions) adjusted for four models of selection (publication) bias

<table>
<thead>
<tr>
<th>Parameter estimate</th>
<th>Standard error</th>
<th>Parameter estimate</th>
<th>( V ) (estimate of) population effect-size variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted ES (SMD)</td>
<td>0.09130349</td>
<td>0.0403020</td>
<td>0.1026520 (SE = 0.1060054)</td>
</tr>
<tr>
<td>Moderate one-tailed selection</td>
<td>0.06115963</td>
<td>-1.018147072</td>
<td>1.5018998</td>
</tr>
<tr>
<td>Severe one-tailed selection</td>
<td>0.02258148</td>
<td>0.095101156</td>
<td>0.1438070</td>
</tr>
<tr>
<td>Moderate two-tailed selection</td>
<td>0.07947753</td>
<td>0.086774433</td>
<td>0.1436267</td>
</tr>
<tr>
<td>Severe two-tailed selection</td>
<td>0.06487532</td>
<td>0.086774433</td>
<td>0.1436267</td>
</tr>
</tbody>
</table>
APPENDIX A: CODING FORM

First Author:  
Year of Publication:  
Title:  

1. **Control group?**  YES  NO  
   If yes:  
   a. **Randomized?**  YES  NO  NOT SPECIFIED  
   b. **Assessments blinded?**  YES  NO  NOT SPECIFIED  
   c. **Number of control groups:**  
   d. **Type of CG1:**  OTHER THERAPY  TAU  MED  WAITLIST  
      i. If other therapy, specify type:  
   e. **Type of CG2:**  OTHER THERAPY  TAU  MED  WAITLIST  
      i. If other therapy, specify type:  
   f. **Researcher allegiance:**  PAI was presented as:  
      i. _____ An active treatment (i.e., a treatment of interest)  
      ii. _____ A control condition  

2. **Sample type:**  DEPRESSION  STRESSOR  CHRONIC ILLNESS  OTHER  

3. **Intended duration of PAI treatment:**  ______ weeks and/or ______ contacts  

4. **Observed average duration of PAI treatment:**  ______ weeks and/or ____ contacts  

5. **Frequency of PAI treatment:**  WEEKLY  MORE THAN WEEKLY  LESS THAN WEEKLY  VARIABLE/UNCLEAR  

6. **Format of PAI treatment:**  GROUP  ONE-ON-ONE  

7. **Mode of delivery of PAI treatment:**  FACE-TO-FACE  PHONE  INTERNET  

8. **Support is:**  MUTUAL  DELIVERED FROM ONE PERSON TO ANOTHER
a. If one person to another, any outcomes for supporters reported?
   YES  NO

9. Peer supporters were selected by:  THEIR OWN RESPONSE TO AN AD
   RECRUITED AS FORMER PATIENTS
   OTHER  NOT SPECIFIED

10. Pre-intervention training of peer supporters:  SOME  NONE  NOT
     SPECIFIED

11. Did peer supporters have ongoing contact with a professional supervisor over
    the course of the intervention?   YES  NO  NOT SPECIFIED

12. Was the PAI treatment co-administered by a health or mental health
    professional, or entirely peer-administered?  PROFESSIONAL CO-
    ADMINISTERED

   ENTIRELY PEER-ADMINISTERED

13. Content of PAI treatment is described as primarily:  SUPPORTIVE
     EDUCATIONAL
     NOT SPECIFIED

14. Total N screened for study: _____

15. Total N enrolled in study (providing pretest data): _____
    a. N enrolled in PAI: _____
    b. N enrolled in CG 1: _____; CG 2: _____

16. N completed study (providing posttest data): _____ (specify whether N or %)
    a. N completed PAI: _____
    b. N completed CG1: _____; N completed CG2: _____

17. Gender breakdown of total sample: _____% female, _____% male
    a. Gender breakdown of PAI group: _____% female, _____% male
    b. Gender breakdown of CG1: _____% female, _____% male
    c. Gender breakdown of CG2: _____% female, _____% male

18. Mean age of total sample: _____ years
    a. Mean age of PAI group: _____ years
    b. Mean age of CG1: _____ years
c. **Mean age of CG2:** _____ years

19. **Instrument(s) used to measure depression (check all that were used)***:
   a. _____ HRSD (a.k.a., HAM-D)
   b. _____ BDI or BDI-II
   c. _____ CES-D
   d. _____ Other: ______________________________________

*If more than one measure was used, write “primary” or “secondary” next to each one checked. HRSD is always primary if it was used, followed in order of priority by BDI, CES-D, EPDS, HADS, and finally any other measure.

20. **PRETEST depression score of PAI group**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

21. **POSTTEST depression score of PAI group**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

22. **PRETEST depression score of CG1**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

23. **POSTTEST depression score of CG1**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

24. **PRETEST depression score of CG2**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

25. **POSTTEST depression score of CG2**:
   a. Primary measure name: ______ Mean: ______ SD:______
   b. Secondary measure name: ______ Mean: ______ SD:______

26. If reported, % **considered “improved” at posttest**:
   PAI: ______    CG1: ______    CG2: ______
27. Are any follow-up data reported (including in a separate manuscript)?
   YES   NO
   If yes:
   a. Number of months post-treatment: ______
   b. Was help-seeking behavior during the follow-up period reported?
      YES   NO
      i. If Yes, % who sought help: PAI: ____  CG1: ____  CG2: ____
   c. Was % remaining improved at follow-up reported?  YES   NO
      i. If Yes, % remaining improved: PAI: ____  CG1: ____  CG2: ____
   d. FOLLOW-UP depression score of PAI group:
      i. Primary measure mean: ______  SD: ______
      ii. Secondary measure mean: ______  SD: ______
   e. FOLLOW-UP depression score of CG1:
      i. Primary measure mean: ______  SD: ______
      ii. Secondary measure mean: ______  SD: ______
   f. FOLLOW-UP depression score of CG2:
      i. Primary measure mean: ______  SD: ______
      ii. Secondary measure mean: ______  SD: ______
APPENDIX B: METHODOLOGICAL QUALITY SCORE (MQS) CODING SYSTEM

(adapted from Miller et al., 1995)

<table>
<thead>
<tr>
<th>Methodological feature</th>
<th>Points available</th>
<th>Points awarded</th>
</tr>
</thead>
</table>
| Group allocation       | 2 = randomized comparison group  
                        | 1 = nonrandomized comparison group  
                        | 0 = no comparison group            | |
| Quality control        | 1 = treatment standardized by manual, specific training, etc.  
                        | 0 = no standardization specified   | |
| Immediate posttreatment assessment rate | 2 = 85-100% of post assessments completed  
                                        | 1 = 70-84.9% of post assessments completed  
                                        | 0 = <70% of post assessments completed, or not specified | |
| Additional posttreatment follow-ups | 1 = at least one additional post-treatment follow-up completed  
                                          | 0 = no additional post-treatment follow-up, or not specified | |
| Dropouts               | 1 = treatment dropouts included in at least some outcome data (e.g., intent-to-treat analysis or compared on dependent variable)  
                        | 0 = treatment dropouts not discussed or accounted for | |
| Blinding               | 1 = posttreatment assessment done by blinded interviewer  
                        | 0 = posttreatment assessment nonblinded, or unspecified | |
| Analyses               | 1 = acceptable statistical analyses  
                        | 0 = no statistical analyses, inappropriate analyses, or unspecified | |
| Multisite              | 1 = parallel replications at two or more sites with separate research teams  
                        | 0 = single site or comparison of sites offering different treatments | |

**SUM QUALITY SCORE:**
REFERENCES

Asterisked entries denote studies that were included in the meta-analysis.


