Change in reflective functioning in interpersonal psychotherapy and cognitive behavioral therapy for major depressive disorder

Annika Ekeblad, Rolf Holmqvist, Gerhard Andersson & Fredrik Falkenström

Published online: 11 Aug 2022.

Submit your article to this journal

Article views: 588

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

RESEARCH ARTICLE

Change in reflective functioning in interpersonal psychotherapy and cognitive behavioral therapy for major depressive disorder

ANNIKA EKEBLAD[1,2], ROLF HOLMQVIST [1], GERHARD ANDERSSON[1,3], & FREDRIK FALKENSTRÖM [1]

[1]*Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden;* [2]*Psychiatric Clinic, Västernorrland County Council, Sundsvall, Sweden &* [3]*Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden*

Abstract

Background Patients with Major Depressive Disorder (MDD) have been found to have restricted capacity for mentalization, and it is possible that this constitutes a vulnerability factor for developing depression. Due to its focus on linking depressive symptomatology to emotions and interpersonal relations, it was hypothesized that Interpersonal Psychotherapy (IPT) would improve mentalization more than Cognitive Behavioral Therapy (CBT).
Methods In a randomized controlled trial of 90 patients undergoing IPT and CBT for MDD, Reflective Functioning (RF) was rated from Adult Attachment and from Depression-Specific Reflective Functioning (DSRF) Interviews before and after therapy. Treatment outcome was assessed using the Beck Depression Inventory-II.
Results The interaction between time and treatment approach was statistically significant, with RF improving significantly more in IPT than in CBT. Change in RF was not correlated with change in depression. The difference in DSRF ratings before and after therapy was not statistically significant for any of the treatments.
Conclusions IPT may improve mentalization more than CBT. However, although RF increased significantly in IPT, the mean level was still low after therapy. A limitation of the study is the large amount of post-treatment missing data. More research is needed to understand the potential role of mentalization in symptom reduction.

Keywords: mentalization; reflective functioning; major depressive disorder; therapy process; interpersonal psychotherapy; cognitive-behavior therapy

Clinical or methodological significance of this article: Limited mentalization may imply a vulnerability to psychopathology, including depression. Interpersonal Psychotherapy was shown to improve mentalization more than cognitive-behavioral therapy in patients with major depressive disorder. If replicated, these findings may point to a specific mechanism of change for Interpersonal Psychotherapy.

Mentalization is defined as the capacity to understand human behavior in terms of underlying mental states, i.e. thoughts, feelings, wishes, needs, and intentions (Fonagy et al., 2002). Since a capacity for mentalizing is fundamental for interpersonal relationship functioning, it has been hypothesized that deficient mentalization might be a common underlying factor of many, if not most, forms of psychopathology (Katznelson, 2014). Indeed, a number of studies have found reduced mentalizing capacity in persons with various psychopathologies; e.g., borderline personality disorder

Correspondence concerning this article should be addressed to Fredrik Falkenström, Department of Behavioral Sciences and Learning, Linköping University, Linköping, SE-581 83, Sweden. Emails: fredrik.falkenstrom@liu.se; fredrik.falkenstrom@lnu.se

(Bateman & Fonagy, 2004), eating disorders (Skårderud, 2007a, 2007b, 2007c) and depression (Ekeblad, Falkenström, & Holmqvist, 2016; Fischer-Kern et al., 2008, 2013). Researchers have also found low capacity for mentalization in persons with other types of social and psychiatric problems like criminal offenders (Levinson & Fonagy, 2004; Möller et al., 2014).

Some studies have found lower levels of mentalization, as measured by the Reflective Functioning scale (RF; Fonagy et al., 1998) rated on the Adult Attachment Interview (George et al., 1985), in depressed patients than in non-clinical samples (Ekeblad, Falkenström, & Holmqvist, 2016; Fischer-Kern et al., 2008, 2013). The direction of potential causation for this relationship is, however, not established. It is possible that depression impairs the capacity for mentalization, since it is well-known that depression results in impaired cognitive functioning (e.g., McIntyre et al., 2013). A study by Ekeblad, Falkenström, and Holmqvist (2016) addressed this by showing that RF in depressed patients was uncorrelated with neuropsychological tests of executive functions and verbal fluency, indicating that mentalization deficits were unrelated to more general cognitive deficits. Another possibility is that an impaired capacity to mentalize is a risk factor for developing depression. If this would be the case, low RF might, for instance, cause problems in interpersonal relationships, which in turn might generate depressive symptoms (or other psychopathological symptoms). A third alternative is a bi-directional influence, implying for instance that depression lowers the capacity for mentalization and that lowered mentalization contributes to interpersonal problems that may in turn enhance depression (Luyten et al., 2012).

There is one published study that showed close to normal RF in a group of depressed patients (Taubner et al., 2011). The difference between the studies that showed low RF in depression and this study seems to be mainly in the level of functioning and comorbidity, with the studies showing low RF in depressed patients having more severely comorbid and low functioning samples (Ekeblad, Falkenström, & Holmqvist, 2016; Fischer-Kern et al., 2008, 2013). This may imply that low RF in depression is an indication of low functioning in general rather than depression per se, potentially meaning that improving RF is an important secondary target of depression treatment that may reduce risk of relapse.

There are no studies of change in mentalization after treatment of depression. Levy et al. (2006) found that borderline patients who had received one year of Transference-Focused Psychotherapy (Clarkin et al., 2006) increased their RF level from 2.86 to 4.11. No change was found for borderline patients in Supportive therapy or Dialectical Behavior Therapy (Linehan, 1993). This indicates that some types of therapy may promote change in RF more than other therapies.

Interpersonal Psychotherapy (IPT; Klerman et al., 1984) is an evidence-based psychotherapy for depression, focusing on improving interpersonal relationships primarily by relating depressive symptoms to interpersonal issues. The most central process in IPT concerns exploring and elaborating on links between depressive symptoms and one of four interpersonal focus areas; (1) problematic grief, (2) role transition, (3) role conflict, or (4) interpersonal sensitivity. As such, it has been suggested that improved reflective functioning may be a central, although implicit, goal for IPT (Markowitz et al., 2009; Rudden et al., 2009) whereas in Cognitive Behavioral Therapy (CBT), change in RF may be less likely due to the focus on change in behavior and/or making cognitions more realistic.

In a previous paper from our group, we found that RF predicted both the quality of the working alliance and change in depressive symptoms during treatment (Ekeblad, Falkenström, & Holmqvist, 2016). Patients with higher baseline RF experienced the alliance as stronger and improved more than patients with lower RF. However, the questions whether RF was improved in these therapies, and, if so, if that was equal across both therapies, were not addressed, and neither was the question whether potential RF change was correlated with change in depressive symptoms.

The main purpose of the present study was to assess whether mentalization levels would increase more after treatment with IPT than after CBT. Specifically, we hypothesized that RF and depression-specific RF would improve more in IPT compared to CBT. A second purpose was to explore whether change in mentalization was related to change in depression symptoms.

Methods

Participants

Data for this study was taken from the Cognitive and Interpersonal Psychotherapy at the Psychiatric clinic in Sundsvall, a randomized non-inferiority trial comparing depression outcome in CBT and IPT. This trial was pre-registered for comparing improvement between CBT and IPT in depression and in return to work, hypothesizing that IPT would be non-inferior to CBT in depressed psychiatry patients (Ekeblad, Falkenström, Andersson, et al., 2016). Data on mentalization were collected as part of the

design as (unregistered) secondary outcome, and although several other papers have been published from this trial (mostly process studies), none has focused on secondary treatment outcomes.

The participating patients were all diagnosed with Major Depressive Disorder (MDD) by experienced psychiatrists and clinical psychologists using the Structured Clinical Interview for DSM IV (SCID) Module A (First & Pincus, 1999). All patients had previously received treatment for depression, most commonly medication in primary care, with none or only partial response, and they were referred to the psychiatric clinic for further treatment. The inclusion criteria were age 18–65 years and MDD diagnosis. Exclusion criteria were psychosis, ongoing substance addiction, serious neuropsychiatric disorder, or active self-harm behavior. To be included in the study the patient had to accept random allocation to the therapy methods and video filming of all sessions. A total of 96 patients were included in the trial, out of 99 who were asked to participate. Patients who met the inclusion criteria were fully informed about the study and gave their written informed consent. A randomization list was generated by an independent statistician. The randomization procedure was done by a psychologist at the clinic not otherwise involved in the project. In this secondary data analysis, we used data from the completer sample, with 63 patients who had completed therapy. The study was approved by the Regional Ethical Review Board in Linköping (2010/348-31).

Therapists

All participating therapists worked on a regular basis at the psychiatric clinic. There were nine therapists providing IPT: six psychiatric nurses, one nurse assistant, one occupational therapist and one social worker. The mean number of patients treated by each IPT therapist was 4.9 (range 2–8). There were 25 therapists providing CBT: 14 psychologists, four psychiatric nurses, two social workers, one nurse assistant, one occupational therapist and three physicians. The CBT therapists treated on the average 1.9 patients (range 1–5).

Training and Supervision

All therapists had basic training in psychotherapy and specialist training in the treatment method they provided. The therapists delivered either IPT or CBT, no therapists provided both therapy forms. During the trial, all therapists had regular supervision (once to twice a month) with certified supervisors in the respective methods and were given the opportunity to attend more days of training in their respective therapy forms.

Treatments

In this trial, the format for both treatments was 14 weekly sessions. IPT was delivered according to the manual by Weissman et al. (2000). CBT was delivered according to two manuals for treatment of depression, one cognitive and one behavioral (Beck et al., 1979; Martell et al., 2010). The CBT therapists were proficient in both manuals and used them to different degrees according to their clinical judgment. Some therapists also included mindfulness-based interventions (Segal et al., 2013) in their CBT treatment according to their clinical judgment. This is how CBT is usually delivered in clinical practice in psychiatric care in Sweden.

Measures

The pre-registered primary outcomes of the study were Montgomery Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979) and return to work. The present study is based on RF, which was a non-registered secondary outcome.

Beck Depression Inventory-II (Beck & Steer, 1996) is a widely used instrument to self-assess depressive symptoms. The scale consists of 21 items, each item rated from 0 to 3. The BDI-II has showed high reliability, capacity to discriminate between depressed and non-depressed subjects, and improved concurrent, content, and structural validity (Wang & Gorenstein, 2013).

The Adult Attachment Interview (AAI), developed by George et al. (1985), is a semi-structured interview in which respondents are asked to describe their childhood attachment experiences and to evaluate the possible impact of these experiences on their adult personality. The interview was transcribed verbatim, and the transcriptions were rated according to the RF scale. In the present study, a slightly shortened version of the interview was used (questions 1–11). Since this version includes most of the so-called demand questions (Taubner et al., 2013), it is possible to use it for rating RF. The AAI was administered by a research assistant who was not involved in the present study.

Reflective Functioning (Fonagy et al., 1998) is used to rate mentalization, usually from responses to the AAI. High RF is characterized by AAI interview passages showing explicit mentalization, especially when interview questions are posed to which a response without any reflection on mental states would seem odd (so-called demand questions,

e.g., "Why do you think your parents behaved the way they did during your childhood?"). The four overarching categories of responses that are scored as Reflective Functioning are:

A. Understanding of the nature of mental states
B. Explicit efforts to tease out mental states underlying behavior
C. Recognizing developmental aspects of mental states
D. Mental states in relation to the interviewer

Scores are given to individual passages throughout the interview, but in the end these are weighed together to one final RF score between −1 and 9, where −1 is negative or antireflective RF, 5 is considered ordinary RF, and 9 is exceptional RF. The Reflective Functioning scale has shown the expected one-factor structure, good reliability, and stability over time (Taubner et al., 2013). In the present study, pre-therapy interviews for both RF and DSRF were rated by the first and last authors, with an inter-rater reliability ICC = .84 (two-way mixed model). Post-therapy DSRF was also rated by both authors, with ICC = .79. The first author rated all post-therapy RF interviews, while the last author only rated six of these. However, inter-rater reliability for these six ratings was good (ICC = .84). All ratings were done blindly to treatment condition.

Depression-Specific Reflective Functioning (Ekeblad, Falkenström, & Holmqvist, 2016). Rudden et al. (2006) developed a Panic-Specific Reflective Functioning interview to test the hypothesis that in specific Axis I disorders, mentalization in general may not be impaired but there may still be specific mentalization deficits around a certain symptom area. In this trial, we have included a Swedish version of this interview, adapted slightly to be used with depression, therefore called Depression-Specific Reflective Functioning (DSRF; Ekeblad, Falkenström, & Holmqvist, 2016). The interview was scored using the original RF manual (Fonagy et al., 1998).

Treatment Adherence

Treatment adherence was assessed by rating sessions 3 and 7 using the Collaborative Study Psychotherapy Rating Scale-6 (Evans et al., 1984) from videotaped therapy sessions. The therapists were not aware of which sessions would be rated. The IPT ratings were higher in IPT-therapies than in CBT therapies (2.15 vs. 1.35, $t(65) = 5.72$, $p < .001$), and the CBT ratings were higher in CBT therapies than in IPT therapies (2.35 vs 1.39, $t(65) = 7.43$, $p < .001$). IPT-therapists were more adherent to IPT than to CBT (2.15 vs 1.39, $t(31) = 8.40$, $p < .001$), and CBT therapists were more adherent to CBT than to IPT (2.35 vs 1.34, $t(34) = 7.21$, $p < .001$). Thus, adherence was considered adequate.

Procedure

Interviews were conducted at baseline by an external interviewer, and after completion of therapy by the same external rater as at baseline. Ratings of BDI-II were made by the patient before every therapy session. Many of the dropouts did not return for post-therapy assessments, resulting in large attrition on post-therapy interviews in this subsample. However, when analyzing mechanisms of change, it can be argued that focusing on completers is more appropriate than intention-to-treat analysis, since only completers have been exposed to a full therapy dose and this may be required for the mechanisms to work properly. This is especially so in the case of RF, since this is likely to change only slowly. Thus, only the patients who completed treatment are considered as the study population for the present study. Figure 1 shows the flow of dropout and data attrition in the study. As can be seen in the figure, 63 patients completed treatment. Of these, we had complete data at both pre- and post-treatment for 41 patients for RF, and 39 patients for DSRF. The initial RF level did not differ between treatment completers and dropouts ($t(83) = 0.29$, ns).

Statistical Analysis

Power analysis for the CIPPS study was based on the non-inferiority hypothesis for the depression severity outcome (Ekeblad, Falkenström, Andersson, et al., 2016). Post-hoc power analysis for the RF outcome, based on the completer sample size (with complete pre- and post data; $n = 41$, and a pre-post correlation of $r = .71$), showed that the study would have 80% power to find an effect of at least $d = .45$ at alpha = .05.

The primary hypotheses were analyzed with repeated measures ANOVA (focusing on the time × treatment interaction). However, since there was a lot of attrition, especially on post-treatment interviews, any statistically significant analyses were re-estimated using multiple imputation. In multiple imputation, all available data is used, and it is also possible to generate observations for cases with no data on the dependent variable as long as there are other variables in the dataset with information on that subject that are correlated with the dependent variable. With multiple imputation, the assumption
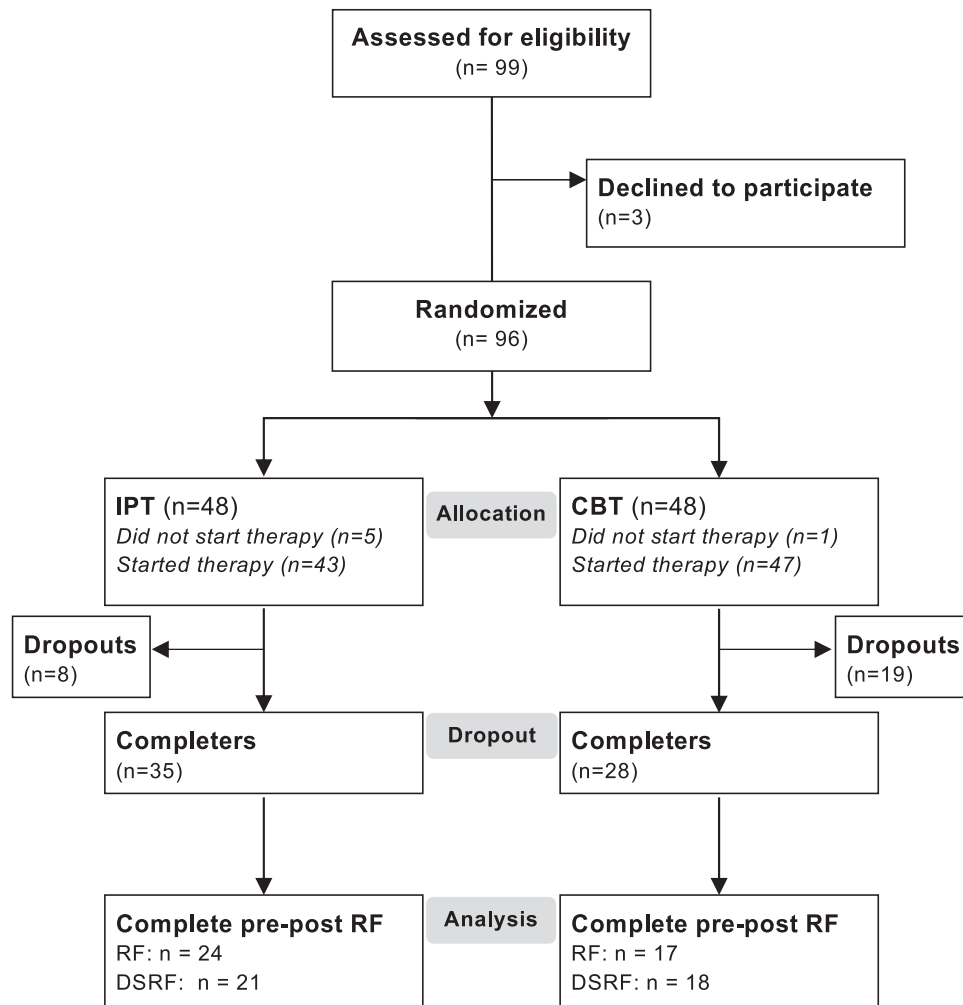
Figure 1. Flowchart of patient inclusion, dropout, and data attrition on RF/DSRF.

for data missingness is missing at random—which is less restrictive than missing completely at random that is assumed when cases with missing data are excluded (Enders, 2010).

Imputation was conducted using a linear regression model, separately for the IPT and CBT groups to preserve interaction effects (Enders, 2010). All variables having a significant correlation with RF in either group were used as auxiliary variables in the imputation procedure. These variables were time-point (pre/post therapy), patient gender, pre- and post-treatment BDI-II score, pre/post treatment sick leave, and pre-treatment personality disorder. Altogether, 27 missing observations out of a total of 126 possible observations (63 completers × 2 time-points) were imputed (i.e., 21% of observations were imputed). One hundred samples were generated and analyzed. All analyses were conducted using Stata 16.0 (StataCorp, 2019).

Results

Descriptive Statistics

Table I shows RF means, standard deviations, ranges, skewness, and kurtosis for patients completing IPT and CBT. The Doornik and Hansen (2008) test of multivariate normality was nonsignificant both for RF (χ^2 (4) = 9.14, p = .06) and for DSRF (χ^2 (4) = 8.21, p = .08), indicating that multivariate normality held. Little's (1988) test for missing completely at random was nonsignificant (χ^2 (5) = 4.18, p = .52), indicating no evidence for nonrandom missingness among treatment completers.

The correlation between RF and DSRF for the initial ratings was r = .55 (p < .001, n = 53). The correlation between RF and DSRF at the end of therapy was r = .42 (p = .005, n = 43). The correlation between the change scores for RF and DSRF was r = .30 (p = .07, n = 37).

Table I Reflective functioning means, standard deviations, and ranges for patients completing Interpersonal Psychotherapy (IPT; $N = 35$) and Cognitive Behavior Therapy (CBT; $N = 28$).

IPT	n	Mean	Std. Dev.	Min	Max	Skewness	Kurtosis
Pre-therapy RF	30	2.63	1.31	1	6	0.93	2.99
Post-therapy RF	26	3.35	1.35	1.5	7	0.87	3.45
Pre-therapy DSRF	28	2.48	1.04	1	5	0.85	3.44
Post-therapy DSRF	25	2.68	0.92	1	5	0.31	3.11
CBT	n	Mean	Std. Dev.	Min	Max		
Pre-therapy RF	26	2.83	0.94	1	4.5	0.11	1.99
Post-therapy RF	18	2.83	1.06	1.5	6	1.41	5.53
Pre-therapy DSRF	25	2.40	0.95	1	4	−0.31	2.04
Post-therapy DSRF	20	2.35	0.88	1.5	5	1.44	5.15

Change in Reflective Functioning

The mean RF level before therapy was 2.72 ($n = 56$, median = 2.50, SD = 1.15, range 1–6), and after therapy it was 3.14 ($n = 44$; median = 3.00, SD = 1.25). The difference between pre- and post-ratings was 0.32 ($n = 41$, se = 0.15, z = 2.16, $p = .03$, 95% CI 0.03, 0.60). The correlation between change in RF and change in BDI-II was $r = .11$ (ns). The mean DSRF level before therapy was 2.44 ($n = 53$, median = 2.50, SD = 0.99, range 1–5) and at termination 2.53 ($n = 45$, median = 2.50, SD = 0.90). The pre- post difference was not statistically significant ($t = -0.26$, $p = .79$).[1]

We also ran multilevel models including therapist effects modeled as a random intercept as sensitivity tests. For RF, the estimated therapist effect was so close to zero that its standard errors could not be computed, while for DSRF it was statistically significant (random effect variance = 0.18, se = .13, 95% CI [0.04, 0.77]). However, including the random intercept for therapists for DSRF did not make any differences for the results.

Differences Between Treatments

A repeated measures ANOVA for RF ratings showed that the interaction effect for time × treatment was statistically significant ($n = 41$, $F(1,39) = 6.08$, $p = .018$). Multiple imputation confirmed the significant time × treatment interaction ($p = .03$). Marginal analysis indicated that if all patients had been treated with CBT, their RF level would be essentially unchanged after treatment (starting at 2.88 and ending at 2.79, a difference of −.09, se = 0.21, $p = .68$, 95% CI −0.52, 0.35). In contrast, had all patients been treated with IPT, RF would—according to this analysis—have increased from 2.67 to 3.27, a difference of .60 (se = 0.18, $p = .002$, 95% CI 0.24, 0.97). The difference in change over time between treatments (i.e., CBT change—IPT change) was −.67 (se = .28, $p = .018$, 95% CI −1.26, −.12). This

represents a medium sized effect (Cohen's $d = 0.61$). The interaction for time × treatment for DSRF did not show any difference between the treatments (z = −0.42, $p = .67$), and including random intercepts for therapists did not change this.

The correlation between pre–post change in RF and pre–post change on the BDI-II was $r = -.06$ for CBT patients and $r = .24$ for IPT patients (both ns). For DSRF, the correlation between pre–post change in DSRF and pre–post change BDI-II for 21 patients in the IPT group was $r = .38$ ($p = .09$). For 17 CBT patients, the correlation was $r = .17$ (ns). None of these results were changed if random intercepts for therapists were included in the model.

Discussion

The purpose of this study was to analyze change in RF and DSRF from before to after IPT and CBT. The analyses indicated that RF, but not DSRF, improved during treatment. However, improvement in RF was only found among patients who received IPT. The fact that RF increased in IPT and not in CBT suggests that the two treatments, although similarly effective in reducing depression symptoms, differ in their capacity to improve mentalization. Although neither treatment had the explicit purpose of improving mentalization, IPT includes processes that may be thought of as more mentalization-enhancing than CBT (Markowitz et al., 2019). For instance, the focus on connecting depressive mood to interpersonal relationships, the encouragement to develop the social network or to solve interpersonal problems, and the emphasis on role playing in sessions, might be promotive of mentalization to a higher extent than behavior activation, cognitive restructuring etc. It would be of value in process-oriented studies to analyze more precisely in what way interventions and therapeutic stance differ in these treatments with respect to their capacity to contribute to increased mentalization.

There was no correlation between change in depressive symptoms and change in RF. On the one hand, this is important since it indicates that improvement of mentalization was not simply an effect of improvement in depression (i.e., it is an indication of discriminant validity of the RF measure). However, it also means that it is uncertain to what extent and in what way improvement in RF might be important for the depressed patients. On theoretical and clinical grounds, it can be argued that the capacity for mentalization is important for the depressed person's realistic evaluation of his or her relationships and self-image; and without mentalization, depressive images of self, others, and the world (Beck et al., 1979) will take the place of more complex, mentalized understanding of the self.

The mean RF level at end of therapy in this study, even in the IPT group, was lower than the level attained in transference-focused therapies with borderline patients reported by Levy et al. (2006). The remaining low level of RF at therapy termination indicates that the potentially preventive function of RF has not been attained by these relatively short therapies. However, if longer therapy, especially longer IPT, would entail that RF would increase to more normal levels remains to be seen.

There were some apparent limitations in this study. We chose to analyze only data from patients who completed treatment, as we considered that only those patients who had received the full treatments would have a fair chance of improving their capacity for mentalization. Even 14 sessions, which was the complete treatment according to the current study protocol, is probably a small "dose" of therapy for the purpose of enhancing mentalization. However, there was considerable attrition of patients who could be interviewed with the AAI/DSRF interview even among treatment completers. To some extent the use of multiple imputation guards against possible sources of bias due to attrition. Multiple imputation rests on the assumption that data is missing at random, which implies that some kinds of non-random missingness are allowed —in contrast to missing completely at random, which is assumed when subjects with missing data are deleted.

Another limitation concerns the somewhat eclectic form of CBT practiced in this study. Therapists were allowed to choose between behavioral and cognitive interventions according to their clinical intuition. Although this is how CBT is usually practiced in Sweden, it limits the internal validity of the study since we do not know if the treatments in the CBT group were more cognitive or more behavioral in orientation. In contrast, the IPT therapies were likely more homogenous. Finally, reliability of RF ratings was not determined for all interviews at post-treatment. However, we believe that since reliability was good both at pre-treatment for both measures, and at post-treatment for DSRF, it seems likely that reliability was acceptable overall.

In future studies in might be valuable to use interviews that are less focused on life history, as such interviews may be experienced as less stimulating to repeat with short time intervals. It will also be important to study RF as a buffer against future relapse/recurrence by testing whether improvement in RF during treatment predicts post-treatment symptom trajectories.

Note

[1] We also tested this using a two-level growth model, in which BDI was regressed on log(session) at Level-1, using an unstructured covariance matrix at Level-2. The pre-post difference score of RF/DSRF was entered at Level-2, correlated with random intercept and slope for the regression of BDI on log (session). We got the same result; that is, no significant correlation between change in RF/DSRF and BDI over time.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

ROLF HOLMQVIST ⬤ http://orcid.org/0000-0003-2093-2510
FREDRIK FALKENSTRÖM ⬤ http://orcid.org/0000-0002-2486-6859

References

Bateman, A. W., & Fonagy, P. (2004). *Psychotherapy for borderline personality disorder: Mentalization-based treatment*. Oxford University Press.

Beck, A. T., Rush, A., Shaw, B., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.

Beck, A. T., & Steer, R. A. (1996). *Beck depression inventory. Manual (Swedish version)*. Psykologiförlaget.

Clarkin, J. F., Yeomans, F. E., & Kernberg, O. F. (2006). *Psychotherapy for borderline personality: Focusing on object relations*. American Psychiatric Publishing, Inc.

Doornik, J. A., & Hansen, H. (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, *70*(s1), 927–939. https://doi.org/10.1111/j.1468-0084.2008.00537.x

Ekeblad, A., Falkenström, F., Andersson, G., Vestberg, R., & Holmqvist, R. (2016). Randomized trial of interpersonal psychotherapy and cognitive behavioral therapy for major depressive disorder in a community-based psychiatric outpatient clinic. *Depression and Anxiety*, *33*(12), 1090–1098. https://doi.org/10.1002/da.22495

Ekeblad, A., Falkenström, F., & Holmqvist, R. (2016). Reflective functioning as predictor of working alliance and outcome in the treatment of depression. *Journal of Consulting and Clinical Psychology*, 84(1), 67–78. https://doi.org/10.1037/ccp0000055

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Evans, M. D., Piasecki, J. M., Kriss, M. R., & Hollon, S. D. (1984). *Raters manual for the collaborative study psychotherapy rating scale*, form 6. Vol. 22161.

First, M. B., & Pincus, H. A. (1999). Classification in psychiatry: ICD-10 v. DSM-IV: A response. *The British Journal of Psychiatry*, 175(3), 205–209. https://doi.org/10.1192/bjp.175.3.205

Fischer-Kern, M., Fonagy, P., Kapusta, N. D., Luyten, P., Boss, S., Naderer, A., Blüml, V., & Leithner, K. (2013). Mentalizing in female inpatients with major depressive disorder. *Journal of Nervous and Mental Disease*, 201(3), 202–207. https://doi.org/10.1097/NMD.0b013e3182845c0a

Fischer-Kern, M., Tmej, A., Kapusta, N. D., Naderer, A., Leithner-Dziubas, K., Löffler-Stastka, H., & Springer-Kremser, M. (2008). Mentalisierungsfähigkeit bei depressiven patientinnen: Eine pilotstudie [The capacity for mentalization in depressive patients: A pilot study]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 54(4), 368–380. https://doi.org/10.13109/zptm.2008.54.4.368

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self*. Other Press.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual* (Version 5). University College London.

George, C., Kaplan, N., & Main, M. (1985). *The adult attachment interview*. Unpublished manuscript.

Katznelson, H. (2014). Reflective functioning: A review. *Clinical Psychology Review*, 34(2), 107–117. https://doi.org/10.1016/j.cpr.2013.12.003

Klerman, G., Weissman, M. M., Rounsaville, B., & Chevron, E. (1984). *Interpersonal psychotherapy of depression*. Basic Books.

Levinson, A., & Fonagy, P. (2004). Offending and attachment: The relationship between interpersonal awareness and offending in a prison population with psychiatric disorder. *Canadian Journal of Psychoanalysis*, 12(2), 225–251. https://doi.org/10.1002/j.2167-4086.2009.tb00406.x.

Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, 74(6), 1027–1040. https://doi.org/10.1037/0022-006X.74.6.1027">https://doi.org/10.1037/0022-006X.74.6.1027

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Luyten, P., Fonagy, P., Lemma, A., & Target, M. (2012). Depression. In A. W. Bateman & P. Fonagy (Eds.), *Handbook of mentalizing in mental health practice* (pp. 385–417). American Psychiatric Publishing, Inc.

Markowitz, J. C., Milrod, B., Bleiberg, K., & Marshall, R. D. (2009). Interpersonal factors in understanding and treating posttraumatic stress disorder. *Journal of Psychiatric Practice*, 15(2), 133–140. https://doi.org/10.1097/01.pra.0000348366.34419.28

Markowitz, J. C., Milrod, B., Luyten, P., & Holmqvist, R. (2019). Mentalizing in interpersonal psychotherapy. *American Journal of Psychotherapy*, 72(4), 95–100. https://doi.org/10.1176/appi.psychotherapy.20190021

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2010). *Behavioral activation for depression: A clinician's guide*. Guilford Press.

McIntyre, R. S., Cha, D. S., Soczynska, J. K., Woldeyohannes, H. O., Gallaugher, L. A., Kudlow, P., Alsuwaidan, M., & Baskaran, A. (2013). Cognitive deficits and functional outcomes in major depressive disorder: Determinants, substrates, and treatment interventions. *Depression and Anxiety*, 30(6), 515–527. https://doi.org/10.1002/da.22063

Möller, C., Falkenström, F., Holmqvist Larsson, M., & Holmqvist, R. (2014). Mentalizing in young offenders. *Psychoanalytic Psychology*, 31(1), 84–99. https://doi.org/10.1037/a0035555

Montgomery, S. A., & Åsberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134(4), 382–389. https://doi.org/10.1192/bjp.134.4.382

Rudden, M. G., Milrod, B., Meehan, K. B., & Falkenstrom, F. (2009). Symptom-specific reflective functioning: Incorporating psychoanalytic measures into clinical trials. *Journal of the American Psychoanalytic Association*, 57(6), 1473–1478. https://doi.org/10.1177/00030651090570060804

Rudden, M. G., Milrod, B., Target, M., Ackerman, S., & Graf, E. (2006). Reflective functioning in panic disorder patients: A pilot study. *Journal of the American Psychoanalytic Association*, 54(4), 1339–1343. https://doi.org/10.1177/00030651060540040109

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2013). *Mindfulness-based cognitive therapy for depression* (2nd ed.). Guilford Press.

Skårderud, F. (2007a). Eating one's words, part I: 'Concretised metaphors' and reflective function in anorexia nervosa–an interview study. *European Eating Disorders Review*, 15(3), 163–174. https://doi.org/10.1002/erv.777

Skårderud, F. (2007b). Eating one's words, part II: The embodied mind and reflective function in anorexia nervosa—theory. *European Eating Disorders Review*, 15(4), 243–252. https://doi.org/10.1002/erv.778

Skårderud, F. (2007c). Eating one's words: Part III. Mentalisation-based psychotherapy for anorexia nervosa—an outline for a treatment and training manual. *European Eating Disorders Review*, 15(5), 323–339. https://doi.org/10.1002/erv.817

StataCorp. (2019). *Stata statistical Software: Release 16*. StataCorp LLC.

Taubner, S., Hörz, S., Fischer-Kern, M., Doering, S., Buchheim, A., & Zimmermann, J. (2013). Internal structure of the reflective functioning scale. *Psychological Assessment*, 25(1), 127–135. https://doi.org/10.1037/a0029138

Taubner, S., Kessler, H., Buchheim, A., Kächele, H., & Staun, L. (2011). The role of mentalization in the psychoanalytic treatment of chronic depression. *Psychiatry: Interpersonal and Biological Processes*, 74(1), 49–57. https://doi.org/10.1521/psyc.2011.74.1.49

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the beck depression inventory-II: A comprehensive review. *Brazilian Journal of Psychiatry*, 35(4), 416–431. https://doi.org/10.1590/1516-4446-2012-1048

Weissman, M. M., Markowitz, J. C., & Klerman, G. L. (2000). *Comprehensive guide to interpersonal psychotherapy*. Basic Books.