

Running Head: MENTALIZING AS A MECHANISM OF CHANGE IN BPD

Mentalizing as a Mechanism of Change in the Treatment of Patients with Borderline Personality
Disorder:
A Parallel Process Growth Modeling Approach

Celine De Meulemeester^a, Kristof Vansteelandt^b, Patrick Luyten^{a,c}, and Benedicte Lowyck^b

in press in: *Personality Disorders: Theory, Research and Treatment*

^a Faculty of Psychology and Educational Sciences, University of Leuven, Leuven, Belgium

^b University Hospitals Leuven, Leuven, Belgium

^c Research Department of Clinical, Educational and Health Psychology, University College
London, UK

Corresponding author:

Patrick Luyten, PhD

Faculty of Psychology and Educational Sciences

University of Leuven

Tiensestraat 102 pobox 3722

3000 Leuven

Belgium

Abstract

Although a number of effective psychotherapeutic treatments have been developed for borderline personality disorder (BPD), little is known about the mechanisms of change explaining the effects of these treatments. There is increasing evidence that impairments in mentalizing or reflective functioning—the capacity to reflect on the internal mental states of the self and others—are a central feature of BPD. To date, no study has directly investigated the core assumption of the mentalization-based approach to BPD, that changes in this capacity are associated with treatment outcome in BPD patients. This study is the first to directly investigate this assumption in a sample of 175 patients with BPD who received long-term hospitalization-based psychodynamic treatment. Using a parallel process growth modeling approach, this study investigated whether (a) treatment was related to changes in mentalizing capacity as measured with the Reflective Functioning Questionnaire; (b) these changes could be explained by pretreatment levels of mentalizing and/or symptomatic distress; and (c) changes in mentalizing capacity over time were associated with symptomatic improvement. Mentalizing and symptomatic distress were assessed at admission, 12 and 24 weeks into treatment, and at discharge. Results showed that treatment was associated with significant decreases in mentalizing impairments (i.e., uncertainty about mental states) and symptomatic distress. Pretreatment levels of mentalizing and symptomatic distress did not predict these changes. However, improvements in mentalizing were strongly associated with the rate of decrease in symptomatic distress over time ($r = .89$). These findings suggest that increases in mentalizing may indeed in part explain therapeutic change in the treatment of BPD, but more research is needed to further substantiate these conclusions.

Keywords: borderline personality disorder, mentalizing, psychodynamic psychotherapy, parallel process growth model

Borderline personality disorder (BPD) is described in DSM-5 (American Psychiatric Association, 2013) as a pervasive pattern of instability in interpersonal relationships, self-image, and affect, and marked impulsivity. The disorder is often comorbid with other mental illnesses and is related to high levels of self-harm and suicidality, resulting in large direct and indirect personal and economic costs ((Black, Blum, Pfohl, & Hale, 2004; Grant et al., 2008; Soeteman, Hakkaart-van Roijen, Verheul, & Busschbach, 2008).

The mentalizing approach to BPD pathology has attracted increasing attention in the past two decades. *Mentalizing* or *reflective functioning* is a form of social cognition that refers to the capacity to reflect on internal mental states (e.g., feelings, wishes, attitudes, and goals) of the self and others (Choi-Kain & Gunderson, 2008; P Fonagy & Bateman, 2008; P Fonagy & Luyten, 2009). Studies investigating mentalizing in BPD have found a paradoxical combination of hypersensitivity to emotional states of others based on their external features (e.g., facial expression) and gross impairments in the capacity to reflect on internal mental states of both self and others (P Fonagy & Luyten, 2009; Sharp & Vanwoerden, 2015).

The mentalization-based approach to BPD provides a theoretically comprehensive framework for understanding BPD and has led to the development of Mentalization-Based Treatment (MBT) for BPD patients (P Fonagy & Bateman, 2008). In MBT, the focus is on strengthening patients' mentalizing capacity, which is thought to improve affect regulation and interpersonal functioning (Fonagy & Bateman, 2006).

Moreover, the mentalizing approach argues that mentalizing is a fundamental psychological process that functions as a common factor in treatments for BPD patients, and that improvement in mentalizing capacity is therefore a central mechanism of change in any successful treatment for BPD (Fonagy, Luyten, & Bateman, 2015).

Although there is some preliminary evidence for this latter assumption, no study to date has directly investigated whether changes in mentalizing capacity parallel improvements in symptoms and complaints. Antonsen, Johansen, Rø, Kvarstein and Wilberg (2016) for example, reported that impairments in mentalizing were associated with symptomatic and interpersonal distress, and lower psychosocial and personality functioning at baseline, but were not related to clinical outcome at 6-year follow-up. However, in this study, trajectories of change in mentalizing and outcome were not directly compared. Bo et al. (2016), in turn, reported that mentalization-based group therapy was associated with improvements in mentalizing in a sample of adolescent female BPD patients. Specifically, in a sample of 34 Danish, female adolescent (age 15-18 years) BPD patients, 1 year of structured mentalization-based group treatment was associated with improvements in both BPD symptoms, as measured with the Borderline Personality Features Scale for Children (BPFS-C, Crick, Murray-Close, & Woods, 2005), and in mentalizing, as measured with the Reflective Functioning Scale for Youth (RFQ-Y, Sharp et al., 2009). Within-person changes from baseline to end of treatment were evaluated with paired sample t-tests. However, this study did not directly evaluate the association between improvements in mentalizing and improvement in outcome.

The Present Study

This study is the first to directly investigate the association between changes in reflective functioning and outcome in a sample of 175 BPD patients using data from a naturalistic study of the effectiveness of hospitalization-based treatment for BPD. Patients were assessed at four time points: at admission, 12 and 24 weeks into treatment, and at discharge. Mentalizing was measured using the Reflective Functioning Questionnaire (RFQ), a self-report questionnaire comprising two scales (Fonagy et al., 2016). The first scale assesses uncertainty about mental

states (RFQ-U) and is assumed to capture *hypomentalizing* or an inability to create mental models of the mind of self and others. The second subscale captures individuals' certainty about mental states (RFQ-C), which in the extreme characterizes *hypermentalizing* or excessive mentalizing (Fonagy et al., 2016).

Using a parallel process multilevel growth modeling approach, this study investigated whether (a) there is a significant decrease in mentalizing problems and symptomatic distress over the course of treatment, (b) pretreatment levels of mentalizing are associated with symptomatic change, and (c) changes in mentalizing capacity are related to change in symptoms during treatment.

Methods

Participants and Procedure

This study is part of a larger, ongoing process-outcome study. The data used in this study were collected between July 2010 and April 2014. During this period, 207 patients consecutively admitted to an intensive psychodynamic hospitalization-based treatment for personality disorders were studied. Inclusion criteria were (a) a principal diagnosis of a personality disorder given by a trained and experienced psychiatrist based on an extensive intake interview, and (b) Dutch literacy. Individuals with an acute psychotic episode or severe substance abuse were excluded. For this study, we selected only patients who fulfilled criteria for DSM-IV-defined BPD (American Psychiatric Association, 2000) based on the self-report questionnaire of the Structured Clinical Interview for DSM-IV Axis II disorders (SCID-II-SQ, Spitzer, Williams, Gibbon & First, 1990) (i.e., scores ≥ 5) that was administered at intake. This resulted in a sample of 175 BPD patients.

Their mean age was 30 years old ($SD = 9.13$) and 104 patients (59.4%) were female. The patients stayed in treatment for a mean of 6.76 months ($SD = 4.56$). Patients were relatively well educated: most of them (82.9%) completed secondary school and 48.6% completed some form of higher education (20% had a university degree). Thirty three percent of the patients lived alone, 35.4% lived with their parents, and 23.8% were married or living with a partner.

Comorbidity of BPD with symptom disorders as assessed with the Psychiatric Diagnostic Screening Questionnaire (PDSQ; Zimmerman & Mattia, 2001) was high. The most commonly reported psychiatric symptoms were depression (82.8% of patients scoring above the clinical cutoff), social phobia (79.5%), generalized anxiety disorder (78.2%) and somatization (63.6%). Comorbid personality disorders were assessed with SCID-II-SQ. Comorbidity in this sample was very high, with 80% of patients scoring above the cutoff for more than 4 different personality disorders. Depressive (83%), avoidant (80.7%) and obsessive-compulsive (73.3%) personality disorder were most common.

After receiving an explanation about the study (provided by a research assistant) and giving informed consent, patients entered the study and were assessed at admission, 12 and 24 weeks into treatment, and at discharge. This study was approved by the ethical committee of KU Leuven.

Treatment

The treatment provided has been described in detail elsewhere (Vermote, Lowyck, & Vandeneede, 2011). It is an intensive, hospitalization-based treatment for patients with personality disorder, consisting of the following elements: (a) group psychodynamic psychotherapy (three times a week), (b) nonverbal therapies (music therapy, psychomotor therapy, and creativity therapy, each twice a week), (c) family therapy (on indication), (d)

psychiatric consultation, (e) weekly individual sessions with a nurse, (f) group sessions with nurses (twice a week), (g) social work, and (h) a weekly patient–staff meeting. The treatment focuses on three central features of patients with (borderline) personality disorder: (a) difficulties with trust, (b) mentalizing impairments, and (c) impairments in the capacity for relatedness. It integrates mentalization-based interventions and principles with more traditional, insight-oriented psychoanalytic work.

Measures

Reflective Functioning Questionnaire (RFQ). The RFQ (Fonagy et al., 2016) is a self-report measure for reflective functioning or mentalizing. The RFQ consists of eight items in total that are answered on a 6-point Likert scale, ranging from 1 (*I do not agree at all*) to 6 (*I very much agree*). The RFQ has two subscales, Certainty and Uncertainty, each containing six items. Four items are used to calculate scores on both subscales, while the other four items are unique to each subscale (see below, and see Fonagy et al., 2016, for a detailed description).

Specifically, the items of the RFQ_Uncertainty (RFQ-U) scale are rescored so that high scores reflect extreme uncertainty about mental states, assumed to assess hypomentaling. For example, the item “*Sometimes I do things without really knowing why*” is recoded 0, 0, 0, 0, 1, 2, since strong agreement reflects a lack of knowledge about mental states. The items of the RFQ_Certainty (RFQ-C) subscale are rescored to capture certainty about mental states. For example, “*I don’t always know why I do what I do*” is recoded 2, 1, 0, 0, 0, 0, because strong disagreement with this statement reflects more certainty about mental states. While the RFQ-U scale has been consistently positively related to other measures assessing impairments in mentalizing (such as alexithymia) and to measures of personality pathology, findings with the RFQ-C scale suggest that it taps into both adaptive and maladaptive features of mentalizing. For

example, the RFQ-C scale has been shown to be positively related to eating disorder symptoms, but also to empathy, and negatively to alexithymia (Badoud et al., 2015; Fonagy et al., 2016). Furthermore, the RFQ-U scale and, to a somewhat lesser extent, the RFQ-C scale have been found to distinguish between normal controls and BPD patients. The internal consistency for both RFQ-U and RFQ-C was good, with Cronbach's alpha values of 0.77 and 0.65, and test-retest reliability over 3 weeks was very good, with $r_s = 0.84$ and 0.75 for RFQ-U and RFQ-C, respectively (Fonagy et al., 2016). In this study, Cronbach's alpha values were 0.68 for RFQ-U and 0.70 for RFQ-C at baseline.

Brief Symptom Inventory. The Dutch version of the Brief Symptom Inventory, the Korte Klachten Lijst (KKL, Lange & Appelo, 2007), is a self-report questionnaire measuring symptomatic distress experienced in the past week, expressed in the following domains: anxiety, memory difficulties, depressive feelings, somatic complaints, irritability, suicidal thoughts, problems in relational functioning, eating disorders, self-harm, sexual problems, sleeping problems, and substance abuse. The questionnaire consists of 13 questions, scored on a 5-point Likert scale, ranging from 0 (*no complaints*) to 4 (*many complaints*). The total score (ranging from 0–52) was calculated by summing the scores for individual questions. The validity and reliability of the inventory have proven to be satisfactory (Lange & Appelo, 2007). In this study, the internal consistency of the KKL at baseline was $\alpha = 0.75$.

SCID-II-SQ. BPD was assessed using a self-report version of the SCID-II (SCID-II-SQ; Spitzer et al., 1990). The SCID-II-SQ consists of 119 yes-or-no questions based on the diagnostic criteria of DSM-IV Axis II. In accordance with the interview version of the SCID-II, a cut-off score (5 or more out of 15 items) for the BPD scale was applied. The agreement between the SCID-II-SQ and the SCID-II interview in terms of number of criteria fulfilled was found to

be high ($r = 0.84$; Ekselius, Lindström, Knorrning, Bodlund, & Kullgren, 1994). There was adequate test–retest reliability (Ouimette & Klein, 1995) and the internal consistency of the SCID-II-SQ in this study was very good ($\alpha = 0.89$).

PDSQ. The Psychiatric Diagnostic Screening Questionnaire (PDSQ; Zimmerman & Mattia, 2001) is a self-report questionnaire consisting of 126 items that can be used to screen for common DSM-IV Axis I psychiatric disorders. The psychometric quality of the PDSQ has been established in two large-scale validation studies (Zimmerman and Chelminsky, 2006; Zimmerman and Mattia, 2001). The total score of the PDSQ was found to be highly reliable in the present study ($\alpha=0.96$).

Data Analysis

This study applied a multivariate multilevel parallel growth model for change (MacCallum, Kim, Malarkey, & Kiecolt-Glazer, 1997; Singer & Willett, 2003) with repeated measurements nested within subjects to study (a) whether mentalizing capacity and symptoms change over the course of treatment, (b) whether pretreatment mentalizing and symptomatic distress are predictive of changes in symptomatic distress over the course of treatment, and (c) whether the process of change in mentalizing and symptomatic distress are related to one another. The first three assessments were scheduled at the same times for all patients: at admission/baseline, after 12 weeks (3 months), and after 24 weeks (6 months). The final assessment, at discharge, was variable. In the model mentioned above, time is a continuous variable expressed in terms of months since admission (see Appendix 1 for more details).

Two separate models were estimated for RFQ-C and symptomatic distress on the one hand, and for RFQ-U and symptomatic distress on the other hand. The best-fitting model for

each of the different outcome variables was selected on the basis of graphical inspection of the data, likelihood ratio tests, and information criteria (Bryk & Raudenbush, 1992; Singer & Willett, 2003; Verbeke & Molenberghs, 2000). For RFQ-C, a model with random subject-specific intercepts and a fixed linear slope for time was selected. For RFQ-U, a model with random subject-specific intercepts and random subject-specific linear slopes for time was selected. Finally, for symptomatic distress, a model with random subject-specific intercepts, random subject-specific linear slopes for time, and a fixed quadratic term for time was used.

Individual differences in patterns of change of mentalizing and symptoms over the course of treatment were studied by examining the fixed effects and the covariance parameters of the random subject-specific intercept and slope parameters of the involved variables. To investigate the parallel processes of the different variables, the model yielded not only estimates of fixed effects, random variance parameters, random covariance parameters, and error terms for each outcome variable separately (see research question (a)), but also random covariances for parameters representing *different* outcome variables (see research questions (b) and (c)). For example, the covariance between the random subject-specific slopes for time for RFQ-U on the one hand, and the random subject-specific slopes for time for symptomatic distress on the other hand, indicates whether subjects are characterized by the same linear change over time on both outcome variables. It should be noted that, as missing data are almost inevitable in longitudinal studies, inference for the proposed models is valid under the assumption of missingness at random (Little & Rubin, 2002). All models were estimated using SAS version 9.4 (Statistical Analysis Software, SAS Institute Inc., Cary, NC, USA).

Finally, effect sizes were calculated to evaluate treatment effects using pre–post standardized mean gain scores (Becker, 1988). These scores can be interpreted in the same way

as Cohen's d (Cohen, 1988): $d \leq 0.20$ is considered a small effect size, $d \approx 0.50$ a medium effect size, and $d \geq 0.80$ a large effect size.

Results

Relationship between Mentalizing and Symptomatic Distress at Baseline

As shown in Table 1, levels of uncertainty about mental states (RFQ-U) were significantly related to symptomatic distress (KKL) at baseline ($r = 0.45, p < 0.001$). Levels of certainty about mental states (RFQ-C) were, however, not related to symptomatic distress at baseline (Table 2; $r = -0.17, p = 0.15$).

Changes in Mentalizing and Symptomatic Distress during Treatment

Levels of uncertainty about mental states (RFQ-U) decreased linearly and significantly during treatment (Table 3; $\beta = -0.13, SE = 0.03, p < .0001$; see thick average line in Figure 1a). There were substantive individual differences both in the level of RFQ-U at baseline (i.e. the intercepts) and in the decrease in RFQ-U during treatment (i.e. the slopes), as shown in Figure 1a. The pre–post treatment standardized mean gain score for RFQ-U had a medium effect size of 0.52. The model for the levels of certainty about mental states (RFQ-C) had random intercepts and a fixed (but nonsignificant) slope for time. As a result, there were individual differences in the RFQ-C at baseline but not in the change—either on average, or individually—of RFQ-C over time (Table 4; $\beta = 0.009, SE = 0.02, p = 0.63$). In line with these results, the pre–post treatment effect size for RFQ-C was -0.10.

Symptomatic distress decreased significantly over the course of treatment; however, a quadratic effect of time was observed, which indicates that the linear decrease of symptoms over the course of treatment leveled off toward the end (Table 3; β for time = -1.25, $SE = 0.20, p < .0001$; β for time² = 0.04, $SE = 0.02, p = 0.02$; see thick average line in Figure 1b). There were

also substantive individual differences both in symptomatic distress at baseline and in the linear decrease of symptoms during treatment (Figure 1b). The pre–post standardized treatment main gain score was large ($d = 1.07$).

Role of Pretreatment Levels of Mentalizing and Symptomatic Distress on Rate of Change during Treatment

Results showed no significant associations between patient-specific (PS) intercepts of RFQ-U and PS slopes of KKL, which indicates that pretreatment levels of RFQ-U were not associated with changes in symptoms during treatment (see Table 1). There was, however, a trend toward significance ($r = -0.40, p = 0.07$) of associations between pretreatment levels of RFQ-U and changes in RFQ-U. In other words, there was a tendency for patients with higher levels of RFQ-U at baseline to show a faster decrease of RFQ-U during treatment and for patients with lower RFQ-U at baseline to show a lower decrease during treatment.

As shown in Table 2, pretreatment levels of RFQ-C did not predict changes in symptoms. Similarly, no association was found between pretreatment levels of symptomatic distress and changes in symptoms during treatment.

Mentalizing and Symptomatic Distress over the Course of Treatment

As shown in Table 1, with regard to RFQ-U, results showed a strong and highly significant positive correlation between PS slopes for RFQ-U and PS slopes for symptomatic distress ($r = .89, p = 0.01$). In other words, a linear decrease in RFQ-U was associated with a decrease in symptomatic distress during treatment.

To illustrate this finding, the parallel process between level of RFQ-U and symptomatic distress is plotted for four patients in Figure 2. Patients who showed a strong decrease in RFQ-U (e.g., patients 3 and 4 in Figure 2a) also showed a strong decrease in symptoms (Figure 2b),

whereas patients who showed less or no reduction in RFQ-U (e.g., patients 1 and 2 in Figure 2a) also showed a smaller reduction in symptomatic distress (Figure 2b).

Results showed no change in RFQ-C during treatment either on average or at an individual patient level. As a result, there was no evidence for a parallel process between RFQ-C and symptomatic distress, as RFQ-C did not change over time.

Discussion

Findings of this study suggest that an increase in the capacity for mentalizing in the course of an intensive specialized treatment explained in part changes in symptomatic distress, consistent with the assumptions of mentalization-based approaches to BPD.

In particular, results showed that treatment was associated with a significant decrease in uncertainty about mental states (hypomentalizing), which indicates an improved capacity to reflect on the internal mental states of the self and others. However, in contrast to the mentalization-based approach to BPD (Fonagy & Bateman, 2008), the current treatment was not associated with a decrease in certainty about mental states. On the one hand, this could mean that there was no improvement in hypermentalizing, which is typical of many BPD patients (Sharp & Vanwoerden, 2015). However, it could also be the case that the RFQ-C scale of the RFQ does not succeed in capturing the excessive certainty about mental states that is assumed to characterize hypermentalizing in BPD. As noted, previous research with the RFQ has shown that scores on the RFQ-C were related to eating disorder symptoms, which have been associated with a tendency toward hypermentalizing (Fonagy et al., 2016; Skårderud, 2007). However, the RFQ-C score has also been inversely related to borderline symptoms, depressive symptoms, and severity of personality pathology (Fonagy et al., 2016). For this reason, it remains somewhat unclear whether the RFQ-C assesses hypermentalizing or whether it also partly captures features

of genuine mentalizing. Further studies are needed to investigate the relationship between the RFQ-C subscale and other measures of hypermentalizing, such as the excessive theory of mind scale of the Movie Assessment for Social Cognition (MASC; Dziobek et al., 2006; Sharp & Vanwoerden, 2015).

Furthermore, pretreatment levels of mentalizing (as well as symptomatic distress) did not predict symptom change over the course of treatment. So, even though there were substantial individual differences between BPD patients in symptom change, suggesting that the treatment was more effective for some patients than for others, these differences could not be predicted by the variables included in this study. There was, however, a tendency for patients with higher levels of RFQ-U at baseline to show a faster decrease in RFQ-U score during treatment, suggesting that the patients' mentalizing capacity at the start of treatment did have an effect on how fast their mentalizing capacity improved during treatment. As this finding might reflect regression to the mean, further research is needed to replicate these findings.

Finally, and most importantly, results showed that the rate of decrease in RFQ-U was associated with the rate of decrease in symptoms, indicating that improvement in mentalizing capacity coincides with improvement in symptoms. This relationship was very strong ($r = 0.89$). Although further research is needed to replicate this finding, the improved capacity to reflect on internal mental states of oneself and others (expressed as less uncertainty about mental states) appears to make patients feel better in terms of experienced symptoms over the course of treatment. Together, these findings suggest that mentalizing may be an important mechanism of change in the treatment of BPD patients.

Despite its strengths, this study had several important limitations. First, the primary outcome was a brief self-report measure of general symptomatic distress in general (the KKL)

rather than specific BPD pathology. Although this measure includes some characteristic BPD symptoms, such as self-harm, suicidality, and interpersonal problems, future studies should include a more direct measure of BPD features as primary outcome. In spite of this limitation, the study results are nevertheless informative for both research and clinical practice, particularly as a number of studies have suggested that general symptomatic distress may be a good proxy of the general severity of personality disorder pathology, the so-called psychopathology or ‘p’ factor (Caspi et al., 2014; Patalay et al., 2015). Second, there was no control group in this study, and so it is impossible to rule out the possibility that the effects may be due to factors other than the treatment itself. Third, BPD diagnosis was based on patients’ self-report on the SCID-II-SQ (after clinical diagnosis of a personality disorder by a psychiatrist). Although there is some evidence suggesting high concordance with interview-based measures of personality disorder diagnosis (Ekselius et al., 1994), the reliance on patients’ self-reported BPD symptoms may have led to an overestimation of the true prevalence of BPD. Finally, mentalizing was assessed using only a self-report questionnaire. It would be interesting to see whether the treatment results in improved performance on other measures of mentalizing (e.g., the Movie for the Assessment of Social Cognition, Dziobek et al., 2006) and whether these improvements also predict outcome. Furthermore, the use of self-report questionnaires to measure both the outcome (symptomatic distress) and the mechanism of change (mentalizing) in this study may have inflated their association due to shared method variance. Finally, this study does not allow us to determine the relative importance of mentalizing as a mechanism of change because there was no assessment of other putative mechanisms of change. This is the focus of a future study. Further studies are also needed to investigate to what extent these findings may generalize to other treatments. The fact that the treatment in the current study resulted in changes in mentalizing, however, provides

some preliminary evidence that mentalizing may be a common factor producing therapeutic change in any successful intervention for BPD (Fonagy et al., 2015).

In conclusion, the results of this study show that in BPD patients, intensive hospitalization-based psychodynamic treatment is associated with significant improvements in mentalizing capacity expressed as less uncertainty about mental states, which coincides with, and may lead to, important reductions in symptomatic distress in these patients.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders DSM-IV* (4th ed., text rev.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed., text revision ed.). Washington, DC: American Psychiatric Association.
- Antonsen, B. T., Johansen, M. S., Rø, F. G., Kvarstein, E. H., & Wilberg, T. (2016). Is reflective functioning associated with clinical symptoms and long-term course in patients with personality disorders? *Comprehensive Psychiatry*, *64*, 46–58.
<https://doi.org/10.1016/j.comppsy.2015.05.016>
- Badoud, D., Luyten, P., Fonseca-Pedrero, E., Eliez, S., Fonagy, P., & Debban??, M. (2015). The French version of the reflective functioning questionnaire: Validity data for adolescents and adults and its association with non-suicidal self-injury. *PLoS ONE*, *10*.
<https://doi.org/10.1371/journal.pone.0145892>
- Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*, 257–278. <https://doi.org/10.1111/j.2044-8317.1988.tb00901.x>
- Black, D. W., Blum, N., Pfohl, B., & Hale, N. (2004). Suicidal behavior in borderline personality disorder: Prevalence, risk factors, prediction, and prevention. *Journal of Personality Disorders*, *18*, 226-239. <https://doi.org/10.1521/pedi.18.3.226.35445>
- Bo, S., Sharp, C., Beck, E., Pedersen, J., Gondan, M., & Simonsen, E. (2016). First empirical evaluation of outcomes for mentalization-based group therapy for adolescents with BPD. *Personality Disorders: Theory, Research, and Treatment*. Advance online publication.
<https://doi.org/10.1037/per0000210>
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park, CA: Sage.

- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, 2, 119–137.
<https://doi.org/10.1177/2167702613497473>
- Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: ontogeny, assessment, and application in the treatment of borderline personality disorder. *American Journal of Psychiatry*, 165, 1127–1135. <https://doi.org/10.1176/appi.ajp.2008.07081360>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Crick, N. R., Murray-Close, D., & Woods, K. (2005). Borderline personality features in childhood: a short-term longitudinal study. *Development and Psychopathology*, 17, 1051–1070. <http://doi.org/10.1017/S0954579405050492>
- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., ... Convit, A. (2006). Introducing MASC: A movie for the assessment of social cognition. *Journal of Autism and Developmental Disorders*, 36, 623–636. <http://doi.org/10.1007/s10803-006-0107-0>
- Ekselius, L., Lindström, E., Knorrning, L. V., Bodlund, O., & Kullgren, G. (1994). SCID II interviews and the SCID Screen questionnaire as diagnostic tools for personality disorders in DSM-III-R. *Acta Psychiatrica Scandinavica*, 90, 120-123.
<http://doi.org/10.1111/j.1600-0447.1994.tb01566.x>
- Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology*, 62, 411-430. <http://doi.org/10.1002/jclp.20241>
- Fonagy, P., & Bateman, A. W. (2008). The development of borderline personality disorder--a mentalizing model. *Journal of Personality Disorders*, 22, 4–21.
<http://doi.org/10.1521/pedi.2008.22.1.4>

- Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, *21*, 1355–1381. <http://doi.org/10.1017/S0954579409990198>
- Fonagy, P., Luyten, P., & Bateman, A. (2015). Translation: Mentalizing as Treatment Target in Borderline Personality Disorder. *Personality Disorders: Theory, Research and Treatment*, *6*, 380–392. <http://doi.org/10.1037/per0000113>
- Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., ... Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The reflective functioning questionnaire. *PLoS ONE*, *11*,. <http://doi.org/10.1371/journal.pone.0158678>
- Grant, B. F., Chou, S. P., Goldstein, R. B., Huang, B., Stinson, F. S., Saha, T. D., ... Ruan, W. J. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV borderline personality disorder: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*, *69*, 533–545. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/18426259>
- Lange, A., & Appelo, M. T. (2007). *Korte Klachten Lijst: KKL*. Houten, The Netherlands: Bohn Stafleu Van Loghum.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). London, UK: Wiley.
- MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215–253. <http://doi.org/10.1207/s15327906mbr3203>
- Ouimette, P. C., & Klein, D. N. (1995). Test-Retest Stability, Mood-State Dependence, and Informant-Subject Concordance of the SCID-AXIS II Questionnaire in a Nonclinical Sample. *Journal of Personality Disorders*, *9*, 105–111. <http://doi.org/10.1521/pedi.1995.9.2.105>

- Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *The British Journal of Psychiatry: The Journal of Mental Science*, 207(1), 15–22. <http://doi.org/10.1192/bjp.bp.114.149591>
- Sharp, C., & Vanwoerden, S. (2015). Hypermentalizing in Borderline Personality Disorder: A Model and Data. *Journal of Infant, Child, and Adolescent Psychotherapy*, 14, 33–45. <http://doi.org/10.1080/15289168.2015.1004890>
- Sharp, C., Williams, L. L., Ha, C., Baumgardner, J., Michonski, J., Seals, R., ... Fonagy, P. (2009). The development of a mentalization-based outcomes and research protocol for an adolescent inpatient unit. *Bulletin of the Menninger Clinic*, 73, 311–338. <http://doi.org/10.1521/bumc.2009.73.4.311>
- Singer, J. D., & Willett, J.B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Skårderud, F. (2007). 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, 323–339. <http://doi.org/10.1002/erv.817>
- Soeteman, D. I., Hakkaart-van Roijen, L., Verheul, R., & Busschbach, J. J. (2008). The economic burden of personality disorders in mental health care. *Journal of Clinical Psychiatry*, 69, 259–265. <https://www.ncbi.nlm.nih.gov/pubmed/18363454>
- Spitzer, R. L., Williams, J. B. W., Gibbon, W., & First, M. B. (1990). User's guide for the Structured Clinical Interview for DSM-III-R (SCID). *Washington, DC: American Psychiatric Press*.
- Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data. Springer series in statistics*. <http://doi.org/10.1007/978-1-4419-0300-6>
- Vermote, R., Lowyck, B., & Vandeneede, B. (2011). Mentalization-based treatment and its implementation in existing psychodynamically-oriented therapeutic settings: A dynamic interaction. In: Bateman, A. & Fonagy, P. (Eds.), *Handbook of mentalizing in mental*

health practice (pp. 247–269). Arlington, VA: American Psychiatric Publishing, Inc.

Zimmerman, M., & Chelminski, I. (2006). A scale to screen for DSM-IV Axis I disorders in psychiatric out-patients: performance of the Psychiatric Diagnostic Screening Questionnaire. *Psychological Medicine*, *36*, 1601–1611.

<http://doi.org/10.1017/S0033291706008257>

Zimmerman, M., & Mattia, J. I. (2001). A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Archives of General Psychiatry*, *58*, 787–794. <http://doi.org/10.1001/archpsyc.58.8.787>

Table 1. Correlation matrix for the random effects (i.e., patient-specific slopes and patient-specific intercepts) of a multivariate multilevel model for change: Uncertainty about Mental States and Symptomatic Distress.

		RFQ-U		Symptomatic Distress (KKL)	
		PS intercepts	PS slopes	PS intercepts	PS slopes
RFQ-U	PS-intercepts	1	-0.40	0.45***	-0.05
	PS-slopes		1	-0.25	0.89**
Symptomatic Distress (KKL)	PS-intercepts			1	0.03
	PS-slopes				1

Note: RFQ-U = Reflective Functioning Questionnaire – Uncertainty subscale, KKL= Korte

Klachten Lijst, PS = patient-specific, ** $p < .01$;*** $p < .001$.

Table 2. Correlation matrix of the random effects (i.e., patient-specific slopes and patient-specific intercepts) of a multivariate multilevel model for change: Certainty about Mental States and Symptomatic Distress.

		RFQ-C	Symptomatic Distress (KKL)	
		PS-intercepts	PS-intercepts	PS-slopes
RFQ-C	PS-intercepts	1	-0.17	-0.40
Symptomatic	PS-intercepts		1	0.10
Distress (KKL)	PS-slopes			1

Note: RFQ-C = Reflective Functioning Questionnaire – Certainty subscale, KKL= Korte

Klachten Lijst, PS = patient-specific, ** $p < .01$;*** $p < .001$.

Table 3. Regression coefficients for the fixed effects in a multivariate multilevel model for change: Uncertainty about Mental States and Symptomatic Distress

Fixed Effects	Regression Coefficients (SE)	
	RFQ-U	Symptomatic Distress (KKL)
Overall intercept	4.18 (0.20) ***	26.28 (0.65) ***
Past months (linear)	-0.13 (0.03) ***	-1.25 (0.20) ***
(Past months) ² (quadratic)	NA	0.04 (0.02) *

Note: RFQ-U = Reflective Functioning Questionnaire – Uncertainty subscale, KKL= Korte

Klachten Lijst, NA = Not applicable; * $p < 0.05$, *** $p < .001$.

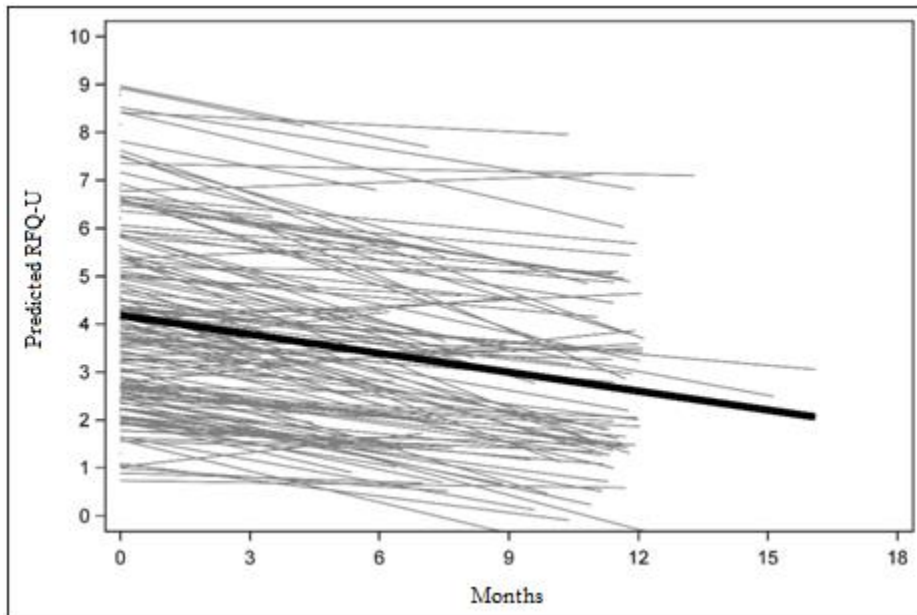
Table 4. Regression coefficients for the fixed effects in a multivariate multilevel model for change: Certainty about Mental States and Symptomatic Distress.

Fixed Effects	Regression Coefficients (SE)	
	RFQ-C	Symptomatic Distress
Overall intercept	1.67 (0.15) ***	26.27 (0.65)***
Past months (linear)	0.009 (0.02)	-1.27 (0.20)***
(Past months) ² (quadratic)	NA	0.04 (0.02) **

Note: RFQ-C = Reflective Functioning Questionnaire – Certainty subscale, KKL= Korte

Klachten Lijst, ** $p < 0.01$ *** $p < .001$; NA = Not applicable;.

(a) Change in RFQ-U score during treatment.



(b) Change in KKL score during treatment.

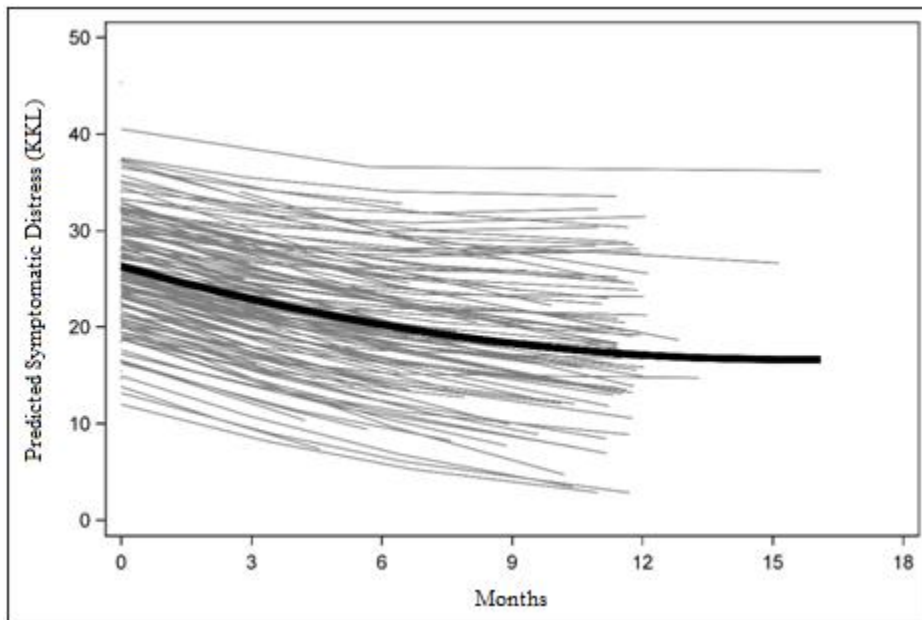
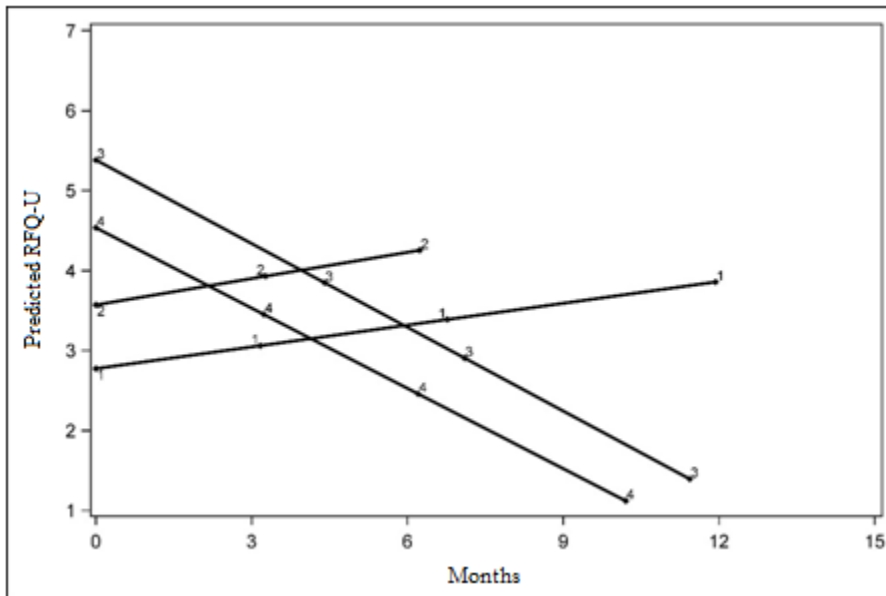


Figure 1. Average (thick black line) and patient-specific (thin black lines) changes in scores of 50 randomly selected patients for (a) RFQ-U and (b) symptomatic distress (KKL) during treatment.

(a) Patient-specific change in RFQ-U score during treatment.



(b) Patient-specific change in symptomatic distress (KKL) score during treatment.

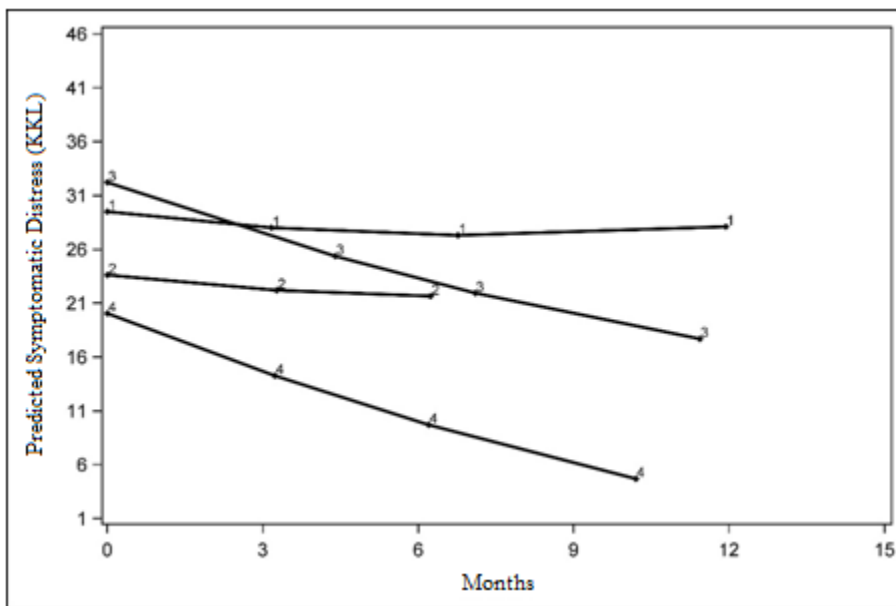


Figure 2. Change in (a) RFQ-U score and (b) symptomatic distress (KKL) score during treatment for four individual patients.