THESIS PROPOSAL

The dose of feedback as a moderator of treatment outcome:

Preliminary findings

Author:
Meropi Simou

Supervisors:
Dr. Kim de Jong, Assistant Professor
Dr. Eiko Fried, Assistant Professor
INTRODUCTION

Despite strong empirical support for the effectiveness of psychotherapy, not all patients seem to benefit from it (Lambert & Ogles, 2004; de Jong, 2012). Providing outcome monitoring feedback to therapists has been suggested as a viable means of improving treatment outcome, especially for clients who do not seem to progress well (Lambert et al., 2003; Bickman, 2008). Outcome monitoring typically consists of regular measurements of patients’ symptoms during treatment, hence allowing therapists to identify when their patients are not progressing, adjust their treatment plan early in therapy, and prevent negative therapy outcomes (de Jong, 2016; Bickman, 2008).

Although a growing body of empirical support over the last two decades has provided encouraging effects for the use of feedback on treatment outcome (Kazdin, 2008) and diagnostic accuracy (Carlier et al., 2010), the results in clinical practice appear to be less optimistic. Not only do clinicians often seem to overestimate their patients’ progress (Hannan et al., 2005), but also the effects of feedback on treatment outcome vary from large to moderate or even small (Sapyta, 2005; Shimokawa, 2010). Moreover, recent studies suggest that outcome monitoring is not equally effective under all circumstances and that several factors related to the implementation of feedback may be reducing its effectiveness (de Jong, 2016; Simon et al. 2012).

Bickman et al. (2011) found that patients whose therapists received session-by-session feedback improved faster than in cases where feedback was provided every three months. Furthermore, they found a dose–response effect, indicating stronger effects when therapists had more frequent access to the available patient measures. Another study by De Jong et al. (2012) showed that almost half of the therapists did not use the feedback given to them by their clients at all, and that a series of “therapist factors” seem to influence the use of feedback by therapists, such as the therapist’s feedback propensity, commitment to use the feedback, or perceived self-efficacy.

In order to investigate the moderating role of feedback dose on treatment outcome, Bickman et al. (2015) calculated an Implementation Index, taking into account the potential amount of feedback received by therapists. The Implementation Index consists of two dimensions: (1) questionnaire completion rate of the patients
and (2) viewing rate of the feedback by the therapists. Failure to fill out the
questionnaires and failure to view the feedback can both be seen as a sign of
implementation failure. An important result from Bickman’s study (2015) was that the
Implementation Index moderated the outcomes of the feedback intervention.

Building upon previous findings, a Dutch prediction model for patients in the
Netherlands was created, based on almost 2000 patients in four mental health care
organizations (de Jong, in revision). The Dutch model uses the initial severity of
patient’s dysfunction, as well as patient’s expectancies on treatment outcome as
predictors for progress. At the same time, preliminary results showed that this was
also a significant predictor for progress in the data that was collected. Bearing in mind
the variance in empirical results regarding the effects of feedback on treatment
outcome and the main considerations about feedback implementation mentioned by
Bickman et al. (2015), the current study used the Dutch prediction model and
investigated whether stronger treatment effects can be achieved when feedback is
implemented more frequently by therapists.

Research questions and hypotheses

The main objective of this study was to examine whether the dose of feedback,
assessed with the Implementation Index, moderates the effectiveness of outcome
feedback. Using patient data from a randomized controlled trial on the effectiveness
of feedback (de Jong, in revision), two forms of feedback and one control condition
were compared.

The first feedback condition consisted of standard treatment with routine
outcome monitoring (ROM), where feedback on the patient’s progress was given to
the therapist in the form of a graph. In the graph, the patient’s progress on the OQ-45
was monitored and the therapist received feedback on the OQ total score, the subscale
scores and the “critical items” (suicidal ideation, substance abuse, violence at work).

The second feedback condition consisted of standard treatment with feedback
on the patient’s progress, based on an expected treatment result (ETR) model.
According to the Dutch ETR model, which was used in the main study by de Jong et
al., therapists receive an alert when patients deviate from their expected track. An
error bound around the expected treatment result for the patient indicates when the
therapist is signaled. If 75% failure boundary is crossed by the patient, the therapist

...
receives an orange warning signal, indicating that the patient has an increased chance of poor treatment outcome. If the 95% failure boundary is crossed, the therapist receives a red warning signal, indicating that the patient is most likely to deteriorate if he/she continues being “off track”. In case of an orange or red alert, an additional questionnaire regarding possible problematic domains (Assessment for Signal Clients) is administered to the patients. Feedback including this information is then provided to the therapist, in addition to the patient’s OQ scores.

In the control condition, therapists received no feedback about their patient’s progress at all. Based on previous feedback literature, we hypothesized the following:

i. Patients in the two feedback conditions were expected to have lower OQ-45 scores (better outcomes) than patients in the control group at post-treatment evaluation.

ii. Patients in the ETR feedback condition were expected to have lower OQ-45 scores (better outcomes) than patients in the ROM feedback condition at post-treatment evaluation.

iii. The Implementation Index was expected to have a predicting/moderating effect on treatment outcome; higher implementation was expected to result in lower OQ-45 scores.

METHOD

This study was part of a two-year randomized controlled clinical trial on the effectiveness of feedback interventions for therapists working in three treatment facilities in the Netherlands. Psychiatric patients between 18 and 65 years who applied for outpatient individual psychiatric treatment in one of the four research locations (Schagen, Amici, Deventer, Twello) were asked to participate in the study and then randomly allocated to one of the three research conditions. The design in the study was single-blinded due to the nature of the intervention and hence therapists were aware of the research condition of their patients.

Participants

During the inclusion period, 1301 adult patients seeking treatment were assessed for eligibility, of which 790 did not continue to the randomization phase
either because they did not meet the inclusion criteria, they refused to participate in the study, or were not eligible for other reasons. The remaining 511 participants were then randomly assigned in the three study conditions (Figure 1), with some participants not receiving the experimental manipulation (n=17) and some others providing invalid or incomplete data that could not be used for the present analyses (n=70), resulting in a final sample of 424 patients who were finally included and reported in this paper. Of those, 142 were allocated in the control/no-feedback condition, 155 in the ROM feedback condition, and 126 in the ETR feedback condition (Table 1). The total sample included 168 male (39.6%) and 255 female (60.1%) participants. Gender distributions in the three study conditions were not significantly different [$\chi^2(2) = 1.24, p = .539$].

Patients were excluded from the study if they had a psychotic disorder as main diagnosis, had a current severe manic episode, a high risk of decompensation, or insufficient language skills in Dutch reading and speaking. Patients could have supplement treatments beside their individual treatment and stay in the study, as long as the individual treatment remained the main treatment.

The investigator could also withdraw patients from the study for urgent medical reasons, or if the patient’s diagnosis or treatment changed in a way that the patient no longer fitted the inclusion criteria of the study, but this was not the case for any patients in the study. Patients could leave the study at any time without any consequences, however there is no information available regarding whether some patients discontinued their participation or simply ended therapy, nor information about the reasons therefore.

The original study examining the effectiveness of feedback included approximately sixty therapists, consisting of psychotherapists, psychologists or psychiatric nurses who worked at the four research sites. Therapist characteristics, such as years of clinical experience, were analyzed and discussed elsewhere (de Jong, in revision), thus they were not the focus of the present study.
Figure 1. Flowchart of participants

Excluded from the analyses because of incomplete or invalid data ($n = 21$)

Included in the analyses
Table 1. Characteristics of the participants

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Male n (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female n (%)</td>
</tr>
<tr>
<td>Total sample</td>
<td>423</td>
<td>168 (39.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>255 (60.1)</td>
</tr>
<tr>
<td>No-feedback group</td>
<td>142</td>
<td>58 (40.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>84 (59.2)</td>
</tr>
<tr>
<td>ROM feedback group</td>
<td>155</td>
<td>65 (41.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90 (57.7)</td>
</tr>
<tr>
<td>ETR feedback group</td>
<td>126</td>
<td>45 (35.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>81 (64.3)</td>
</tr>
</tbody>
</table>

**Instruments**

**Patient Measures**

_Demographic questionnaire_. Upon inclusion in the study, patients were asked to fill out a questionnaire that assessed demographic characteristics, including information such as patient’s age, gender, nationality, marital status, living and working situation, educational level, prior treatment, pre-treatment use of medication, the main complaint and the duration of the main complaint. Due to administrative issues, participants’ demographic information was not available in the present study, with the exception of patients’ gender, which was also included in the outcome questionnaire.

_Psychiatric assessment_. As part of the standard intake procedure, all patients who were asked to participate in the study were interviewed with the MINI-Plus (Sheehan et al., 1998). Patients who agreed to participate were also administered with the SCID-II PQ (First et al., 1997) by the research assistant, before being randomly assigned to one of the three research conditions. Both the MINI-Plus and the SCID-II PQ are described in detail elsewhere (de Jong, in revision).

_Outcome Questionnaire-45 item version (OQ-45; Lambert, 1996)._ One of the main outcomes of the study was the level of the patient’s psychological dysfunction, measured by the Dutch version of the Outcome Questionnaire-45 (OQ-45). The OQ-45 is a self-report instrument consisting of 45 items, asking how often the respondent
has felt in a specific way over the last week on a 5-point Likert scale, ranging from 0 (never) to 4 (almost always). The OQ-45 comprises of three subscales that assess different domains of client functioning: Symptom Distress, Interpersonal Relations and Social Role, and includes three “critical items” which alert the therapist about potential risk for suicide, substance abuse and violence. The 25-item Symptom Distress domain (25 items) assesses common psychological symptoms in highly prevalent mental disorders; an example item of this subscale is “I blame myself for things”. The 9-item Interpersonal Relations domain (9 items) assesses the patient’s functioning in interpersonal relationships; an example item is “I have trouble getting along with friends and close acquaintances”. Finally, the 11-item Social Role domain assesses the patients functioning in social roles, such as work and school; an example item is “I feel stressed at work/school”. Possible scores on the OQ-45 range from 0 to 180, with higher scores reflecting more severe distress or dysfunction.

The Dutch OQ-45 has good psychometric properties. The internal consistency for the Total score ranges between 0.92 and 0.96 in university, community, and mixed samples. For the subscales, the international consistency is 0.90-0.95 for the Symptom Distress scale, 0.74-0.84 for the Interpersonal Relations subscale and 0.53-0.72 for the Social Role subscale (de Jong, Nugter, Lambert & Burlingame, 2008). In the present study, Cronbach’s alpha was 0.95.

Assessment for Signaling Clients (ASC; White et al., 2015). The Assessment for Signaling Clients (ASC) is a 40-item self-report instrument used to measure the therapeutic alliance, motivation for treatment, social support and stressful life events. In the present study, the ASC was administered to those patients in the ETR feedback condition who were indicated as “off-track” and most likely to deteriorate if they followed their predicted therapy progress. Feedback on patients’ ASC scores and their specific problematic domains was forwarded to therapists, in addition to practical tips for improving these therapeutic domains. Because the present study did not focus on “off track” patients, information on the ASC was not analyzed or reported here.

Therapist Measures

Use of Feedback (UOF). A self-constructed questionnaire (de Jong, in revision) consisting of three items assessed what the therapist has done with the available feedback; a) to what extend the feedback was in concurrence with what the therapist expected, b) if the therapist discussed the feedback with the patient, and c)
what the therapist has done with the treatment as a result of the feedback. When therapists answer the questionnaire, it is assumed that they have previously viewed the client’s feedback. Therefore, the frequency of UOF completion is considered indicative of the questionnaire viewing rate, and from now on will be mentioned as feedback viewing rate. The feedback viewing rate was used in this study to calculate the Implementation Index (see next) according to Bickman’s model.

*Implementation Index (II).* One of the main goals in the current study was to investigate whether the dose of feedback can moderate the effectiveness of treatment. In order to examine this, an Implementation Index was calculated, based on Bickman’s (2015) model. The Implementation Index consists of two dimensions: (1) questionnaire completion rate of the patients and (2) viewing rate of the feedback by the therapists. If the questionnaires are not completed, practically feedback cannot be created, and if the questionnaires are completed but feedback is not viewed by the therapist, then the intervention being evaluated (i.e. implementation of feedback) is not actually delivered. In order to calculate the Implementation Index, the OQ-45 questionnaire completion rates are averaged over all sessions, and the therapist’s feedback viewing rates are averaged for each patient. The two averaged scores are then multiplied together and divided by 100, thus the implementation index can range from 0 to 100. Thus, higher scores indicate increased implementation of feedback.

*Feedback intervention*

In the two feedback conditions the therapists were automatically provided with feedback on their patients’ progress. In the ROM condition, therapists received feedback on their patients’ progress in charts and tables, and the progress of the patient is provided by e-mail at session 1, 3, 5, 10 and 15, but can also be viewed by the therapist at all times, by logging on to the feedback system. In the ETR condition, the actual treatment course of the patient, based on their OQ scores, is compared with the predicted treatment course, which is calculated by a formula. The progress of the patient is provided to the therapist by e-mail when the patient is not progressing well, but it can also be viewed by the therapist at all times, by logging on to the feedback system. In the ETR condition, therapists were also provided with the patients’ scores on the ASC, in combination with a set of Microsoft Word documents, the Clinical
Support Tools, which consisted of practical tips on improving therapeutic alliance, motivation and social support.

**Procedure**

Following the standard intake assessment, eligible participants completed written informed consent procedures and were randomly assigned to one of the three research conditions. As the randomization took place at the patient’s level, therapists could be in any of the three conditions.

As soon as treatment started, patients were asked to complete the OQ-45 before each session, for a maximum of 15 sessions. In addition, the OQ-45 was administered 3 and 6 months after treatment (or study) termination, however the follow-up results of the patients are not reported here. In the ETR feedback condition, patients were asked to fill out the ASC when they were going “off track”. In the ROM feedback condition and control condition, patients were asked to fill out the ASC when they would signal, but the therapist would not receive any feedback on how the patient responded. The UOF was administered to therapists after sessions 5, 10 and 15 or at the end of treatment in case treatment lasted less than 15 sessions.

**Statistical Analysis**

Data was analyzed using the Statistical Package for the Social Sciences (IBM SPSS). Descriptive statistics were computed for the pre- and post-treatment OQ-45 scores (mean scores and standard deviations), and for the available demographic characteristics of the patients. Due to methodological reasons only the gender variable could be used in the present analyses. Information about participants’ characteristics such as patient’s age, previous treatment, or patient’s diagnosis is reported elsewhere (de Jong, in revision). Internal consistency of the outcome scale was calculated with the use of Cronbach’s alpha.

In order to compare treatment outcome for patients in the three study conditions, we conducted a one-way analysis of variance (ANOVA), including post-hoc tests with Bonferroni correction. In order to investigate whether the dose of
feedback has a predictive/ moderating effect on treatment outcome, we conducted a series of step-wise regression analyses.

**RESULTS**

Before testing the effectiveness of the feedback interventions, preliminary analyses were completed to test for baseline group differences. A one-way analysis of variance (ANOVA) was conducted to test for statistically significant differences in the mean OQ-45 scores at baseline for the three treatment groups. The mean pretreatment OQ score was 76.19 (SD=22.63) for the no-feedback/control group, 73.79 (SD=22.41) for the ROM feedback group, and 73.44 (SD=20.18) for the ETR feedback group. No statistically significant between-groups differences were found at pretreatment ($F=0.66$, $p > .5$). These results suggest that randomization was effective in creating groups that did not have dissimilar levels of initial distress or dysfunction.

Concerning the corresponding treatment variables, namely the number of OQ questionnaires completed by the patients during treatment, the number of therapy sessions, and the number of times that the therapists viewed the feedback given by the patients, there were no significant differences among the three groups (Table 2). When we controlled for gender differences in the corresponding treatment variables, there were no differences found in regard to pre- and post-treatment outcome scores between men and women. However, female participants completed more questionnaires in comparison to male participants ($F = 4.33$, $p = .04$), and there was a trend for women towards more sessions compared to men ($F = 2.98$, $p = 0.08$).

In regard to the feedback viewing rate of the therapists, the two feedback groups were not significantly different ($p = .8$), with therapists in the ROM group viewing their patients’ feedback a mean of 1.57 (SD=1.16) times, and therapists in the ETR group viewing feedback a mean of 1.44 (SD=1.16) times during treatment. There were no significant differences regarding feedback viewing when controlling for patients’ gender.
When we examined the mean pre- and post-treatment outcome scores on the OQ, all patients demonstrated lower scores at post-test, indicating that their functioning had improved over the course of therapy. However, between the three groups there were no significant differences at post-treatment in regard to treatment outcome (Table 3).

Next, in order to test our third hypothesis, we examined the average OQ completion rates by the patients and the average feedback viewing rates by the therapists, in order to calculate the Implementation Index. Since both the completion of questionnaires by the patients and the viewing of the feedback by the therapists are the two necessary elements for the implementation of the intervention being examined, we examined the mean rates for both variables based on condition (Table 4). The three groups did not differ in terms of questionnaire completion, and
differences regarding feedback viewing were only significant between the experimental and the control condition. The latter was a reasonable finding, as therapists in the control condition were not supposed to receive any feedback and thus could not participate in the feedback viewing procedure.

Table 4. Means and Standard Deviations in Questionnaire completion and Feedback viewing rates by Treatment Group

<table>
<thead>
<tr>
<th></th>
<th>No-feedback (n=142)</th>
<th>ROM feedback (n=156)</th>
<th>ETR feedback (n=126)</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire completion</td>
<td>75.69 (32.04)</td>
<td>77.60 (30.30)</td>
<td>73.09 (31.79)</td>
<td>0.72</td>
<td>.49</td>
</tr>
<tr>
<td>Feedback viewing</td>
<td>.00 (.00)</td>
<td>50.85 (35.81)</td>
<td>47.35 (37.32)</td>
<td>130.01</td>
<td>.00</td>
</tr>
</tbody>
</table>

The implementation index was significantly correlated with the number of sessions and with the outcome score difference between pre- and post-treatment (Table 5). These results suggest that there is a positive relationship between the amount of feedback implemented in therapy and on one hand the amount of sessions completed, ad on the other hand the change in outcome scores from the first to the last session of the client.

The number of sessions also demonstrated a small but significant positive correlation with pre- and post-treatment outcome scores, which reasonably suggests that patients with higher baseline scores, namely more severe initial distress or dysfunction completed a larger amount of sessions until end of treatment. Highly positively correlated were also the pre- and post-treatment scores, indicating that patients who started therapy with higher baseline scores also ended treatment with (decreased but still) higher post-test scores than patients who started treatment with a lower baseline dysfunction. The negative correlation between post-treatment scores and pre-post score difference suggests that patients with higher OQ scores at the end of therapy had a considerably smaller outcome difference as a result of therapy.
Table 5. Pearson Correlations Between Study Variables (n=423)

<table>
<thead>
<tr>
<th></th>
<th>Implement. Index</th>
<th>Number of sessions</th>
<th>Pre-treat. OQ scores</th>
<th>Post-treat. OQ scores</th>
<th>OQ pre-post difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation Index</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of sessions</td>
<td>.29**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Pre-treatment OQ scores</td>
<td>.07</td>
<td>.28**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Post-treatment OQ scores</td>
<td>-.075</td>
<td>.23**</td>
<td>.71**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OQ pre-post difference</td>
<td>.17**</td>
<td>-.00</td>
<td>.15**</td>
<td>-.58**</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. **p < 0.01

Given that the implementation index was significantly correlated with the number of sessions completed and the pre-post difference in outcome scores, we next conducted regression analyses to examine whether there was an additional predictive effect between these variables. According to the prediction model, the implementation index was not predictive of the pre-post difference in outcome, whereas on the other hand both the pre- and post-treatment OQ scores could partially predict the difference in treatment outcome. As a result, the possible predictive effect of the implementation index that we had anticipated was not supported by our results. Overall, although the reasons for this inconsistency are not fully clear, there are various possibilities and conceptual suggestions could illuminate the present findings. These suggestions as well as the limitations of the present study are discussed below.

**DISCUSSION**

Given that previous research has demonstrated the additive effects of feedback to psychotherapy, we anticipated that patients in the two feedback groups would show better treatment outcomes at the end of treatment compared to patients who were in the
no-feedback/ control group. Contrary to our expectations, this hypothesis was not supported by the results of our study, as no statistically significant differences were found between the three groups in regard to treatment outcome. Despite a trend toward better outcomes at post-treatment for the routine outcome group, the feedback groups did not appear to be superior to the no-feedback/control group based on the OQ scores of the participants. Therapy length as expressed by the total number of sessions completed in treatment did not differ either between the three groups, contrary to previous literature suggesting that feedback decreased therapy length as it accelerated positive treatment change. The results of our study suggest that therapy was rather equally effective for patients in the three groups, and feedback as implemented by the therapists did not provide an additional benefit to treatment outcome or duration of therapy.

A possible explanation for the “non-superiority” of the two feedback groups compared to treatment as usual/ control, can possibly be found in the presence of several, uncontrolled confounding factors, including several characteristics of the patients and/or the therapists, that could potentially create “noise” and affect treatment outcome. These confounding variables include patients’ age, main diagnosis, prior treatment, use of medication or other supplementary treatment, comorbidity or duration of main complaint. As aforementioned, patients participating in the study were not excluded for receiving co-interventions (i.e. other type of psychological treatment or pharmacotherapy) as a supplement to the individual treatment. Therapist factors include characteristics such as age, gender, level of experience, etc. However, the aforementioned variables were not the focus of this preliminary study and their effect on the examined relationships could not be accounted for. Nevertheless, this possible confounding could result in a change in our effect estimates and future research should include these variables in the analyses and control for their effect on the relationship between treatment outcome and feedback use.

Our results were not consistent with the findings of previous research (Lambert, 2003; Hawkins et al., 2004; Bickman, 2008), suggesting that providing feedback to therapists about their patients’ progress strengthens treatment outcomes. Furthermore, we had hypothesized based on previous studies that the addition of a more complex feedback system, based on an individualized prediction model for patients’ treatment course, would have a cumulative effect on treatment outcome, which would be superior
to standard feedback as provided in routine outcome monitoring. However, our results did not support this hypothesis, as patients in the ETR feedback group did not show better outcome results compared to the standard ROM feedback or to the no-feedback control group. Some considerations may be raised about the possible effectiveness of the experimental manipulation being delivered and measured in this study, namely the use of feedback. As previously highlighted, the implementation of feedback entails and presupposes the active participation of both the patient and the therapist. According to Bickman (2016), “Both of these actions must happen to at least some reasonable degree for successful implementation to occur”. In other words patients need to actively fill in the outcome questionnaires and therapists need to actively view the feedback provided to them by their patients. Hence, one could assume that the use of feedback was either not effective in improving treatment outcome in the present study, or the way feedback was implemented in this study did not have a significant additive effect on treatment outcome. Nevertheless, further research is needed to illuminate the circumstances under which feedback to patients can result in better treatment benefits, in addition to qualitative information (i.e. ASC about problematic domains, UOF about therapists’ use of feedback) about the utility and perceived impact of feedback in therapy.

Considering that many studies in the past have been examining ways to improve feedback implementation by increasing the likelihood of therapists using the available information in therapy to reduce treatment failure (Slade et al., 2008), it would be rather useful to track therapists’ use of in-therapy feedback more closely, and use frequent reminders to increase the likelihood of therapists routinely viewing patients progress and creating interventions that prevent patients from deteriorating.

STUDY LIMITATIONS & FUTURE DIRECTIONS

In addition to the above considerations, several limitations of the study need to be considered when interpreting the results. First of all, although we employed Bickman’s model about the implementation index, for the present preliminary analyses, we did not make a distinction between ‘on-track’ and ‘off-track’ patients. Patients making expected progress as well as patients whose progress could have deviated from their expected change trajectory were included in the same sample, and we solely examined patients’ pre- and post-treatment outcome scores and difference,
as an indicator of treatment change. Probst et al. (2013) found no differences between feedback and no-feedback/control conditions before the timepoint at which patients first signaled as ‘off-track’ (i.e., therapists received a warning signal about the patient for the first time), however patient’s progress was influenced by this signal from that point onwards. Therefore, further analyses of our data could examine the slopes prior to and after that point/session at which patients started going off-track. At the same time, patients who had a ‘crisis’, i.e. a relapse / acute increase in symptomatology, were also included in the analyses, which could have resulted in data contamination and consequently further limited the generalizability of our results. Hence, as a future step, it would be interesting to look into the crisis-cases and examine the influence of feedback on the incidence of crisis and vice versa.

Moreover, due to the naturalistic nature of the study, patients could terminate treatment at any time, and considering that off-track patients could drop-out of treatment before they could experience some significant change, this could have resulted in attrition before the effect of feedback intervention was visible. At the same time, it is not clear based on our preliminary data if there were patients who discontinued their participation in the study or dropped out of treatment. Therefore, it would be very insightful in future steps if we could follow both participants and non-participants and acquire qualitative data regarding their treatment progress, reasons for therapy (dis)continuation and feedback on how their overall experience was. Moreover, a differential loss of participants to follow-up or due to early termination of treatment/drop-out would increase the risk for selection bias, which would be important to know.

Another consideration about the present study is that we solely examined the effect on treatment outcome based on the total scores of the OQ-45, and did not further look into patients’ scores at the individual subscales/domains of the measure. The OQ-45 is an instrument with very good concurrent validity coefficients with a variety of self-report measures (Vermersch et al., 2000) and has been demonstrated to change over time as a result of psychotherapy sessions (Kadera et al., 1996). However, it would be an interesting future step to look into the three different domains of function depicted in the OQ-45 subscales (Symptom Distress, Interpersonal Relations and Social Role) and examine change over time.
Last but not least, some information could not be utilized in the present study, as participants seemed to have provided incomplete or invalid data (e.g. incomplete or duplicate questionnaires). Irregularities in the data were corrected before proceeding to the analyses, however valuable information could potentially have been lost and affected the present findings. In order to reduce human error and ensure both patients’ and therapists’ adherence to questionnaire completion, better planning and briefing of the participants at the beginning of the study would be necessary, in addition to using reminders about filling in the questionnaires.

Concluding, it is certain that further research is needed about the effects of feedback and the specific functions of routine outcome monitoring. Given that feedback interventions are considered a highly promising approach to improve clinical outcomes (de Jong et al., 2017; Shimokawa et al., 2010), and that research findings about feedback effects vary from large to moderate or even small in some studies (Sapyta, 2005; Shimokawa, 2010), further research should explore potential factors related to patient, therapist, or therapy characteristics in order to shed more light into this increasingly important area of clinical and research interest.
REFERENCES


