



Universiteit
Leiden
The Netherlands

A chance for change : building an outcome monitoring feedback system for outpatient mental health care

Jong, K. de

Citation

Jong, K. de. (2012, April 17). *A chance for change : building an outcome monitoring feedback system for outpatient mental health care*. Retrieved from <https://hdl.handle.net/1887/18691>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/18691>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/18691> holds various files of this Leiden University dissertation.

Author: Jong, Kim de

Title: A chance for change : building an outcome monitoring feedback system for outpatient mental health care

Date: 2012-04-17

Understanding the differential
impact of outcome monitoring:
therapist variables that moderate
feedback effects in a randomized
clinical trial

Chapter 5

de Jong, K., van Sluis, P., Nugter, M.A.,
Heiser, W.J. & Spinhoven, P. (in press)

Psychotherapy Research.

Abstract

Providing outcome monitoring feedback to therapists seems to be a promising approach to improve outcomes in clinical practice. This study aims to examine the effect of feedback and investigate whether it is moderated by therapist characteristics. Patients ($n = 413$) were randomly assigned to either a feedback or a no feedback control condition. There was no significant effect of feedback in the full sample, but feedback was effective for not on track cases for therapists that used the feedback. Internal feedback propensity, self-efficacy, and commitment to use the feedback moderated the effects of feedback. The results demonstrate that feedback is not effective under all circumstances and therapist factors are important when implementing feedback into clinical practice.

Introduction

A large body of research, performed over 40 years, has demonstrated that psychotherapy can effectively improve functioning in patients (Lambert & Ogles, 2004). In randomized controlled clinical trials (RCTs) an average of 67% of patients is statistically reliably improved at the end of treatment (Hansen, Lambert, & Forman, 2002). In clinical practice, the success rates are much lower: only 35% of the patients were improved and the effect sizes of improvement were less than half the effect sizes of RCTs (Barkham, et al., 2008; Hansen & Lambert, 2003; Weisz, Donenberg, Han, & Weiss, 1995). These differences in outcomes may in part be due to selection criteria used in RCTs. However, Blais et al. (2011) found that in clinical practice the improvement rate is 57% when patients are selected that would qualify for inclusion in RCTs, which is still lower than rates found in RCTs.

Bickman (2008) considers feedback interventions a promising approach to improve clinical practice. Kluger and DeNisi (1996) define feedback interventions as actions taken by external agents to provide information regarding some aspect of one's task performance." In psychotherapy research, a common example is the monitoring of patients' progress during treatment and providing feedback to therapists on that progress. According to Bickman, clinicians need to have more systematic and reliable information about the status of their patients, in order to adjust their treatment if necessary, thus improving outcomes. In a recent review article, Carlier et al. (2010) concluded that feedback appears to have a positive impact on diagnosis and communication between patient and therapist, but effects on outcome were less clear. Meta-analyses show effects of feedback on outcome in the range of very small to very large (Knaup, Koesters, Schoefer, Becker, & Puschner, 2009; Sapyta, 2004 in Sapyta, Riemer, & Bickman, 2005; Shimokawa, Lambert, & Smart, 2010). Feedback appears to be most effective for patients that are not progressing well in therapy, the so-called not on track (NOT) cases (Lambert, et al., 2003; Sapyta, 2004 in Sapyta, et al., 2005). Carlier et al. found that feedback did not have a positive effect in 16 of the 45 (36%) mental health trials they included in their sample and Knaup et al. showed that feedback even had a negative effect in three of the twelve studies they included in their analyses.

The largest effects of feedback in mental health care have been found by the research group of Michael Lambert. They have performed five controlled studies in which therapists received feedback about a patient's improvement through the use of progress charts and warning signals about NOT cases. Results showed that NOT patients in the feedback condition had significantly more improvement than in the no-feedback control condition. The effect sizes of the various feedback conditions compared to the control conditions ranged from 0.16 to 0.70 in the full sample (Shimokawa, et al., 2010). Feedback did not have a significant effect in the on track (OT) cases (Lambert, 2007).

The feedback that the Lambert group provides to therapists is very specific. A patient's change is compared to an expected treatment course based on a statistical model and the therapist gets a warning signal when the patient deviates too much from the expected course. Most outcome monitoring systems are not as advanced. Worldwide, there are many large initiatives (e.g. Burgess, Pirkis, & Coombs, 2006; Evans, et al., 2002; Howard, Moras, Brill, Martinovich, & Lutz, 1996; Kraus, Seligman, & Jordan, 2005; Miller, Duncan, Sorrell, & Brown, 2005; Wing, et al., 1998). Some systems have developed expected treatment recovery curves, but most do not include them in feedback on the individual patient. The existing systems vary greatly in frequency of assessment, content of the feedback and the way in which feedback is provided (Trauer, 2010). Often, feedback is not provided on a session-by-session basis, but at treatment evaluations, for instance every three months. Most feedback systems do not have signals for patients that are not progressing well in therapy. It is often assumed that all types of feedback will be effective in improving outcomes, but in fact, not much controlled research has been done on the subject (Marshall, Haywood, & Fitzpatrick, 2006).

Although it seems that feedback has potential to enhance outcomes in clinical practice, there are still many unanswered questions about how feedback works. In order to explain why outcome monitoring feedback leads to improvement in some cases and not in others, more insight is needed in the underlying processes of feedback. Characteristics of the therapists and the way in which they use feedback may play a central role in the effectiveness of feedback. After all, if therapists do not use feedback constructively, it is unlikely that it will improve outcomes. There is not much empirical knowledge on the effects that recipient characteristics have on the effectiveness of feedback (Kluger & DeNisi, 1996).

Riemer and Bickman (2011) propose the contextualized feedback intervention theory (CFIT) to explain how feedback is interpreted and used in clinical practice. CFIT focuses on the way that feedback gets attention and is accepted by therapists. When a person receives feedback a comparison is made between the content of the feedback and a goal. So if a therapist receives progress feedback, a comparison is made between the goal (recovery) and the feedback (current health status and progress so far). This comparison creates a positive or negative evaluation of the therapist's performance relative to the goal. When a discrepancy is noted, people are motivated to reduce it (Kluger and DeNisi, 1996).

This implies that behaviour change as a result of feedback will only occur if therapists attend the feedback and accept the feedback as valid (Riemer & Bickman, 2011). Feedback is more likely to be accepted if it comes from a source that has credibility and has personal relevance to the receiver (Claiborn & Goodyear, 2005). This concept is referred to as perceived validity. Another factor that seems important in acceptance is feedback

orientation (Herold & Fedor, 2003; Herold, Parsons, & Rensvold, 1996). External feedback propensity reflects the preference for externally mediated feedback as well as greater faith in such information than in what one can self-generate, whereas internal feedback propensity reflects preference for internally generated feedback as well as the tendency to reconcile differences between internal and external feedback in the direction of internally generated information. An external feedback propensity is associated with more feedback seeking behavior and better performance on novel tasks (Herold & Fedor, 2003).

Self-efficacy is another recipient characteristic that influences the feedback process. It refers to a person's beliefs concerning his or her ability to successfully perform a given task or behavior (Bandura, 1977). In case of negative feedback, people with high self-efficacy are motivated to increase their effort to reach the goal, whereas people with low self-efficacy tend to lower the goal (Kluger and DeNisi, 1996). People who have high self-efficacy also tend to consider negative feedback as more desirable than positive feedback (Claiborn and Goodyear, 2005).

Therapists' commitment to use the feedback in therapy might also be an important factor. Australian research showed that 44% of therapists thought outcome monitoring was a waste of time (Aoun, Pennebaker, & Janca, 2002) and two-thirds of the therapists were not willing to use the monitoring feedback, even if it would lead to demonstrably better outcomes (Walter, Cleary, & Rey, 1998). Riemer and Bickman (2011) state that therapists will be more committed to use the feedback if they link it to higher-level personal goals, such as being a good therapist. An a priori commitment to use feedback is expected to be highly related to actual use of the feedback.

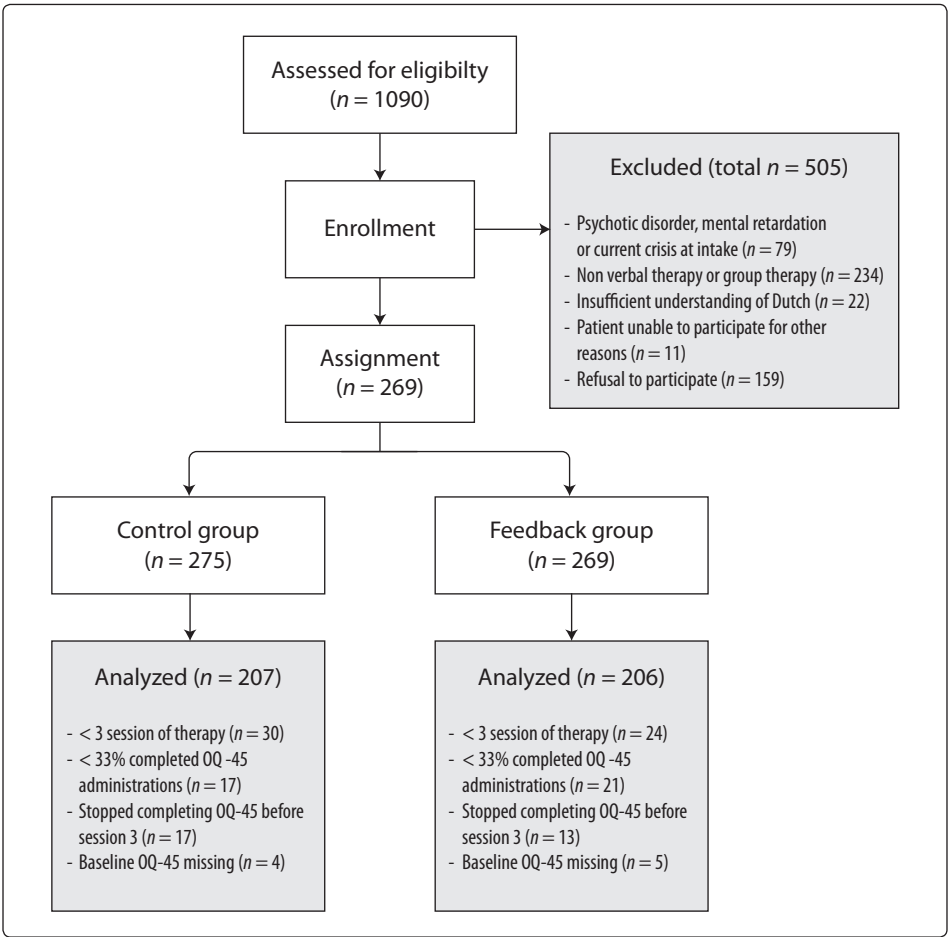
In summary, outcome monitoring has potential to improve outcomes, especially feedback with expected recovery curves and alarms for 'not on track' cases tending to result in positive effects. Most outcome monitoring systems used in clinical practice do not have these features and not much is known about the effectiveness of these systems. The effectiveness of feedback may also be related to therapist characteristics. If therapists do not accept the feedback and are not inclined to use it, feedback is not likely to be effective. In this study, we aim to research the efficacy of 'simple' (no warning signals or expected recovery curves) feedback in clinical practice compared to no feedback. Patients will be randomly assigned to a no-feedback control group or the feedback condition. We expect that patients in the feedback condition will have faster progress, compared to the no feedback control group. Although no alarms are used, therapists may be able to identify NOT cases themselves. Therefore, NOT patients will be identified post-hoc, based on reliable deterioration during the course of treatment, and it is expected that feedback will be most effective for this group. A secondary aim is to investigate whether therapist characteristics moderate the effect of feedback.

Method

Patients

During the inclusion period 1090 outpatients were screened for participation in the study in the three participating treatment departments. The treatment departments were part of two medium sized mental health care institutions in the Netherlands and typically treated a wide range of psychiatric disorders, including mood, anxiety, adjustment and personality disorders with an outpatient population. Exclusion criteria were: psychotic disorder, mental retardation, a current crisis at the time of referral, non-verbal treatment (e.g., internet therapy, pharmacological therapy, art therapy), group therapy as main treatment, re-referral within the same treatment centre within

Figure 1 Flowchart of participants



six months, and an insufficient level of understanding of Dutch. Of the remaining 703 patients, 159 declined to participate in the study.

In total, 544 patients were randomly assigned to the feedback group or control group. The first progress feedback was provided immediately before session 3; therefore, patients who had less than three sessions of therapy or stopped completing questionnaires before session 3 were excluded from analysis. Patients that had missing baseline measurements or completed less than a third of the measurements were also excluded from analysis. The flow of participants through the study is presented

Table 1 Characteristics of the patients in the feedback and control group

	Sample entered study (<i>n</i> = 544)				Sample in analysis (<i>n</i> = 413)			
	Not in analysis		In analysis		Control		Feedback	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Sex								
- Female	131	75 (57%)	413	252 (61%)	207	124 (60%)	206	128 (62%)
Age	131	M = 37.0 SD = 12.3	413	M = 36.8 SD = 11.9	207	M = 36.9 SD = 11.8	206	M = 36.7 SD = 12.1
Marital status								
- Single	128	58 (45%)	410	178 (43%)	206	94 (46%)	204	84 (41%)*
- Living together		10 (8%)		32 (8%)		14 (7%)		18 (9%)
- Married		39 (31%)		136 (33%)		62 (30%)		74 (36%)
- Divorced		21 (16%)		56 (14%)		35 (17%)		21 (10%)
- Widowed		0 (0%)		8 (2%)		1 (0.4%)		7 (3%)
Education								
- Low	106	39 (37%)	380	94 (25%)*	193	48 (25%)	187	46 (25%)
- Medium		57 (54%)		226 (60%)		113 (59%)		113 (60%)
- High		10 (9%)		60 (16%)		32 (17%)		28 (15%)
Main DSM-IV disorder								
- Mood	131	35 (27%)	413	96 (23%)	207	47 (23%)	206	49 (24%)
- Anxiety		17 (13%)		79 (19%)		34 (16%)		45 (22%)
- Adjustment		32 (24%)		92 (22%)		50 (24%)		42 (20%)
- Personality		9 (7%)		31 (8%)		16 (8%)		15 (7%)
- Eating		0 (0%)		10 (2%)		8 (4%)		7 (3%)
- Usually first diagnosed in childhood		7 (5%)		13 (3%)		6 (3%)		7 (3%)
- Substance related		2 (2%)		12 (3%)		7 (3%)		5 (2%)
- Somatoform disorder		2 (2%)		13 (2%)		5 (2%)		8 (4%)
- Impulse control		2 (2%)		10 (2%)		7 (3%)		3 (2%)
- Other		24 (18%)		46 (11%)		24 (12%)		22 (11%)
- No Axis I or II disorder		1 (1%)		6 (2%)		3 (1%)		1 (2%)
Comorbidity								
- Multiple Axis I disorders	131	52 (40%)	413	154 (37%)	207	76 (37%)	206	78 (38%)
- Comorbid Axis I and II disorders		26 (20%)		97 (24%)		50 (24%)		47 (23%)
QQ-45 intake score	127	M=72.9 SD= 23.4	413	M= 76.7 SD= 22.1	207	M=76.7 SD= 22.0	206	M= 76.7 SD= 22.3
Prior treatment								
- Yes	127	75 (59%)	410	237 (58%)	206	124 (60%)	204	113 (55%)

Note

* $p < .05$

in Figure 1. The 413 patients that were included in the analysis included 252 females (62%), aged 18-64 years ($M = 36.8$; $SD = 11.9$). Patient characteristics are reported in Table 1.

Therapists

There were 57 therapists who participated in this study, 21 males (37%) and 36 females (63%), aged from 26 to 60 years with a mean age of 45.3 years ($SD = 9.7$). Therapists were psychologists (49%), psychiatric nurses (39%), social workers (7%) or other mental health care professionals (5%). The therapists had 0 to 35 years of experience after getting licensed, with a Mean of 13.8 years ($SD = 11.3$). Therapists had between 1 and 23 clients in the study ($M = 7.3$; $SD = 5.6$). Therapies provided included cognitive behavioral therapy, interpersonal therapy, brief solution focused therapy and counseling. Most therapies were integrative and did not represent a single therapy orientation.

Instruments and manipulation

Outcome Questionnaire-45 item version (OQ-45)

The Outcome Questionnaire-45 (OQ-45) was used to measure patient progress during treatment. The OQ-45 (Lambert, et al., 2004) is a self-report instrument and has 45 items, 9 of which are reversed, asking how the respondent has felt over the last week on a 5 point rating scale, ranging from 0 (never) to 4 (almost always). The OQ-45 consists of three subscales: Symptom Distress, Interpersonal Relations, and Social Role. The Symptom Distress domain consists of 25 items relating to psychological symptoms that are common in highly prevalent mental disorders. The Interpersonal Relations domain consists of 9 items that assess functioning in interpersonal relationships, and the Social Role domain consists of 11 items that assess functioning in social roles, such as work and school. The cutoff score for normal functioning is 55 for the Dutch OQ-45 and the reliable change index is 14. The internal consistency for the Dutch version of the OQ-45 is between 0.92 and 0.96 for the Total Score in university, community, patients and community and patients combined samples. For the subscales, the internal 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, 2009).

Demographic questionnaire

The demographic questionnaire is a 19-item self-constructed questionnaire that assesses the demographic characteristics of the patient. It asks for the patient's date

of birth, gender, postal area code, nationality, country of birth, country of birth of the patient's parents, marital status, living and working situation, educational level, prior treatment, pretreatment use of medication, the main complaint, and the duration of the main complaint.

Feedback User Questionnaire

This questionnaire consisted of the Internal and External Feedback Propensity Scales and an adaptation of the CFIT User Survey.

The *Internal and External Feedback Propensity Scales* (IEFPS; Herold, et al., 1996) are used to measure feedback propensity. The instrument consists of two subscales that measure internal and external feedback propensity. Each subscale consists of six items that are answered on a five-point rating scale that varies from strongly disagree to strongly agree. An item from the External Feedback Propensity Scale is 'It is very important to me to know what people think of my work.' A sample item from the Internal Feedback Propensity Scale is 'How other people view my work is not as important as how I view my own work.' The reliability of the IEFPS was 0.71 for the external feedback propensity scale and 0.73 for the internal feedback propensity scale (Herold, Parsons, & Fedor, 1997). In our sample, the internal feedback propensity scale had a Cronbach's α of 0.71 and the external feedback propensity scale had an α of 0.62.

An adaptation of the *CFIT User Survey*, designed by the Center for Evaluation and Program Improvement of Vanderbilt University, was used to measure commitment to use the feedback, self-efficacy and perceived validity of the feedback. The items are scored on various five-point rating scales. The commitment to use the feedback was measured with a scale based on the *Goal Commitment Scale* (Hollenbeck & Klein, 1987) that consists of seven items. A sample item is 'It is hard to take the idea of using these measures in my clinical practice seriously.' The self-efficacy scale consists of eight items. A sample item from that scale is 'To what extent do you feel confident in your ability to know what to do if a client is not progressing in treatment.' Perceived validity of the feedback was measured by a six-item scale. A sample item is: 'I think that feedback based on the OQ-45 will be helpful for my counseling.' The internal consistency in the current sample was 0.90 for the commitment scale, 0.88 for the perceived validity scale, and 0.82 for the self-efficacy scale.

Use of feedback

In the original study design the use of feedback was asked per patient, but due to problems in the software, this questionnaire was not administered. Therefore, the use of feedback by the therapist was assessed post hoc by asking by e-mail if the therapists had used the feedback with their patients (yes/no) and in what way (open question). Therapists that had used the feedback usually did so in multiple ways, including

discussing the feedback with patients, giving homework assignments, and using the feedback to end the therapy when sufficient progress was made.

Feedback intervention

In the feedback condition, the therapist received e-mails that contain a progress report after sessions 1, 3, 5, and subsequently every fifth session. The patient's progress on the OQ-45 Total Score was shown in a graph. A table showed the patient's baseline score, the last available measurement, the change in scores on the OQ-45, and the clinical and reliable change status (see Appendix B). Patients were classified as deteriorated if their OQ-45 score increased 14 points or more compared to baseline on the OQ-45 Total Score and classified as reliably improved if their Total Score had decreased 14 points or more. If patients improved reliably and crossed the cut-off point for normal functioning (55), they were classified as clinically significantly changed. Patients that did not meet these criteria were considered unchanged. Positive changes were shown in green, negative changes in red. A second graph and table displayed the subscale scores. The critical items on the OQ-45 that alert the therapist to suicidal thoughts, aggression and drugs and alcohol use, were presented if patients answered them with a score of 1 (seldom) or higher. Prior to the study, all therapists were given training on how to interpret the feedback, but were given no specific guidelines to identify 'not on track' patients, consistent with the concept of simple feedback. They also received an instruction card that explained all elements of the feedback report.

Procedure

Patients were screened for eligibility after intake and contacted by phone if they did not meet the exclusion criteria. If patients agreed to participate in the study, they signed an informed consent form and received explanation using the on-site test computer. Patients completed the OQ-45 on the computer prior to each of the first five sessions of therapy, and subsequently every fifth session for a maximum period of one year. At the first session, patients also completed the demographic questionnaire. If patients were assigned to therapies that were excluded from the study, data collection was stopped and measurements up until that point were used in analysis. Therapists completed the Feedback Questionnaire prior to the study. The use of feedback questionnaire was e-mailed to the therapists after completion of the study.

Table 2 Therapist variables (*n* = 57)

	<i>% missing</i>	<i>Mean</i>	<i>SD</i>
Self-efficacy therapist role	21%	27.2	2.6
Internal feedback propensity	23%	19.6	3.4
External feedback propensity	23%	19.7	2.4
Perceived validity	18%	21.2	3.4
Commitment to use feedback	16%	23.9	3.9

Analysis

Missing data

Missing data on the therapist level were imputed using the Multiple Imputation procedure in PASW Statistics 18.0 (SPSS, 2009). The multiple imputation procedure is based on the Multiple Imputation by Chained Equations (MICE; van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, 2006) algorithm. Since multiple imputation is not supported for three-level models yet, single imputation of the missing values was selected by setting the imputation number as one. Only missing values on the therapist variables were imputed. Table 2 reports the percentage of missing data and the mean score and standard deviation in the original data.

Definition of not on track

Patients that were NOT were identified post hoc. Patients were considered NOT if they deteriorated, defined by an increase in the Total Score at least as large as the reliable change index (14 points) compared to the baseline measurement at any point in their treatment. This criterion was chosen since it was mentioned in the feedback report if the patient deteriorated. A total of 67 patients (16%) were NOT according to this definition.

Main hypotheses

Baseline differences in demographic and clinical characteristics between the treatment conditions were tested with chi-square tests and independent sample t-tests using PASW Statistics 18 (SPSS, 2009). The main hypotheses were tested by two three-level multilevel models, using the PROC MIXED procedure in SAS (SAS Institute Inc, SAS 9.2. Cary, NC, USA, 2008). As time variable the 10log of the session number was used to allow for a linear model (also see Lutz, Martinovich, & Howard, 1999). Maximum likelihood estimation was used to estimate the model parameters, using an unstructured variance structure. A random intercept, random slope model (on both patient and therapist level) was used to test the main hypothesis on the effect

of feedback. First, an unconditional growth model was postulated, and then the main effect of feedback was added to the model, followed by the interaction with being NOT and use of feedback by the therapist. Redundant factors were eliminated from the model, in order to obtain a parsimonious model. To test for therapist effects, a model with a fixed slope on level two and a random slope at level 3 was used (with random intercepts). A backwards procedure was applied, starting with a full model including all relevant level 2 and 3 predictor variables and their interactions and eliminating non-significant factors (using the Wald test for fixed effects) one by one until a parsimonious model was reached, that was not significantly worse than the full model (compared with the deviance test).

To predict which therapist characteristics would predict use of feedback, a logistic regression analysis was performed in PASW statistics (SPSS, 2009), using a backwards procedure.

Results

Baseline differences between groups

Baseline differences on gender, age, marital status, education, DSM-IV disorder, OQ-45 intake score and prior treatment between the patients that were included and excluded in analysis were tested. The groups did not differ on most variables (see Table 1), except for educational level. Patients that were excluded from analysis were more likely to have a low education than patients that were included, $\chi^2(6) = 11.4, p = .039$. For patients that were in analysis, baseline differences between the control group and feedback group were tested. The two groups only differed on marital status: patients in the control group were more likely to be widowed and less likely to be divorced than the feedback group, $\chi^2(2) = 7.2, p = .027$.

Effect of feedback

There was no significant effect of feedback on the rate of change (see Model A, Table 3). The interaction between feedback and being NOT was also not significant. Adding the interaction with use of the feedback to the model revealed that for therapists that used the feedback (46%, representing 57% of the patients), there was a significant positive effect of feedback in NOT cases (see Model B, Table 3), although the effect was not large enough to counterbalance the negative change trajectory that NOT patients typically have.

Table 3 3-level models on the effect of feedback and moderating therapists factors

		<u>Model A</u>	<u>Model B</u>	<u>Model C</u>
Parameter		<i>Estimate (SE)</i>	<i>Estimate (SE)</i>	<i>Estimate (SE)</i>
<u>Fixed effects</u>				
Initial status, β_{0j}	Intercept	77.47 *** (1.11)	77.49 *** (1.11)	77.30 *** (1.11)
Rate of change, β_{1j}	Time	-16.83 *** (1.63)	-17.57 *** (1.64)	-5.11 (9.69)
	Feedback	0.59 (2.19)		
	Not on Track	27.37 *** (3.43)	27.84 *** (2.92)	19.55 *** (1.80)
	Feedback * NOT	-5.87 (4.83)		
	Use *Feedback		3.06 (2.30)	
	Use * Feedback * NOT		-10.77 * (5.15)	
	Internal feedback propensity			0.58 * (0.27)
	Commitment to use feedback			-0.84 ** (0.29)
	Self-efficacy* Feedback			-1.24 *** (0.31)
	Commitment * Feedback			1.33 *** (0.34)
<u>Variance components</u>				
Level 1	Within-person	92.44	92.50	112.91
Level 2	Intercept	440.74	440.16	417.69
	Slope	175.78	172.67	
	Covariance	-56.27	-56.87	
Level 3	Intercept	0.16	0.11	2.87
	Slope	5.02	5.47	21.51
	Covariance	-6.46	-6.48	-16.19

Note: Time is modeled as the 10log of the session number. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

NOT = Not On Track

Negative values in the fixed part of the models correspond with a faster decrease of dysfunctioning over time.

Use, NOT and Feedback were coded as dummy variables (0 = no, 1 = yes).

Table 4 Logistic regression analysis predicting use of feedback

	<i>B</i>	<i>SE</i>	Odds ratio	95% confidence interval	
				Lower bound	Upper bound
Constant	-5.23	2.21	0.01		
Commitment to use feedback	0.25*	0.10	1.28	1.06	1.54
Female	-6.61**	2.48	4.01	1.10	14.45

Note: $R^2 = .18$ (Cox & Snell), $.23$ (Nagelkerke). Model $\chi^2(2) = 10.78, p = .005$. * $p < .05$, ** $p < .01$

Therapist characteristics

The effects of therapist characteristics on the rate of change are presented in Model C, Table 3. The correlation between the therapist characteristics was below $r = 0.45$ for all pairs, except for perceived validity and commitment to use feedback ($r = 0.70$). Having an internal feedback propensity had a negative effect on the rate of change, regardless of whether the therapist received feedback. A higher commitment to use the feedback had a general positive effect on the rate of change, but there was also a significant interaction between commitment and feedback in a negative direction, indicating that when therapists actually received feedback, having a higher commitment was predictive of a slower rate of change in their patients. Finally, there was a positive effect of self-efficacy in the feedback condition. Patients of therapists with higher self-efficacy expectations who received feedback had a higher rate of change than patients of therapists with lower self-efficacy expectations that did not receive feedback. Against our expectations, there was no significant effect of external feedback propensity and perceived validity.

Predicting use of feedback

Since use of feedback by the therapist significantly interacted with the effect of feedback in NOT cases, we were interested in which variables predicted use of the feedback by the therapists. Table 4 shows that a higher commitment to use the feedback increased the odds that therapists would use the feedback. Being a woman also increased the odds of using the feedback: female therapists were four times more likely to use the feedback than male therapists. No significant effects were found for the type of therapist, years of experience of the therapist, self-efficacy, internal and external feedback propensity, and perceived validity.

Discussion

This study aimed to assess the effect of monitoring outcomes and providing feedback to therapists on the rate of change in patients. Contrary to our expectations, for the full sample no beneficial effect of feedback was found and there was no significant interaction between feedback and patients being NOT. However, in NOT cases a positive significant effect was found when therapists indicated that they used the feedback. Therapist variables moderated the effectiveness of feedback. Therapists with a high internal feedback propensity, who are more likely to trust their own opinion than feedback from external sources, had patients with a slower rate of change than therapist with a low internal feedback propensity, whereas therapists who were more committed to use the feedback at the beginning of the study had patients who progressed faster. These two results suggest that therapists with an open attitude towards getting feedback reach faster progress with their patients. Strangely though, when therapists with a high commitment to use the feedback actually received feedback, this slowed down the rate of change in their patients. There also was a positive effect of self-efficacy. Patients in the feedback condition whose therapist had higher self-efficacy progressed quicker in therapy than patients whose therapist had lower self-efficacy or patients whose therapist did not receive feedback. No effect was found for external feedback propensity and perceived validity. Therapists were more likely to use the feedback if they were more committed to use the feedback at the start of the study and if they were female.

Our results demonstrate that feedback may not be effective under all circumstances for all therapists. This is in line with a recent study by Lambert's group, in which they used complex feedback in a hospital-based outpatient clinic and found much lower effects of feedback than in previous studies. Further analysis showed that feedback was only effective for half of the therapists (Simon, Harris, & Lambert, 2011). The therapists that were participating in that study had very heavy caseloads and seemed demoralized by organizational changes. Riemer and Bickman (2011) stress that organizational factors, such as a high administrative workload, can become barriers for therapists to use feedback. In a recent survey that included many of the therapists participating in the current study, therapists indicated lack of time and other tasks that were competing for their attention as important barriers to use the feedback (De Jong, in press). Londen, Smither and Adsit (1997, in Riemer & Bickman, 2011) state that if there is no accountability for using feedback, it will have little impact. Accountability should be handled with care though, as it can also provoke defensive reactions in therapists (Riemer & Bickman, 2011). Although the managers of the participating departments were actively involved in the study, it was still complicated to hold therapists accountable for using the feedback within the context of a research project.

Managers were not allowed to view therapists' progress curves, in order to prevent defensive reactions.

Some of the choices we made in designing the study could have influenced results. The chosen frequency of measurements and feedback reports was not on a session-by-session basis, as Lambert does, which may reduce the chance to signal a patient being not on track and as a result might reduce the effect of the feedback. We encountered a relatively low rate of signal cases (16%) in this study. Another issue is that patients completed questionnaires up until one year or until they were referred to treatments that were excluded from this study. As a result, a portion of the patients in our dataset stopped with the study before the end of treatment, which may have reduced the feedback effect. However, one would still expect an effect of feedback for the sessions on which feedback was provided. A third factor that may have influenced results in a negative way was several problems we occasionally encountered with the feedback software. One of the problems was that the questionnaires we had planned to administer on use of feedback did not work, which forced us to measure therapists' use of feedback post hoc. This may have several disadvantages, as therapists may not always remember accurately if they used the feedback or not. Also, demand characteristic may play a role. Therapists are aware that they should have used the feedback and may be less likely to report that they did not. However, considering that half of the therapists indicated that they had not used the feedback, we believe that the effect of sociably desirable answers was limited in our study. A final issue is that the sample may be selective to a certain degree. It seemed that patients not included in analysis were more likely to have low education, which is consistent with results from dropout studies (Clarkin & Levy, 2004). Therefore, our results may be less representative for lower educated patients.

What implications do these results have for clinical practice? It seems important to realize that not all types of feedback may be equally effective. People often refer to Lambert to justify implementing feedback, but take out elements of his feedback system that may be particularly effective. Warning signals might be effective in getting therapists' attention to look at the feedback and the statistical model underlying the expected recovery curves may cause the therapists to perceive the feedback as more valid. Another implication is that therapists' commitment to use the feedback seems to influence the feedback's effectiveness. It is especially important to pay attention to commitment to use the feedback, which predicts both rate of change and likelihood to use feedback, when implementing feedback into clinical practice. Unwillingness to use feedback may be due to uneasiness regarding receiving feedback on one's performance. After professional training and licensing, therapists no longer receive structured feedback on their performance. Not using the feedback might be a way to cope with the anxiety of not being a good therapist. That is consistent with our finding

that therapists with higher self-efficacy were able to use the feedback to their benefit, although self-efficacy was not a significant predictor of actually using the feedback. An alternative interpretation is that not all therapists might be interested in enhancing their therapeutic skills. In any case, it is important to pay attention to the role of the therapists and their use of outcome monitoring feedback when aiming to use outcome monitoring as a tool to improve clinical outcomes. Under pressure of third parties, such as health insurance companies, many organizations just start measuring and do not pay sufficient attention to how feedback works and how therapists can effectively use it (De Jong, 2012).

This is the first study that has measured therapist factors in the context of a feedback intervention. Our results demonstrated that therapist characteristics are relevant and more research in this area is needed. Therapist characteristics that might be interesting to study include attribution style, locus of control, personality traits of the therapists and emotional stability. Therapist characteristics might be manipulated by training therapists in specific feedback-related skills. In addition, it would be important to get more insight in the dynamics of how feedback works and for whom. Perhaps some groups of therapists or patients perform worse when they are provided with feedback, but so far, we do not know if this is the case. Finally, for the further development of feedback, it is crucial that the premises of feedback theory are tested in a clinical context, since most of these theories originate from social and organizational psychology. Feedback effects are considered context specific and currently the contextualized feedback intervention theory (CFIT; Riemer & Bickman, 2011) is the only available theory that focuses on clinical practice. CFIT is complex and for the largest part untested, therefore alternative theoretical models could be explored as a basis to generate new hypotheses about how feedback works in clinical practice.

Acknowledgements

The authors would like to thank patients and staff from GGZ Noord-Holland-Noord and GGZ Dijk en Duin that have participated in this study, as well as the students that have assisted in the data collection, in particular Lianne Boom. Thanks are also due to Andrew McAleavey for his comments on the draft of this chapter.

