Different Perspectives of Short-Term Changes in the Rehabilitation of Schizophrenic Patients

Stefan Gunkel and Stefan Priebe

In the rehabilitation of schizophrenic patients, changes and outcome are usually assessed by members of multiprofessional therapeutic teams and by the patients themselves. In this study, the correlations of different perspectives of change during rehabilitation were investigated. After admission to a partial hospitalization program (printing workshop), 12 schizophrenic patients were examined weekly; the test period lasted 12 weeks. Ratings concerning various aspects of behavior were independently made by the patients themselves, by a work therapist, and by a psychiatrist, and week-by-week changes were analyzed. Factor analysis and correlations among change scores suggest a moderate to high homogeneity of each psychiatrist's and work therapist's ratings. Correlations among the three perspectives of change show little substantial concordance. Consequences for the practical work of multiprofessional teams in the rehabilitation of schizophrenic patients are discussed.

TREATMENT and rehabilitation of schizophrenic patients require efforts and cooperation of multiprofessional therapeutic teams and of patients. Therapeutic teams consist of caregivers from varying professions and training backgrounds such as psychiatrists, psychologists, social workers, nurses, and occupational therapists. Differences in professional background may lead to different perceptions and assessments of the patients' behavior and to different conceptions of improvement and deterioration. Therapeutic teams meet regularly and frequently—in many institutions once a week. In team discussions, experiences with the patient in different situations are compared and an agreement may be reached as to whether the patient has improved, deteriorated, or remained unchanged.

In the literature diagnosis on several levels, e.g., the multimodal assessment of psychopathological symptoms or abnormal behavior,[1-2] has been suggested. However, correlations of disparate levels of assessment and the clinical relevance of each level have hardly been clarified empirically. The psychiatric research on agreement between judgments of symptoms shows that intercorrelations of cross-sectionally obtained assessments remain mainly low to medium.[3-9] Most authors report coefficients (r) between .20 and .60.[10] Comparisons of self-ratings and clinical ratings have been conducted mainly with depressive patients[11-20] and less often with schizophrenic patients.[21-23]

Agreement between different raters is higher when the symptom to be assessed is defined in more concrete terms.[24-25] Correspondence is also higher in less disturbed samples[16,26-27] and increases with reduction of symptoms[1,13,28-31] and in remission phases.[32-33]

Thus far, just a few process-oriented studies in the synchronicity of different serial ratings have been undertaken, all with samples of depressives. In these studies, the correlation of change over time as assessed by consecutive assessments of different raters was investigated.[34] Von Zerssen et al.[35] report a middle-rank correlation of (r) .92 among sequential self-ratings of mood and specific clinical depression scores over a period of 25 sessions. We do not know of empirical studies with schizophrenic patients in which different serial assessments during the course of treatment or rehabilitation and their correspondence were examined. Yet during rehabilitation, it is change that is more relevant for evaluation of the process so far and for planning of future interventions than is an isolated assessment of patients' behavior at one point in time. Therapeutic teams consider mainly any change that has occurred when they discuss the experience of the previous week and plan care for the following days and weeks. In terms of clinical and practical aspects, the question of how different perspectives of change correspond over time is particularly meaningful. Many therapeutic decisions such as change of medication or referral to other therapeutic institutions depend on judgments made by dif-

---

From the Department of Social Psychiatry, Freie Universität Berlin, Germany.

Address reprint requests to Stefan Gunkel, Dipl.-Psych., Department of Social Psychiatry, Freie Universität Berlin, D-14050 Berlin, Platanenallee 19, Germany.

0010-440X/93/3405-0007$03.00/0

ferent professionals concerning changes of the patient's condition and behavior.

Therefore, in this study we investigated the correlation between independently taken, indirect assessments of change in partially hospitalized schizophrenic patients. The following three different perspectives were considered: (1) self-ratings by the patients, (2) assessments by a work therapist, and (3) clinical ratings by a psychiatrist.

METHOD

Subjects

The sample consisted of 12 chronic or subchronic schizophrenic patients; the primary diagnosis of schizophrenia was made according to DSM-III-R. The patients were consecutively admitted to a therapeutic printing workshop in the Department of Social Psychiatry at the Freie Universität Berlin. The workshop is integrated in a comprehensive community care system with three partial hospitalization programs, community-based services, and various outpatient facilities.[36]

The printing workshop is a partial hospitalization program (15 places) providing psychiatric treatment and vocational rehabilitation for severely disturbed and chronic patients.[37] At the same time, it is a small industrial printing shop equipped to compete in a free market. The workshop is directed by work therapists who have been trained in the techniques of printing and business contracts and who are experienced in the rehabilitation of mentally ill patients. The atmosphere is oriented to the usual world of work, and activities are clearly structured; all patients are regularly seen by clinical case managers. Patients were examined over a period of 12 weeks beginning at the end of the first week after admission. For two patients, treatment in the printing shop was terminated early (after 3 and 5 weeks, respectively). Ratings were made at the end of each week.

Measures

The areas of behavioral functioning to be assessed were selected such that the psychiatrist as well as the work therapist and the patients could rate aspects from their usual perceptual viewpoint. All ratings were made independently, i.e., psychiatrist, work therapist, and patients were blind toward each other's current or previous ratings.

The psychiatrist saw the patients only during the weekly interviews and was otherwise not involved in treatment. He assessed psychopathological symptoms and overall current degree of illness on the following scales: (1) Brief Psychiatric Rating Scale ([BPRS] summary score and five subscales),[38] (2) Hamilton Rating Scale for Depression (HAMD),[39] and (3) Clinical Global Impressions Scale (CGI).[40]

The work therapist also assessed patients on several scales, including (1) the Nurses' Global Impression Scale (NGI)[41] and (2) six 100-mm visual analog scales[42-44] on which the current degree of different abilities was rated between two extreme points (e.g., "not at all" *v* "very strong"). These scales are as follows: difficulty of task (DT) the patient is able to achieve, consistency of work performance (CP), accuracy of work performance (AP), general activity (GA^{w}), frequency of social contact (FC^{w}), and quality of social contact (QC^{w}). The mean score of the first three scales regarding work performance was taken as a performance index (PI).

Patients rated their condition on seven visual analog scales (VAS) concerning the following aspects: general condition (GC), anxiety (A), depression (D), work performance (WP), general activity (GA^{p}), frequency of social contact (FC^{p}), and quality of social contact (QC^{p}).

In the VAS, scores between the extreme points were read and transformed to numerical values from 0 (minimal extent) to 100 (maximal extent) so that the left end was positive in some scales (GC, A, D) and negative in others (WP, GA^{p}, FC^{w}, QC^{w}). The GC scale was formulated according to Aitken[42] with the question "How do you judge your current general condition?" ("generally good" = 0, "generally poor" = 100).

To analyze the concordance of the three perspectives of change assessment, data were transformed in the following way: only differences between scores of a given week and those scores of the same variable assessed one week before were taken into calculation ($X_{tn} - X_{tn-1}$).[9] Computation of those time-sequential differences reduces potential serially correlated errors caused by a repeated measurement procedure.

Thus week-by-week changes indirectly assessed by the psychiatrist, the work therapist, and the patients were analyzed. For these transformed change data, intercorrelations and factor analyses were calculated.[45-46] Because we were not interested in determination of the exact amount and level of significance, but in the direction and approximate size of possible relationships, we took the risk of unreliable and distorted correlation estimates. Methodological difficulties of dealing with dependence and autocorrelation of successive observations are discussed in detail by Wonnacott and Wonnacott (p. 226ff).[47]

RESULTS

Sample

The three female and nine male patients were aged between 18 and 43 years. Eight patients had been referred to the printing shop from psychiatric hospitals and four from other psychiatric services in the community. Seven patients lived alone, two with their parents, and two with a partner. One patient was living in the night clinic of the community care system while attending the printing workshop during the day.

None of the patients had any advanced education, and one had not finished any type of school; six had completed some sort of professional training. On average, the patients had not worked at a regular job for 2.3 years, but one patient had an occupation at the time of the study. The duration of illness ranged from 6 months to 19 years (mean, 4.4 years), and the number of previous full hospitalizations varied between zero and seven (mean, 3.0). Only one patient had been previously treated in the printing workshop. During the 12-week period of the study, patients had a mean of 10.5 single sessions with their case managers; the mean length of sessions was 29.9 minutes. Ten patients received neuroleptic medication at some point during the 12 weeks.

Overall Change and Homogeneity of Ratings

From the beginning to the end of the 12-week study period, mean scores of the whole sample showed positive overall improvement. In the last week, the work therapist assessed patients as being significantly better on each scale than at admission (VAS changes between 22 and 26 scale points, $t = 4.02$ to 6.06, each $P < .001$; NGI reduction from 5.8 to 5.1, $t = 2.69$, $P < .05$). The psychiatrist's ratings also showed a marked improvement by the end of the test period on the CGI (from 5.4 to 4.9, $t = 2.89$, $P < .01$), the HAMD (from 9.3 to 5.2, $t = 2.45$, $P < .05$), and the BPRS (from 35.5 to 29.5, $t = 2.48$, $P < .05$). On only two of five BPRS subscales was the symptom reduction statistically significant (anergia and thought disturbance). Self-ratings by the patients also changed mostly in a positive way during that time (GC, D, WP, GA^P, and QC^P); however, the differences were not statistically significant.

To determine the internal homogeneity of each rater's judgments, three intercorrelation matrices regarding changes since the preceding week were analyzed. All correlations regarding week-by-week changes considered 114 pairs of scores, i.e., all patients by all indirectly assessed short-term changes (11 for 10 patients, three for one patient, and one for one patient). Interscale correlations of weekly changes of the psychiatrist's (rBPRS-HAMD = .56, one-tailed $P < .001$; rBPRS-CGI = .35, $P < .001$; rHAMD-CGI = .15 NS) and work therapist's ratings (e.g., rDT-AP = .69, rGA^w-FC^w = .83, each $P < .001$) showed mainly medium to high homogeneity. Corresponding coefficients of patients' weekly rating changes were lower and less often statistically significant. Despite a poor general homogeneity of patients' ratings some substantial correlations were found, suggesting a logical consistency, e.g., between an increase of depression on the one hand and a decrease of general activity (rD-$GA^P = -.26$, $P < .01$), an increase of anxiety (rD-A = .48, $P < .001$), and a deterioration of general condition (rD-GC = .40, $P < .001$) on the other, as well as between self-rated deterioration of general condition and reduction of work performance (rGC-PW = −.32).

Week-by-Week Changes and Interrater Synchronicity

Interrater correlations were generally low; all remained below an *r* of .34, and most were statistically not significant. The frequency of significant correlations was just slightly higher than chance agreement. Only seven of 56 (patients and psychiatrist), then one of 63 (patients and work therapist), and then two of 72 (psychiatrist and work therapist) coefficients were at the 1% level of statistical significance.

Most associations were found between self-ratings and clinical ratings by the psychiatrist (Fig 1). Week-by-week changes of self-rated anxious feelings (A) and changes on HAMD, CGI, and the BPRS subscale anxiety/depression (rA-HAMD = .28, rA-CGI = .26, rA-BS1 = .26, each $P < .01$) were most clearly correlated. The BPRS sum score was significantly associated with self-ratings of depressive feelings (rD-BPRS = .23, $P < .01$).

Concordance of ratings was markedly lower

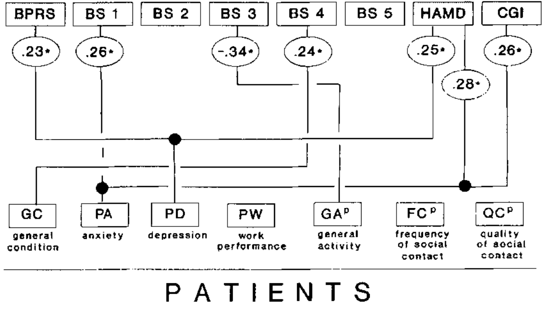

Fig 1. Significant correlations of week-by-week changes in the psychiatrist's ratings and in patients' self-ratings. $^*P < .01$, one-tailed.

regarding judgments by the work therapist and self-ratings. Most coefficients were found to be below an r of .15. Just one statistically significant relation, a positive correlation between changes of self-rated work performance and the frequency of social contact as rated by the work therapist, was found ($r\text{PW-FC}^{p} = .24, P < .01$).

Correlations between work therapist's and psychiatrist's ratings were very low as well. Again, only a few coefficients exceeded an r of .15, and just two were statistically significant. Moreover, those two disclosed an opposite change in the two perspectives; contrary to all expectations and despite using nearly identical scales (NGI, CGI), the work therapist and psychiatrist disagreed in their global clinical impression of illness ($r\text{NGI-CGI} = .23$; $P < .01$).

Results of a common factor analysis based on all intercorrelations of change ratings are shown in Table 1.

The first factor (18% of variance) is formed by ratings of the work therapist concerning working behavior (difficulty of tasks, consistency, and accuracy of performance). With this "performance factor," neither aspects of the patients' self-ratings, nor variables rated by the psychiatrist were substantially correlated. A second factor (14%) is represented by three behavioral aspects that were assessed by the work therapist and not directly related to work performance, i.e., aspects of the patients' social behavior and the level of activity.

On the third factor, only symptom scales (HAMD and BPRS sum scores, BPRS subscale anxiety/depression) rated by the psychiatrist have high loadings. The factor accounts for 11% of the variance and may be seen as a "psychopathology factor." Likewise, the third factor constitutes some subscales of the BPRS. Regarding patients' ratings, three separate latent factorial structures were found; however, the factors 4 through 8 should not be interpreted because of the low level of explained variance.

Table 1. Common Factor Analysis (varimax rotation) for Week-by-Week Changes: Psychiatrist's, Work Therapist's, and Patients' Ratings (loadings ≥ .50)

	Factors							
Perspectives/Variables	I	II	III	IV	V	VI	VII	VIII
Psychiatrist								
BPRS Sum score	.	.	.55	.78	.	.	.	.
BS1 Anxiety/depression	.	.	.87	.	.	.	.	.
BS2 Anergia	.	.	.	.79	.	.	.	.
BS3 Thought disturbance	.	.	.	.51	.	.	.	.
BS4 Activation	.	.	.	.	.	.	.	.
BS5 Hostile suspiciousness	.	.	.	.57	.	.	.	.
HAMD Sum score	.	.	.85	.	.	.	.	.
CGI	.	.	.	.	.	.	.68	.
Work therapist								
NGI	.	.	.	.	.	.	−.69	.
DT Difficulty of tasks	.89	.	.	.	.	.	.	.
CP Consistency of performance	.90	.	.	.	.	.	.	.
AP Accuracy of performance	.85	.	.	.	.	.	.	.
PI Performance index	.99	.	.	.	.	.	.	.
GA^{w} General activity	.	.90	.	.	.	.	.	.
FC^{w} Frequency of social contacts	.	.89	.	.	.	.	.	.
QC^{w} Quality of social contacts	.	.85	.	.	.	.	.	.
Patients								
GC General condition	.	.	.	.	.	.56	.	.
A Anxiety	.	.	.	.	.	.71	.	.
D Depression	.	.	.	.	.	.87	.	.
WP Work performance	.	.	.	.	.69	.	.	.
GA^{p} General activity	.	.	.	.	.67	.	.	.
FC^{p} Frequency of social contacts	.	.	.	.	.	.	.	.55
QC^{p} Quality of social contacts	.	.	.	.	.	.	.	.82
Explained variance (%)	17.7	13.8	10.6	7.2	6.6	5.7	5.4	4.7
Eigenvalue	4.2	3.3	2.6	1.7	1.6	1.4	1.3	1.1

Interpersonal Agreement on a Case Level

Finally, we analyzed only the direction and the global degree of any week-by-week change on a case level. Thus clear correspondences do occur between some variables or during particular weeks, although on the group level regarding week-by-week changes the correlation among the three raters remained limited in general. To illustrate this finding, Table 2 shows direction (improvement, deterioration, or no change) and global degree (slight, moderate, or marked) in week-by-week changes of a few central variables in one patient (no. 6).

In 9 of 11 weeks, CGI and NGI scores changed in the same direction and to the same degree, or both remained unchanged. BPRS score and performance as rated by the work therapist changed likewise in 8 weeks. In a separate analysis of time periods, all three raters generally agreed that the patient deteriorated during the fourth week and improved during the fifth. At other points of time, the raters clearly disagreed as to whether the patient had improved or deteriorated in the last week.

DISCUSSION

The psychiatrist and in particular the work therapist assessed more positive change in the patients' condition over the entire 12-week period than the patients did themselves. This finding is in concordance with results of research in psychotherapy showing that therapists tend to judge treatment effects as being more successful than do patients.[8] Possibly, caregivers wish to confirm the quality and usefulness of their therapeutic efforts.

In agreement with other studies,[8,24] the intraindividual homogeneity of expert ratings proved higher than the homogeneity of self-ratings. The psychiatrist and the work therapist tended to provide rarely differentiated judgments about the patients, although differing aspects of the patients' behavior were assessed by each of them.

When correlations among the three perspectives were considered, just a few substantial associations were found between changes in pairs of different ratings; all coefficients (*r*) were lower than .35. However, all correlations and corresponding levels of significance should be interpreted very carefully because of possible serially correlated errors in sequential data. In analysis of week-by-week changes, most of the significant correlations were between symptom ratings by the psychiatrist (HAMD, BPRS, CGI) and self-rated depression and anxiety. In assessing psychopathological symptoms, the psychiatrist relies on interviews with the patient; what the patient shows and expresses in these situations might be dominated by self-perceived depression and—partly—anxiety. The work therapist can base the judgment on behavior observation and not only on what patients say in an interview. As a result of a different perception and interpretation of behavioral aspects, the work therapist may be more likely to disagree with the patient than the psychiatrist. Although

Table 2. Week-by-Week Changes (improvement, stability, deterioration) in Ratings of Psychiatrist and Work Therapist and Self-Ratings for Patient No. 6

Perspectives (variables)	Direction and Strength of Week-by-Week Changes (reference intervals, weeks)										
	1	2	3	4	5	6	7	8	9	10	11
Psychiatrist											
BPRS	–	–	+	–	+	+	–	(–)	0	(–)	++
CGI	0	–	+	0	0	+	–	+	0	–	+
Work therapist											
NGI	0	–	+	–	+	+	–	+	0	–	+
Performance index	––	+	++	–	+	++	+	–	(–)	––	++
Patient no. 6											
General condition	++	––	++	––	+++	–––	––	–	+	++	––
Depression	–	–	(+)	–	++	–	–	–	+	–	++
Work performance	0	(–)	–	––	++	–	+	+	––	++	–

NOTE. Symbols are as follows: 0, unchanged; (+)/(–), very small improvement/deterioration; +/–, slight improvement/deterioration; ++/––, moderate improvement/deterioration; and +++/–––, marked improvement/deterioration.

there were single interrater correlations of considerable degree, in general correlations between short-term changes in the patients', psychiatrist's, and work therapist's ratings were low.

The moderate intraindividual homogeneity in the psychiatrist's and work therapist's ratings as well as the low interrater concordance are reflected by the results of the factor analysis. All three raters have a different perspective of change. The psychiatrist and work therapist each show a fairly homogeneous perspective, whereas the patients' view is less uniform. Thus the professional perceptions in caregivers are less differentiated and less contradictory in themselves, whereas patients rate and possibly view disparate aspects of behavior in a disparate way. It should be taken into account that the validity of VAS used as a self-rating method in the study remains unclear in schizophrenic patients. Given the long-term nature of rehabilitation, a week may be too short a period of time to expect dramatic improvement that might be obvious from any perspective. However, all three raters did perceive week-by-week changes, some of which were substantial; they just hardly agreed in their perception of change.

Why is the concordance of the three perspectives so low? All three raters based their assessments on different experiences in different periods of time. The psychiatrist saw the patients for just 1 hour a week; the work therapist spent the whole working-week with the patients; finally, the patients could consider experiences during the diagnostic interview with the psychiatrist, during work in the printing shop, and during the remaining time of the week, e.g., when with relatives, with friends, or alone. All three raters assessed partially different aspects of the patients' behavior. Thus the work therapist mainly rated actual work performance in the shop, whereas the psychiatrist assessed psychopathological symptoms on conventional scales.

Different aspects of behavior might change in a different way. However, the interrater correlations were not higher when identical aspects such as depression were assessed by two raters. In addition, the reported homogeneity of caregivers' ratings does not support the hypothesis that different behavioral aspects are viewed differently. The perspectives of the three raters may also be influenced by their special—partly professional—backgrounds and personal experiences in psychiatric settings.[48-49] Their social roles and their actual and emotional involvement in treatment differ. In the interview, the psychiatrist may focus on subjective symptoms and thus to some extent influence the patient's behavior and statements, whereas the work therapist concentrates on performance and might not ask about subjective symptoms. The low synchronicity of patients' and caregivers' perspectives may be affected by an illness-related bias in the patients' view. Paranoid symptoms and lack of insight into illness have been reported as diminishing agreement between self-ratings and clinical ratings.[27] Some patients of our sample showed paranoid symptoms, yet no patient was consistently without insight into illness. It remains open as to whether an illness-related bias has influenced the patients' ratings at some points of time and whether patients' understanding of the items[4,13,50] changed over time.

Any interpretation of the results should take into account the tentative character of the study and the small size of the sample. Although each correlation was based on 114 pairs of score differences from repeated measurements, treatment processes in only 12 patients were examined. For calculation of correlations, changes in different patients and changes of the same patients at different points in time were put together. This method—although believed to be appropriate—is associated with the problem of possible dependence and autocorrelation of successive scores.[47] Additionally, a high number of correlations was computed so that an adjustment of α-power would have clearly diminished the frequency of significant associations. The study was performed in a special treatment and rehabilitation setting and in a highly selective sample; thus the findings cannot easily be generalized to other samples of schizophrenic patients.

Despite these objections, it may be carefully concluded that different members of a multiprofessional therapeutic team or the patients themselves do not necessarily mean the same thing when they assess an improvement or a deterioration. Particularly when a change during the last week is the focus—as is typically the case in team discussion—the association among different perspectives is hardly better than coincidental, at least when the assessments are independently made. Therefore, in rehabilitation of schizophrenic patients, thorough team discussion and an integration of different perspectives—possibly more than just the three examined in this study—are essential and necessary for an adequate assessment of change. Furthermore, it remains open as to whether one of the investigated perspectives is more valid or more relevant within the rehabilitation process than the others.

ACKNOWLEDGMENT

The help of the patients who took part in this study is gratefully acknowledged. The authors also wish to acknowledge the cooperation and support received from A. Daudert, K. Hoffmann, N. Lichey, and M. Schönhoff during the study.

REFERENCES

1. Debus G. Selbst- und Fremdeinstufungen des aktuellen Befindens in psychopharmakologischen Längsschnittuntersuchungen. Arzneimittelforschung 1974;24:996-1001.

2. Seidenstücker G, Baumann U. Multimethodale Diag-

rating scales: a critical review. Arch Gen Psychiatry 1973;28: 361-366.

17. Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bowling TG. Quantitative rating of depressive states: corre-