Combined group and individual schema therapy for borderline personality disorder: A pilot study

Veronique Dickhaut [a], Arnoud Arntz [b],*

[a] Community Mental Health Center Maastricht, The Netherlands
[b] Clinical Psychological Science, Maastricht University, PO Box 616, The Netherlands

ARTICLE INFO

ABSTRACT

Background and Objectives: Schema Therapy (ST) is a highly effective treatment for Borderline Personality Disorder (BPD). In a group format, delivery costs could be reduced and recovery processes catalyzed by specific use of group processes. As patients may also need individual attention, we piloted the combination of individual and group-ST.
Methods: Two cohorts of BPD patients ($N = 8$, $N = 10$) received a combination of weekly group-ST and individual ST for 2 years, with 6 months extra individual ST if indicated. Therapists were experienced in individual ST but not in group-ST. The second cohort of therapists was trained in group-ST by specialists. This made it possible to explore the training effects. Assessments of BPD manifestations and secondary measures took place every 6 months up to 2.5 years. Change over time and differences between cohorts were analyzed with mixed regression.
Results: Dropout from treatment was 33.3% in Year 1, and 5.6% in Year 2, without cohort differences. BPD manifestations reduced significantly, with large effect sizes, and 77% recovery at 30 months. Large improvements were also found on general psychopathological symptoms, schema (mode) measures, quality of life, and happiness. Cohort-2 tended to improve faster, but there were no differences between cohorts in the long term.
Limitations: The study was uncontrolled, training effects might have been non-specific, and the sample size was relatively small.
Conclusions: Combined group–individual ST can be an effective treatment, but dropout might be higher than from individual ST. Addition of specialized group-ST seems to speed up recovery compared to only individual ST.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Borderline personality disorder (BPD) is a severe mental condition, characterized by a pervasive pattern of instability in moods, interpersonal relationships, self-image and behavior. The prevalence is estimated to be 1–2% of the general population and ranges from 10 to 20% among outpatient and inpatient individuals treated in mental health clinics (American Psychiatric Association, 2005). Specific structured therapies have demonstrated efficacy in reducing BPD-symptoms in randomized controlled trials, such as dialectical behavior therapy (Linehan, Armstrong, Suarez, Allmon, & Heard, 1991) and cognitive therapy (Davidson et al., 2006). In the last decade more comprehensive treatments which aim at full recovery have been tested. Various treatments seem promising.

Among them is Schema Therapy (ST; Arntz & van Genderen, 2009; Young, Klosko, & Weishaar, 2003). In a multicenter trial in which ST was compared to Transference Focused Psychotherapy (TFP; Yeomans, Clarkin, & Kernberg, 2002) ST turned out to have better treatment retention and to be more effective on various measures (Giesen-Bloo et al., 2006). ST was also more cost-effective than TFP with lower societal costs and stronger effects (van Asselt et al., 2008). A second study demonstrated that ST can be successfully implemented in regular practice, and that telephone availability outside office hours is not necessary (Nadort et al., 2009).

The duration of ST makes the therapy expensive, and problematic to deliver to all patients requesting it. These are compelling reasons to use a group therapy format. Other advantages of group therapy relate to the curative factors as described by Yalom and Leszcz (2005). Among these are universality, getting and giving emotional support, modeling, sense of belonging, practicing interpersonal skills and bonding. Patients can experience the satisfaction of being helpful to others and by doing so bolster their self-

---

* Corresponding author.
 E-mail address: Arnoud.Arntz@maastrichtuniversity.nl (A. Arntz).

confidence. An important assumption in working with patients with BPD in groups is that they recognize each other's problems faster and easier than their own problems, and as a consequence patients can validate, support, confront and advise one another. Moreover, patients often experience such responses by other patients as more genuine than when made by a therapist. For these reasons, it has been argued that group-ST might "catalyze" the change processes of ST, thus leading to faster and deeper changes than individual ST (Farrell & Shaw, 2012; Farrell, Shaw, & Webber, 2009).

We developed a protocol for outpatient treatment, in which ST in group was combined with individual ST. We assumed that individual treatment was essential for a number of reasons. We considered individual attention and attachment a basic need of the BPD patient, and it was our opinion that trauma processing is preferably offered in individual sessions, where specific techniques can be used to process painful and disturbing memories, that might be too confronting for other group members. An additional argument is that the combination mimics natural development of attachment with different persons (parent and peers).

During the study, we learned that Farrell and Shaw (2012) had developed a specialized group-ST model, of which a first RCT indicated very strong effects (Farrell et al., 2009). The group process is handled in a very specific way, which demands specific behavior and collaboration of the therapist pair. Our therapists were trained in their method; but the first cohort was already halfway through treatment and the second hadn't started when the training took place. This offered us the possibility to explore whether there were any differences between the two cohorts associated with the use of the Farrell and Shaw model.

2. Methods

2.1. Patients

Patients referred to the Community Mental Health Center of Maastricht, with a primary diagnosis of BPD, based on the Structured Clinical Interviews for the DSM-IV, I and II (First, Gibbon, Spitzer, Williams, & Benjamin, 1997; First, Spitzer, Gibbon, & Williams, 1996; Groenestijn, Akkerhuis, Kupka, Schneider, & Nolen, 1999; Weertman, Arntz, & Kerkhofs, 2000) were asked to participate in the study. If they agreed, they were further screened, by using a semi-structured clinical interview, the Borderline Personality Disorder Severity Index, fourth edition (BPDSI-IV) (Giesen-Bloo, Wachters, Schouten, & Arntz, 2010). If scores were >20, further baseline measurement took place.

Inclusion criteria were: a main diagnosis of BPD, BPDSI-IV score >20, age 18–60, IQ >80 and Dutch literacy. IQ was only tested in case of doubt. Exclusion criteria were: an axis-I disorder that generally needs primary treatment. These were psychotic disorder (except short reactive psychotic episodes belonging to BPD), manic episodes, attention deficit/hyperactivity disorder (ADHD), addiction of such severity that detoxification was indicated (other addictions were not excluded), anorexia nervosa, autistic disorder. Also >2 Narcissistic or >2 Antisocial PD traits were exclusion criteria, as these are likely to be disruptive to BPD group treatment. Males were excluded if they would be the only one in the group. Eighteen women with a primary diagnosis of BPD were included (see Fig. 1 for the consort flow diagram).

2.2. Outcome measures/assessment

The primary outcome measure was the score on the BPDSI-IV, a DSM-IV based semi-structured interview that assesses frequency and severity of BPD manifestations during the last 3 months (Arntz et al., 2003; Giesen-Bloo et al., 2010). The interview is appropriate for repeated measurements and therefore for treatment evaluation. This instrument shows excellent psychometric properties (Cronbach's alpha = .85 in a BPD sample, .96 in a heterogeneous sample; interrater reliability .99; excellent validity and sensitivity to change). The BPDSI has a cut-off score of 15 between patients with BPD and controls, with a specificity of .97 and a sensitivity of 1.00 (Giesen-Bloo et al., 2010). The recovery criterion was therefore defined as achieving a BPDSI-IV score of less than 15 and maintaining this score until the last assessment. An independent research assistant, trained in the BPDSI, administered the BPDSI-IV.

Secondary outcome measures were the following self-report instruments. The BPD-checklist inquires for someone's experienced burden of BPD complaints during the last month (Giesen-Bloo, Arntz, & Schouten, 2006). It is complementary to the BPDSI-IV in the way that it reflects the patients' experienced change, where the BPDSI-IV examines the persons' objective change on BPD symptomatology. The SCL-90 (Derogatis, Lipman, & Covi, 1973) measures subjective distress from a range of psychopathological symptoms. Quality of Life was assessed with the mean WhOQoL Group: Development of the World Health Organization (1998), short version item score, and the thermometer scale of the EuroQoL (range 0–100; Brooks 1996), which assesses primarily subjective physical health state. Happiness was assessed with the 1-item happiness question validated in more than 30 countries, with the following response possibilities: (1) completely unhappy; (2) very unhappy; (3) fairly unhappy; (4) neither happy nor unhappy; (5) fairly happy; (6) very happy; (7) completely happy (Veenhoven, 2008). ST-specific measures were the Young Schema Questionnaire (YSQ; Rijkeboer, van den Bergh, & van den Bout, 2005; Young & Brown, 1994) of which the total score was used; and the Schema Mode Inventory (Lobbestael, van Vreeswijk, & Arntz, 2008; Lobbestael, van Vreeswijk, Spinhoven, Schouten, & Arntz, 2010), of which the mean item score of dysfunctional modes and the mean item score of functional modes was used.

2.3. Treatment and therapists

The treatment protocol consisted of weekly *90-min* group sessions led by two therapists, *combined with weekly 1-h* individual sessions. For pragmatic reasons, group therapists did not have to be the same as the individual therapists. After the first year, the frequency of individual sessions could diminish when therapist, peer supervision group and patient agreed. Individual sessions followed the Arntz and van Genderen (2009) protocol and had the specific aim to support group sessions (e.g., to help with problems patients had in dealing with the group), to deal with crises, and to do extensive trauma processing.

Before group therapy started, some individual sessions were required. In these sessions case conceptualizations in terms of the mode model were made. Patients were prepared by addressing worries patients had about the group. The number of pretherapeutic sessions differed per patient, from 2 to 12 sessions. Treatment integrity was monitored by means of peer supervision. Therapists were all experienced schema therapists, but none of them had any experience with doing group-ST.

The treatment protocol addressed the theoretical model of ST, different phases of (group) therapy, and ST techniques. Central to the theoretical model of ST in working with BPD is the assumption of 6 schema modes. The distinguished modes in BPD are vulnerable (abandoned/abused) child, angry/impulsive child, punitive parent, detached protector or any other protective mode, and the functional healthy adult and happy child modes. Schema modes are sets of schemas expressed in pervasive patterns of thinking, feeling and behaving. Important goals of ST are to show empathy with and
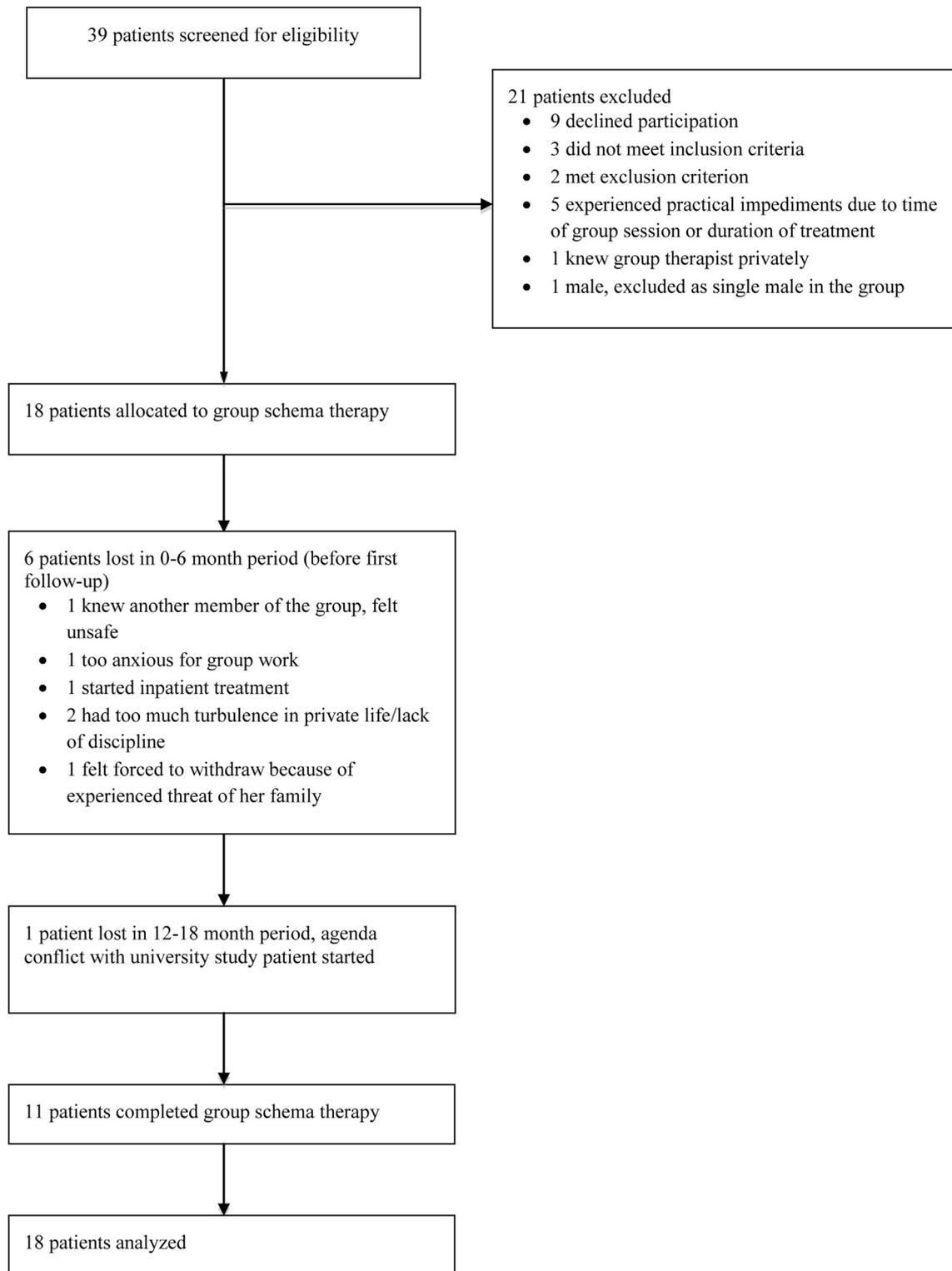
Fig. 1. Patient flow chart.

protect the Vulnerable Child, to challenge the Punitive Parent, to set limits to the Angry/Impulsive child, to reassure the Protective mode, to enhance the Healthy Adult and promote expression of the Happy Child.

Group treatment was time-limited to 2 years. To ensure safety no additional patients were introduced after the first session and structure was maintained on a high level. Sessions did not have an

established program, but every week a session plan was made, depending on the process of the group and individual patients. Main goal for the first stage of treatment was mutual understanding, bonding and cohesion; and to get acquainted with the mode model and the corresponding language. In the working phase more and more behavioral, cognitive and experiential techniques were introduced. The aim of the last phase was preparing for the

end of and the future without the group. If necessary, individual sessions were continued after 2 years.

Group therapists of the first group-ST started without training for working in ST groups. After 9 months of treatment, all involved therapists *received a three day (20.5 h) training* by Farrell and Shaw (2012), Farrell et al. (2009), who had more than 20 years of experience with group-ST. Farrell and Shaw developed a therapeutic style and collaboration between therapists to optimally conduct ST in group format. They stress the need to focus on and balance the individual needs of each participant and the collective need of the group, similar as a parent would for a group of siblings. As a consequence of the training, therapists' style changed, and new therapeutic techniques were used. Most important changes were: stressing the communality between group members and support the connection between group members by constantly broaden conversation (instead of working individually in front of the group), clear division of roles between the therapists (with at least one therapist taking care that all patients stay involved), enhancing security by setting clear rules and limits, offering more structure in sessions, de-escalate in times of crisis, providing written information and psychoeducation, giving homework assignments and paying more attention to the happy child mode.

The group therapists of the second group-ST started right after this training and used the Farrell and Shaw format. The therapists of the first group reported only limited success in changing style and methods. Group and individual therapists of both cohorts had weekly peer supervision and no expert supervision. There was no difference between cohorts in amount of peer supervision.

All group members participated in the study and if patients wanted to quit group therapy, individual ST stopped. If desired, patients could in that case be referred to another treatment.

2.4. Statistical analysis

The statistical analysis was based on the intent-to-treat principle, using all available data from all participants that started treatment. We used mixed regression for longitudinal data as this method can deal with missings and yields more valid estimates of effects than analyzing completers only or using last observation carried forward imputation (Schafer & Graham, 2002). The six 6 months assessments during 2.5 years constituted the repeated measure (i.e., time); and fixed factors/covariates were cohort, time, cohort by time interaction. Any variable significantly different between cohorts at baseline was planned to be added as additional covariate. We assessed linear, quadratic and cubic time trends (except for the analysis of recovery, see below).

We first checked the distribution of the data. One case had extreme scores on baseline BPDSI and BPD-checklist, creating outliers. These were pulled in to 2 SD, and remained the highest scores. No further serious deviations from normal distribution were observed. Next differences at baseline between cohorts were examined. Groups differed significantly on baseline WHOQoL and EuroQoL thermometer (Table 2), thus these were considered as covariates. However, EuroQoL baseline didn't contribute to the discrimination between groups after WHOQoL baseline (the strongest predictor) was entered in a logistic regression, therefore only the (centered) WHOQoL baseline score was used as covariate.

Recovery, defined by BPDSI score <15 (Giesen-Bloo et al., 2010) was analyzed by mixed logistic regression. Complex covariance structures of the repeated measures part led to estimation problems; therefore a simple compound symmetry model was used, with random intercept. As analyses with time modeled as a dimensional covariate failed to converge, time was included as a factor with 6 levels (assessments as levels). Addition of the WHOQoL baseline as predictor also led to estimation problems, the covariate was

therefore not included. Thus, the fixed part consisted of intercept, group, assessment, and group by assessment interaction.

Dimensional outcomes were analyzed with an unstructured covariance structure for the repeated measures part (being the best fitting). For the fixed part, intercept, centered group, centered WHO-QoL baseline, time trends (linear, quadratic and cubic) and group by time trends were analyzed. For the WHOQoL analysis, the baseline was left out of the analysis as WHOQoL baseline was already covariate. When $p \geq .10$, interactions were deleted from the model, in hierarchical order: first cubic, then quadratic, then linear time trend × group. Effect sizes were reported as $r = \sqrt{(t^2/(t^2 + df))}$ for the mixed regression outcomes, and for the conventional treatment effects from baseline to last assessment (30 months) as Cohen's $d = $ mean change/SD, with mean change derived from the mixed regression estimates and pooled baseline SD as numerator (Feingold, 2009).

3. Results

3.1. Patient accrual

The consort diagram is presented in Fig. 1. Thirty-nine patients were referred after intake and informed about the study. Twenty-one were excluded (14 patients after initial contact, 7 after the complete screening procedure). Nine patients declined participation (5 patients felt not being able to commit themselves to a 2 year period of therapy, 4 did not want to have therapy in a group), 3 did not meet inclusion criteria (2 had BPDSI-IV scores <20 and 1 had only 4 traits on BPD), 2 met exclusion criteria (1 had ADHD and 1 had 3 narcissistic traits on SCID II), 5 had the intention to start but experienced practical impediments that made participation impossible, such as moving to another city, or inconvenient time of group sessions, and 1 was male, and was excluded because he would be the only male patient in group. One patient knew one of the group therapists privately, which made it not possible to include her. Eventually 18 patients participated in the study.

The first cohort consisted of 8 patients. In the first week 1 patient dropped out of treatment, after 11 weeks a second patient quit. For a long time the first group consisted of 6 patients. After almost one and a half year a last patient dropped out, because she went to university and could not attend group sessions anymore. The second cohort started treatment 9 months later. This cohort consisted of 10 patients. Within the first 6 months 4 patients dropped out for various reasons, like too much anxiety, too much turbulence in daily life or expected rejection and punishment by (immigrant) family if patient would reveal being in treatment. The difference in dropout (group 1: 37.5%, group 2: 40%) was not significant, Fisher's exact test $p = 1.00$.

3.2. Treatment groups at baseline

Table 1 gives an overview of the main patients' characteristics at baseline. Patients were all female, mainly in their 20s and 30s and with average educational levels. There were no significant differences between the two cohorts in these variables. Table 2 provides an overview of differences between the groups on outcome parameters at baseline. As can be seen, WHOQoL and EuroQoL thermometer differed significantly between groups at baseline. The WHOQoL was used as covariate to correct for this (the EuroQoL being redundant as extra covariate).

3.3. Primary outcome

3.3.1. Recovery

Using the BPDSI <15 criterion, recovery rates by group and assessment were estimated by means of mixed logistic regression.

Table 1
Sociodemographic and clinical characteristics of starters.[a]

	Treatment group (N = 18)
Age, mean (SD)	28.5 (8.7) Range: 19–46
Education	
Graduate (WO en HBO)	2 (11.1)
Some college (LBO, MBO)	6 (33.3)
High school graduate (MAVO, HAVO, VWO)	9 (50.0)
Grades 7–11	1 (5.6)
Employment status	
Housewife	1 (5.6)
Student	5 (27.8)
Employed	4 (22.2)
Disability	4 (22.2)
Welfare	4 (22.2)
Psychotropic medication use at baseline	13 (72.2)
Recent suicide planning, steps or attempts[b]	7 (38.9)
Recent nonsuicidal self-injury[c]	10 (55.6)
Number of Axis I diagnoses	2.38
Number of Axis II diagnoses (incl. BPD)	1.89
No. of patients who had previous treatment	7 (4 patients had 2 and 3 patients had 3 previous treatments)

[a] Data are given as number (percentage) except where otherwise indicated.
[b] According to BPDSI-IV items 5.11–5.13 in the previous 3 months.
[c] According to BPDSI-IV items 5.1–5.8 in the previous 3 months.

Fig. 2 presents the estimated proportions. At 24 months, immediately after the group treatment finished, the proportions differed significantly, $F(1,65) = 4.07$, $p = .048$, with the second group having a higher proportion of recovered patients (66.5% (95%CI [29.2, 90.5]) vs. 18.7% (95%CI [2.8, 65.1])). At FU the difference disappeared and the estimated recovery was 77.4% (95%CI [45.9, 93.3]), that is 14 out 18.

3.3.2. BPDSI total score

Mixed regression revealed significant linear, squared and cubic time effects (Table 3). There was a trend toward a significant group by cubic time effect ($p = .052$), reflecting the relatively slower reduction of BPDSI scores in group 1, and the quick reduction to similar BPDSI levels after the group was finished as in group 2 (Figs. 2 and 3). Cohen's d was $23.14/8.50 = 2.72$ for change from baseline to last observation (average of both groups).

3.4. Secondary outcomes

3.4.1. BPD-checklist

After deleting the cubic time by group interaction ($p > .10$), mixed regression revealed significant linear, squared and cubic time effects, a significant group by quadratic time effect, and a trend toward a group by linear time effect (Table 3). The interactions reflect the relatively slower reduction in BPD-checklist

Table 2
Baseline values of outcome variables of the two groups.

Variable	M (SD)		t/U	p
	Group 1	Group 2		
BPDSI	35.53 (6.25)	34.26 (11.82)	t = .27	.79
BPD-checklist	116.13 (20.43)	114.40 (30.46)	t = .14	.89
SCL-90	256.25 (48.68)	228.70 (72.16)	t = .92	.37
Happiness	2.75 (1.16)	3.60 (1.17)	t = −1.53	.15
EuroQoL thermometer	50.13 (12.38)	63.40 (14.49)	U = 54	.049
WHOQoL	1.80 (.32)	2.24 (.49)	t = 2.19	.043
SMI dysfunctional	3.18 (.57)	3.20 (.66)	t = −.05	.96
SMI functional	2.81 (.68)	3.15 (.67)	t = −1.05	.31
YSQ total	44.70 (9.07)	42.36 (10.32)	t = .50	.62

Note. t = t-value of t-test with df = 16; U = U-value of Mann–Whitney test.

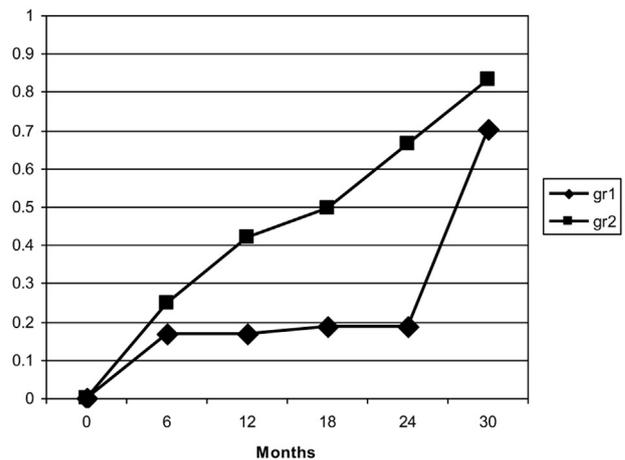**Estimated Proportions Recovered**

Fig. 2. Estimated proportions recovered participants by group from mixed logistic regression. At 24 months (posttest) the groups differed significantly, $F(1,65) = 4.07$, $p = .048$.

scores in group 1 during months 6–24, which disappeared at 30 months (Fig. 2). Cohen's d was $51.72/22.06 = 2.34$ for change from baseline to last observation (average of the two groups).

3.4.2. SCL-90

Mixed regression revealed significant interactions of quadratic and cubic time trends with group, reflecting the relatively slower reduction in BPD-checklist scores in group 1 during months 6–24, which disappeared at 30 months (Fig. 2, Table 3). Cohen's d was $89.79/60.42 = 1.49$ for change from baseline to last observation (average of the two groups).

3.4.3. Happiness

Mixed regression revealed that the cubic time trend by group interaction was n.s. ($p = .39$). The results after deleting this interaction from the model showed significant linear, quadratic and cubic time effects, as well as a significant quadratic time by group interaction (Table 3), reflecting quicker increase in happiness ratings in group 2 (Fig. 2). Cohen's d was $2.07/1.17 = 1.77$ for change from baseline to last observation (average of the two groups). Veenhoven (2008) reports Dutch general population norms: a mean of 5.28 with SD = .83. This implies that at last observation participants fell on average well in the normative range (mixed regression estimated mean = 5.23), and individual estimates showed that all fell within the 95% CI of the population norms.

3.4.4. EuroQoL thermometer

After deleting the n.s. group by cubic and quadratic time interactions, mixed regression revealed a significant group by linear time interaction, next to the main time effects (Fig. 2, Table 3). As baseline WHOQoL was n.s. ($p > .20$) the covariate was deleted from the model (inclusion led to highly similar results). The second group had a significantly stronger increase in thermometer ratings. As conditions differed at 30 months, separate effect sizes of the change with respect to baseline were calculated: group 1 $d = 4.66/13.44 = .35$; group 2 $d = 20.41/13.44 = 1.52$ (average $d = .94$).

3.4.5. WHOQoL

After deleting the group by cubic time trend interaction and the main cubic time effect (p's > .10), significant group by quadratic and group by linear time trend interactions were found (Table 3). As

Table 3
Results of mixed regression analyses.

Variable effect	Estimate	95% CI	S.E.	t	df	p	r
BPDSI							
Intercept	33.43	30.20; 36.67	1.50	22.31	13.13	<.001	.99
WHOQoL_bl	−6.92	−10.06; −3.77	1.41	−4.91	9.82	.001	.84
Group	2.80	−4.30; 9.89	3.34	.84	15.45	.42	.21
Time linear	−13.64	−17.32; −9.96	1.64	−8.33	9.43	<.001	.94
Time quadratic	4.98	3.34; 6.63	.71	7.04	7.73	<.001	.93
Time cubic	−.64	−.86; −.41	.10	−6.35	8.25	<.001	.91
Group × Time lin	3.17	−4.21; 10.54	3.28	.97	9.37	.36	.30
Group × Time qua	−2.75	−6.05; .56	1.42	−1.93	7.62	.09	.57
Group × Time cub	.46	.004; .92	.20	2.28	8.10	.052	.63
BPD-checklist							
Intercept	114.84	109.60; 120.07	2.20	52.30	6.72	<.001	1.00
WHOQoL_bl	−16.33	−21.55; −11.11	1.59	−10.28	2.84	.002	.99
Group	16.08	5.56; 26.60	3.87	4.16	4.22	.013	.90
Time linear	−35.88	−48.61; −23.15	5.58	−6.43	8.53	<.001	.91
Time quadratic	14.88	8.75; 21.02	2.70	5.51	8.83	<.001	.88
Time cubic	−1.96	−2.74; −1.17	.35	−5.67	8.86	<.001	.89
Group × Time lin	−11.93	−24.80; .94	5.77	−2.07	9.93	.066	.55
Group × Time qua	2.70	.00; 5.40	1.17	2.30	8.14	.05	.63
SCL-90							
Intercept	235.06	216.81; 253.31	8.55	27.49	14.82	<.001	.99
WHOQoL_bl	−39.72	−44.88; −34.56	2.08	−19.13	5.63	<.001	.99
Group	4.66	−32.01; 4134	17.24	.27	15.39	.79	.07
Time linear	−33.32	−66.56; −.08	15.03	−2.22	10.57	.050	.56
Time quadratic	13.74	−2.31; 29.80	7.30	1.88	11.04	.086	.49
Time cubic	−2.13	−4.32; .05	.99	−2.17	10.52	.054	.56
Group × Time lin	55.33	−11.15; 121.81	30.06	1.84	10.58	.094	.49
Group × Time qua	−39.69	−71.81; −7.57	14.60	−2.72	11.04	.020	.63
Group × Time cub	6.07	1.70; 10.43	1.97	3.08	10.53	.011	.69
Happiness							
Intercept	3.21	2.88; 3.55	.16	20.62	16.01	<.001	.98
WHOQoL_bl	1.06	1.03; 1.10	.02	67.36	7.06	<.001	1.00
Group	−.18	−.84; .47	.31	−.59	16.10	.56	.15
Time linear	1.50	.68; 2.33	.38	3.94	12.90	.002	.74
Time quadratic	−.66	−1.05; −.26	.18	−3.58	12.55	.004	.71
Time cubic	.09	.04; .14	.02	3.67	12.35	.003	.72
Group × Time lin	.34	−.39; 1.07	.33	1.04	10.54	.32	.30
Group × Time qua	−.10	−.20; −.01	.04	−2.41	10.43	.036	.60
EuroQoL Thermometer							
Intercept	56.83	50.12; 63.54	3.17	17.19	16.50	<.001	.98
Group	7.77	−1.76; 17.30	4.45	1.75	14.30	.10	.42
Time linear	13.19	2.48; 23.90	5.00	2.64	14.22	.019	.57
Time quadratic	−6.73	−13.48; .02	3.14	−2.14	13.81	.050	.50
Time cubic	.92	−.07; 1.90	.46	2.01	13.69	.065	.48
Group × Time lin	3.15	2.25; 4.05	.40	7.86	9.49	<.001	.93
WHOQoL							
Intercept	2.14	1.91; 2.37	.10	20.74	10.74	<.001	.99
WHOQoL_bl	.19	.003; .38	.09	2.24	10.84	.047	.56
Group	−.12	−.64; .40	.24	−.50	14.59	.62	.13
Time linear	−.06	−.21; .10	.07	−.78	11.46	.45	.22
Time quadratic	.03	−.003; .05	.01	2.42	11.53	.033	.58
Group × Time lin	.39	.08; .71	.14	2.75	11.46	.018	.63
Group × Time qu	−.08	−.13; −.02	.02	−3.21	11.53	.008	.69
SMI dysfunctional scales							
Intercept	3.10	2.88; 3.33	.11	28.76	14.16	<.001	.99
WHOQoL_bl	−.42	−.68; −.16	.12	−3.51	13.74	.004	.69
Group	.39	−.14; .91	.25	1.56	15.16	.14	.37
Time linear	−.10	−.23; .02	.06	−1.82	10.19	.099	.49
Time quadratic	−.01	−.03; .01	.01	−.87	9.83	.40	.27
Group × Time lin	−.44	−.69; −.19	.11	−3.95	10.17	.003	.78
Group × Time qu	.09	.04; .13	.02	4.45	9.80	.001	.82
SMI functional scales							
Intercept	2.97	2.81; 3.13	.07	39.78	13.16	<.001	1.00
WHOQoL_bl	.49	.33; .64	.07	6.92	11.40	<.001	.90
Group	−.22	−.57; .13	.16	−1.33	14.77	.21	.33
Time linear	.55	.09; 1.01	.21	2.59	12.24	.023	.60
Time quadratic	−.20	−.43; .03	.10	−1.96	11.17	.076	.50
Time cubic	.03	−.002; .06	.01	2.05	10.60	.066	.53
Group × Time lin	.20	−.18; .58	.17	1.16	11.30	.272	.32
Group × Time qu	−.06	−.12; .008	.03	−1.92	11.48	.081	.49
YSQ total score							
Intercept	43.86	40.29; 46.13	1.34	32.36	14.86	<.001	.99
WHOQoL_bl	−7.64	−10.84; −4.45	1.50	−5.11	14.62	<.001	.80
Group	5.44	−1.11; 11.99	3.09	1.76	15.80	.097	.41
Time linear	−5.10	−9.04; −1.16	1.78	−2.87	10.57	.016	.66

Table 3 (continued)

Variable effect	Estimate	95% CI	S.E.	t	df	p	r
Time quadratic	1.61	.004; 3.22	.72	2.24	9.71	.050	.58
Time cubic	−.25	−.42; −.07	.08	−3.18	9.60	.010	.72
Group × Time lin	2.04	−5.83; 9.92	3.56	.57	10.58	.58	.17
Group × Time qu	−2.77	−5.98; .45	1.44	−1.92	9.72	.084	.53
Group × Time cu	.47	.13; .82	.15	3.06	9.61	.013	.70

Note. Effect size $r = \sqrt{(t^2/(t^2 + df))}$; WHOQoL_bl = centered baseline WHOQoL covariate; lin = linear; qu = quadratic; cu = cubic.

Fig. 2 shows, this reflected the initially stronger increase in WHO-QoL scores in the second group, compared to the first group. Treatment effect size based on the average effect of the two groups was $d = .53/.405 = 1.31$.

3.4.6. Schema Mode Inventory − dysfunctional modes

After deleting the cubic time trend effects with $p > .10$, the interactions of linear and quadratic time trends with group were significant, Table 3. As is shown in Fig. 2, the reduction in dysfunctional schema mode scores took place earlier and faster in the second group, as compared to the first group. The treatment effect size based on the average effect of the two groups was $d = .72/.62 = 1.16$.

3.4.7. Schema Mode Inventory − functional modes

After deleting the cubic time trend by group interaction with $p > .10$, a trend toward significance in the quadratic time by group interaction emerged (Table 3). As is shown in Fig. 2, there was a more rapid improvement in functional SMI scales in group 2 compared to group 1. The final treatment effect size of the two groups combined was $d = 1.004/.677 = 1.48$.

3.4.8. Young Schema Questionnaire (YSQ)

As is clear from Fig. 2 and Table 3, the strong decreases in YSQ scores were qualified by a group by cubic time trend interaction. Group 1 lagged behind in decreases in YSQ scores compared to group 2, but attained the same final effect at month 30. Treatment effect size of the combined groups was $d = 15.865/9.695 = 1.64$.

4. Discussion

The first objective of this pilot study was to explore the effectiveness of a combined group and individual ST format for BPD patients. An important difference between the two cohorts we studied was that therapists of the second cohort received training in group-ST for BPD along the Farrell and Shaw model, whereas the first therapists developed their own methods. Thus, a secondary aim became to *explore* whether training had an effect on treatment outcome.

The results show in both cohorts a significant reduction of BPD symptomatology as measured by the primary outcome measure, the BPDSI, with a very high effect size (Cohen's $d = 2.72$ over both cohorts) at 2.5 years follow-up. Recovery estimated by mixed regression at this point was also very high (77.4%). Results on the BPD-checklist were comparable: burden of BPD-symptoms decreased strongly (Cohen's $d = 2.34$ at 2.5 year). SCL-90 scores indicate that there was also a significant reduction in general psychopathology. Indices of quality of life and happiness indicated strong improvement in wellbeing. Of special interest is the finding that participants reached a normative range of happiness, indicating that the treatment did not only reduce psychopathology but also helped patients to develop a satisfying life.

The clinical impression was that the second cohort improved faster and more than the first. We hypothesize that this was a consequence of training of the group therapists in the Farrell and Shaw model. The statistical analyses indeed indicate that the second cohort improved faster than the first cohort, though in two of the nine dimensional outcomes this pattern was only a trend toward significance. The figures however clearly show that in the 6 months after the group ended, the first cohort caught up with the second cohort, with the exception of the EuroQoL thermometer. Cohort-1 therapists reported more problems with the group process. Perhaps the end of the group and extra individual ST after that helped cohort-1 patients to reach a similar endpoint as cohort-2 patients.

Compared with individual ST reported by Giesen-Bloo et al. (2006), in which many of the patients received 3 years of therapy, we found a higher percentage of recovery on the BPDSI. The same holds when compared to Nadort et al., (2009), who tested results after 1.5 years of treatment in an ST implementation study, but here the different time frame might explain differences. Nordahl and Nysaeter (2005) performed an individual case series in which patients were treated between 18 and 36 months. Three out of six patients were evaluated to be recovered. Farrell et al. (2009) were so far the only ones to use a group-ST format. Their program consisted of thirty weekly group sessions in addition to (individual) usual treatment over a period of 8 months. 94% of the patients in the ST condition would no longer be diagnosed with BPD after treatment. Thus, the strong effects of our combined group−individual ST format support that this model is effective and suggests that it might be even more effective than individual ST, perhaps by the "change catalyzing" factors suggested by Farrell and Shaw (2012).

It should be noted that the group-ST tested in the present study is very different from the more structured and cognitive group protocol evaluated in two previous open trials (Renner et al., 2013; van Vreeswijk, Spinhoven, Eurelings-Bontekoe, & Broersen, 2012). The structured schema cognitive-behavioral therapy for groups (SCBT-g) has only 20 sessions, uses a highly structured program, with less experiential and group-based processes than the Farrell and Shaw protocol. It is therefore not considered as full ST, more as a form of CBT that uses ST concepts. Comparing the effect sizes of the studies investigating the two forms indicates that group-ST has twice as strong effects on the same measures (SCL-90, YSQ, SMI) as SCBT-g. However, it should be noted that samples were not comparable, as SCBT-g studies recruited patients with mixed diagnoses, including many with subthreshold PD; whereas the group-ST studies only included severe BPD patients.

Limitations of the study include the following. First, we could not empirically assess what inclusion criteria patients have to meet. Based on our clinical impression within this small *sample*, we suggest that antisocial and narcissistic traits do indeed interfere with group work: despite our exclusion of patients with >2 traits of these PDs, we had problems with patients acting out in narcissistic or antisocial ways in the *groups*. These patients had trouble to connect and attach to the group, scared other participants, resulting in a disturbing effect on the group process or dropping out of the therapy group. In addition, patients with a lot of psychosocial turmoil had problems dealing with the topics of the group sessions.

A second limitation is the small sample size, restricting the confidence intervals of our findings. For instance, it is unclear
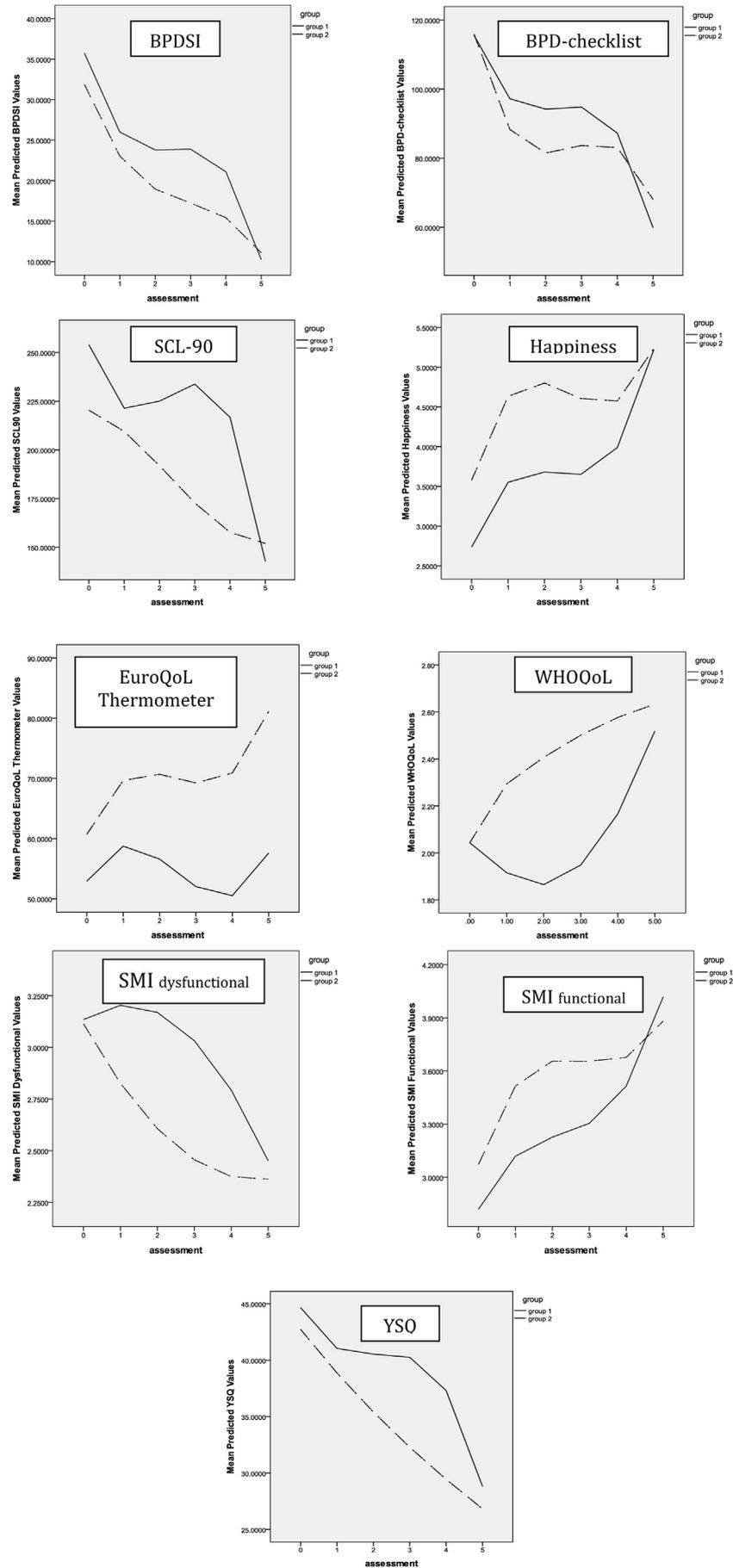
Fig. 3. Estimated means by group and assessment, from the mixed regression analyses. Higher order time trend by group interactions reflect quicker improvement in group 2 in all variables, except for the EuroQoL thermometer, where the linear time effect was stronger in group 2 than in group 1.

whether the higher dropout rates in this study compared to individual ST (40% vs. 27% in the Giesen-Bloo et al., 2006 study) are a chance finding, or inherent to group treatment. Note that it has often been observed that group treatment is associated with a higher dropout rate, for instance because of scheduling problems (MacNair & Corazzini, 1994; Yalom, 1966). A third limitation is that the study was uncontrolled. Future studies are needed to document the effectiveness of group-ST for BPD, and also to compare the two formats that are developed: pure group-ST vs. the combination of group- and individual ST. A large international RCT testing the effects of both formats by comparison to usual treatment is currently executed (Arntz et al., 2013). A fourth limitation is that the explanation of cohort differences by training in a specialized group-ST model is preliminary, as other factors might explain differences as well, e.g., non-specific training effects (any training in a group model might have caused similar effects) and cohort effects (e.g., the first cohort having more difficult patients).

Although group-ST along the lines of Farrell and Shaw's model might indeed catalyze schema change processes, thus speed up recovery and reducing treatment length, there are also limitations to the group format. Scheduling is less flexible compared to individual treatment, thus reducing the number of patients that can participate, and increasing dropout risk. Secondly, only one patient acting in a threatening way might compromise the safety in the group and reduce the effectiveness of the treatment. Thirdly, the closed format is complicated as all patients have to start at the same time and waiting for start can be problematic or create extra care delivery costs. A semi-closed group format could partially solve this problem, but is in itself complicated as safety within the group and attachment among group members can be endangered. A semi-closed format protocol should be developed and tested, as it might be more attractive for clinical practice. Lastly, quite some patients refuse group treatment as they are too distrustful or afraid, though the positive results of studies on group-ST might help them to overcome this resistance.

We learned that for ST for BPD-patients individual therapy can support group therapy, and vice versa. Individual therapists reported that the group offered possibilities for their patients that are otherwise difficult to organize. Both group and individual therapists reported a positive experience of sharing responsibilities for the patient, which reduced the burden of treating BPD patients. It is an empirical issue to what degree pure group-ST is as effective as the combined format, but we hypothesize that the combined format is superior as it offers complimentary situations and meets the needs of patients for individual attention and attachment, as wellbeing a member of a safe group.

An unsolved issue we encountered was how to deal with lack of discipline, which is often a symptom of the pathology. This was seen in no show, being late or refusing to do homework assignments. In therapy we often experienced tension between understanding and accepting the lack of discipline as being part of the BPD on the one hand, and limit setting on the other hand. Limit setting might lead to have the patient stop the group. This might lead to too small groups, though a semi-closed group model can compensate for that.

In conclusion, this study shows that it is possible to treat BPD-outpatients with a combined individual–group-ST format, and it offers a replication of the findings by Farrell et al. (2009). BPD symptomatology decreased in the patients who participated in the study. Effect sizes were high. Dropout risk seems however increased with the group format compared to individual treatment, and might be a disadvantage of group-ST. On the other hand, the positive effects of group-ST create opportunities to make ST accessible for more BPD patients and improve cost-effectiveness. The results suggest that it is important to train the

group therapists in Farrell and Shaw's group-ST model to achieve faster treatment results, though further research is needed to test how specific the effects of this model are. More research is needed to replicate and validate current findings in a larger sample, and to investigate whether a combined group–individual ST format or a pure group-ST format is to be preferred. Future research also needs to identify factors for treatment indication for individual vs. group-ST.

Funding

Declaration of interest

The authors report no conflicts of interest.

References

American Psychiatric Association. (2005). *Diagnostic and statistical manual of mental disorders, text revision* (4th ed.). Washington, DC: American Psychiatric Association.
Arntz, A., Farrell, J. M., Jacob, G., Lee, C. W., Fassbinder, E., Fretwell, H., et al. (2013). Multi-site RCT of group schema therapy for borderline personality disorder. Work in progress Dutch trial register NTR2392 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=2392.
Arntz, A., & van Genderen, H. (2009). *Schema therapy for borderline personality disorder*. Chichester: Wiley.
Arntz, A., Marije van den Hoorn, M., Cornelis, J., Verheul, R., van den Bosch, W. M. C., & de Bie, A. J. H. T. (2003). Reliability and validity of the borderline personality disorder severity index. *Journal of Personality Disorders, 17*, 45–59.
van Asselt, A. D. I., Dirksen, C. D., Arntz, A., Giesen-Bloo, J. H., van Dyck, R., Spinhoven, P., et al. (2008). Outpatient psychotherapy for borderline personality disorder: cost effectiveness of schema-focused therapy versus transference focused psychotherapy. *The British Journal of Psychiatry, 192*, 450–457.
Brooks, R. (1996). EuroQol the current state of play. *Health Policy, 37*, 53–72.
Davidson, K., Norrie, J., Tyrer, P., Gumley, A., Tata, P., Murray, H., et al. (2006). The effectiveness of cognitive behavior therapy for borderline personality disorder: results from the borderline personality disorder study of cognitive therapy (BOSCOT) trial. *Journal of Personality Disorders, 20*, 450–465.
Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: an outpatient psychiatric rating scale: preliminary report. *Psychopharmacology Bulletin, 9*, 13–28.
Farrell, J. M., & Shaw, I. A. (2012). *Group schema therapy for borderline personality disorder: A step-by-step treatment manual with patient workbook*. New York: Wiley.
Farrell, J. M., Shaw, I. A., & Webber, M. A. (2009). A schema-focused approach to group psychotherapy for outpatients with borderline personality disorder: a randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 40*, 317–328.
Feingold, A. (2009). Effect sizes for growth-modelling analysis for controlled trials in the same metric as for classical analysis. *Psychological Methods, 14*, 43–53.
First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *User's guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Press.
First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P), Version 2.0*. New York: Biometrics Research Department, New York State Psychiatric Institute.
Giesen-Bloo, J. H., Arntz, A., & Schouten, E. (2006a). The borderline personality disorder checklist: psychometric evaluation and factorial structure in clinical and nonclinical samples. In J. H. Giesen-Bloo (Ed.), *Crossing borders: Theory, assessment and treatment in borderline personality disorder* (pp. 85–102).
Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., et al. (2006b). Outpatient psychotherapy for borderline personality disorder: randomized trial of schema-focused therapy vs. transference-focused psychotherapy. *Archives of General Psychiatry, 63*, 649–658.
Giesen-Bloo, J. H., Wachters, L. M., Schouten, E., & Arntz, A. (2010). The borderline personality disorder severity index-IV: psychometric evaluation and dimensional structure. *Personality and Individual Differences, 49*, 136–141.
Groenestijn, M. A. C., Akkerhuis, G. W., Kupka, R. W., Schneider, N., & Nolen, W. A. (1999). *Structured Clinical Interview for DSM-IV Axis I Disorders*. Lisse: Swets & Zeitlinger BV.
Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 48*, 1060–1064.

Lobbestael, J., van Vreeswijk, M., & Arntz, A. (2008). An empirical test of schema mode conceptualizations in personality disorders. *Behaviour Research and Therapy, 46*, 854—860.

Lobbestael, J., van Vreeswijk, M., Spinhoven, P., Schouten, E., & Arntz, A. (2010). Reliability and validity of the short Schema Mode Inventory (SMI). *Behavioural and Cognitive Psychotherapy, 38*, 437—458.

MacNair, R. R., & Corazzini, J. G. (1994). Client factors influencing group therapy dropout. *Psychotherapy: Theory, Research, Practice, Training, 31*, 352—362.

Nadort, M., Arntz, A., Smit, J. H., Giesen-Bloo, J., Eikelenboom, M., Spinhoven, P., et al. (2009). Implementation of outpatient schema therapy for borderline personality disorder with versus without crisis support by the therapist outside office hours: a randomized trial. *Behaviour Research and Therapy, 47*, 961—973.

Nordahl, H. M., & Nysaeter, T. E. (2005). Schema therapy for patients with borderline personality disorder: a single case series. *Journal of Behavior Therapy and Experimental Psychiatry, 36*, 254—264.

Renner, F., van Goor, M., Huibers, M., Arntz, A., Butz, B., & Bernstein, D. (2013). Short-term group schema cognitive-behavioral therapy for young adults with personality disorders and personality disorder features: associations with changes in symptomatic distress, schemas, schema modes and coping styles. *Behaviour Research and Therapy, 51*, 487—492. http://dx.doi.org/10.1016/j.brat.2013.05.011.

Rijkeboer, M. M., van den Bergh, H., & van den Bout, J. (2005). Stability and discriminative power of the Young Schema-Questionnaire in a Dutch clinical versus a non-clinical population. *Journal of Behavior Therapy and Experimental Psychiatry, 36*, 129—144.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*, 147—177.

van Vreeswijk, M. F., Spinhoven, P., Eurelings-Bontekoe, E. H. M., & Broersen, J. (2012). Changes in symptom severity, schemas and modes in heterogeneous psychiatric patient groups following short-term schema cognitive—behavioural group therapy: a naturalistic pre-treatment and post-treatment design in an outpatient clinic. *Clinical Psychology and Psychotherapy*. http://dx.doi.org/10.1002/cpp.1813.

Veenhoven, R. (2008). http://worlddatabaseofhappiness.eur.nl/hap_quer/hqs_fp.htm.

Weertman, A., Arntz, A., & Kerkhofs, M. L. M. (2000). *Structured Clinical Interview for DSM-IV, Axis II Personality Disorders*. Lisse: Swets & Zeitlinger BV.

WhOQoL Group: Development of the World Health Organization. (1998). WHO-QOL-BREF quality of life assessment. *Psychological Medicine, 28*, 551—558.

Yalom, I. D. (1966). A study of group therapy dropouts. *Archives of General Psychiatry, 14*, 393—414.

Yalom, I. D., & Leszcz, M. (2005). *The theory and practice of group psychotherapy* (5th ed.). New York: Basic Books.

Yeomans, F. E., Clarkin, J. F., & Kernberg, O. F. (2002). *A primer for transference focused therapy for the borderline patient*. Northvale: Jason Aronson Inc.

Young, J. E., & Brown, G. (1994). Young Schema-Questionnaire (2nd ed.). In J. E. Young (Ed.), *Cognitive therapy for personality disorders: A schema-focused approach* (Rev. ed.) (pp. 63—76). Sarasota, FL: Professional Resource Press.

Young, J. E., Klosko, J., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. New York: Guilford Press.