Group schema therapy for cluster-C personality disorders

A multicentre open pilot study

Wibbelink, C.J.M.; Venhuizen, A.-S.S.M.; Grasman, R.P.P.P.; Bachrach, N.; van den Hengel, C.; Hudepohl, S.; Kunst, L.; de Lange, H.; Louter, M.A.; Matthijssen, S.J.M.A.; Schaling, A.; Walhout, S.; Wichers, K.R.; Arntz, A.

DOI
10.1002/cpp.2903

Publication date
2023

Document Version
Final published version

Published in
Clinical Psychology and Psychotherapy

Received: 2 July 2022 | Revised: 9 August 2023 | Accepted: 14 August 2023

DOI: 10.1002/cpp.2903

RESEARCH ARTICLE

WILEY

Group schema therapy for cluster-C personality disorders: A multicentre open pilot study

Carlijn J. M. Wibbelink[1] | Anne-Sophie S. M. Venhuizen[1] | Raoul P. P. P. Grasman[2] | Nathan Bachrach[3,4] | Charlotte van den Hengel[1] | Sandy Hudepohl[5] | Liselotte Kunst[6,7] | Hinde de Lange[8] | Mark A. Louter[9] | Suzy J. M. A. Matthijssen[10] | Arita Schaling[11] | Simone Walhout[12] | Karen (Renske) Wichers[13] | Arnoud Arntz[1]

[1]Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands

[2]Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands

[3]Department of Medical and Clinical Psychology, Tilburg University, Tilburg, The Netherlands

[4]GGZ Oost Brabant Helmond, Helmond, Netherlands

[5]PsyQ Amsterdam, Amsterdam, The Netherlands

[6]PsyQ Haarlem, Haarlem, The Netherlands

[7]NPI, Amsterdam, The Netherlands

[8]Emergis, Kloetinge, The Netherlands

[9]Viersprong Institute for Studies on Personality Disorders, Rotterdam, The Netherlands

[10]Altrecht Academic Anxiety Centre, Altrecht GGz, Utrecht, The Netherlands

[11]Pro Persona, Ede, The Netherlands

[12]GGZ Oost Brabant Oss, Oss, The Netherlands

[13]IPGGZ, Veendam, The Netherlands

Correspondence
Carlijn J. M. Wibbelink, Department of Clinical Psychology, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WS, Amsterdam, The Netherlands.
Email: c.j.m.wibbelink@uva.nl

Funding information
This study did not receive funding.

Abstract

Background: Group schema therapy (GST) is increasingly popular as a treatment for personality disorders (PDs), including Cluster-C PDs. Individual ST has proven to be effective for Cluster-C PD patients, while the evidence for GST is limited. This study aimed to investigate the effectiveness of GST for Cluster-C PD. Moreover, differences between the specific Cluster-C PDs (avoidant PD, dependent PD and obsessive-compulsive PD) were explored.

Methods: A multicentre open trial was conducted, including 137 patients with a Cluster-C PD (avoidant PD: $n = 107$, dependent PD: $n = 11$ and obsessive-compulsive PD: $n = 19$). Patients received 30 weekly GST sessions with a maximum of 180 min of individual ST and five optional monthly booster sessions. Outcome measures including Cluster-C PD severity, general psychopathological symptoms, quality of life, functional impairment, happiness, PD-related beliefs, self-esteem, self-ideal discrepancy, schemas and schema modes were assessed at baseline until 2-year follow-up with semi-structured interviews and self-report measures. Change over time and differences between the specific Cluster-C PDs were analysed with mixed regression analyses.

Results: The outcome measures showed significant improvements for all Cluster-C PDs, with medium to large effect sizes after 2 years. A treatment dropout rate of 11.7% was found. There were some indications for differences between the Cluster-C PDs in severity at baseline, change trajectories and effectiveness of GST.

Conclusions: This study demonstrated that GST is a promising treatment for Cluster-C PDs. The following step is a randomized controlled trial to further document the (cost-)effectiveness of GST.

KEYWORDS
avoidant, dependent and obsessive-compulsive personality disorders; cluster-C personality disorders; effectiveness; group schema therapy; pilot study

1 | BACKGROUND

Personality disorders (PDs) are serious mental health disorders that make up about a third of diagnoses in psychiatric outpatients (Zimmerman et al., 2005). A large part of these diagnoses represent Cluster-C PDs, including avoidant PD (APD), dependent PD (DPD) and obsessive-compulsive PD (OCPD). Their prevalence in the Western population is 6.7% (Volkert et al., 2018). In clinical samples, the prevalence is estimated at 10.5% for OCPD, 15.0% for DPD and 24.6% for APD (Torgersen, 2012). Cluster-C PDs are characterized by anxiety, avoidance, interpersonal dysfunction and a need for control. Moreover, Cluster-C PDs are associated with severe functional impairment (Skodol et al., 2002), economic inactivity (Coid et al., 2006) and high co-morbidity (Martinussen et al., 2017). In addition, the quality of life of people suffering from a Cluster-C PD is severely impaired (Soeteman et al., 2008). Considering the high prevalence and serious impairments, it is surprising as well as worrying that there is a relative lack of attention for Cluster-C PD in the scientific literature when compared to for instance borderline PD (BPD; Hutsebaut et al., 2018). A meta-analysis on 15 treatment effectiveness studies has shown that cognitive-behavioural and psychodynamic approaches (along with social skills training) were beneficial for Cluster-C PDs, although contradictory results were found and uncertainty remains regarding the relative effectiveness for DPD, APD and OCPD (Simon, 2009). More recent studies exploring new approaches show promising results but point to difficulty comparing established to newer treatments due to differing methodologies (Simonsen et al., 2022). Moreover, studies are often underpowered, focused solely on APD or did not study differences between PDs (Popolo et al., 2022; Simonsen et al., 2022; Skewes et al., 2015). Regarding the differences between the Cluster-C PDs in treatment effectiveness, previous studies are inconclusive; one study found that OCPD patients improved most (Barber et al., 1997), while other studies favoured APD patients (Karterud et al., 2003) or DPD patients (Seivewright et al., 2002). Finally, a necessity for more research into effective treatment for the three Cluster-C PDs is further exemplified by a trial study that found that patients in their sample (90% Cluster-C PD) had 2.33 previous treatments with an average total duration of 2.67 years (Bamelis et al., 2014), suggesting that Cluster-C PDs are not dealt with effectively by existing treatments.

One treatment that was mainly designed to treat PDs such as Cluster-C PD is schema therapy (ST; Young, 1999). ST combines cognitive, behavioural and experiential techniques to target pervasive patterns of emotions, behaviour and bodily sensations that were formed in childhood called early maladaptive schemas (Young et al., 2003). Activation of these schemas together with coping styles can give rise to schema modes, which are temporary mental states that influence a person's cognitions, emotions and behaviour. These schema modes, often functional in childhood, become problematic in later life. The main goal of ST is to reduce maladaptive schemas and schema modes and strengthen functional ones. Most of the evidence for the effectiveness of ST comes from research into BPD, for example by showing superiority of individual ST (IST) to transference-focused therapy (Giesen-Bloo et al., 2006) and treatment as usual (TAU; Nadort et al., 2009). However, evidence also exists for the effectiveness of IST for Cluster-C PDs. One large randomized controlled trial (RCT) found IST to be superior to TAU regarding PD recovery in a Dutch sample consisting of 90% Cluster-C PD patients (Bamelis et al., 2014). Additionally, this study also found superiority with regard to cost-effectiveness for IST (Bamelis et al., 2015).

A form of ST presumed to be even more cost-effective is group ST (GST). In group treatment larger numbers can be treated in less time using fewer resources, posing an attractive solution for long waiting lists in health care. A GST model developed by Farrell and Shaw has become increasingly popular for more severe PD patients (Farrell et al., 2016). Their specific use of group dynamics is hypothesized to 'catalyse' schema change processes. Moreover, group therapy offers a unique opportunity for vicarious learning, practicing new behaviours and gaining peer-support and a sense of belonging. In BPD, GST was proven superior over TAU (Arntz et al., 2022) and resulted in fewer treatment drop-outs (Farrell et al., 2009). However, a large group of patients currently receiving GST in the Netherlands has a primary Cluster-C PD diagnosis. The evidence for GST as a treatment for Cluster-C PDs is limited to impressions derived from clinical practice, preliminary observations of a running RCT comparing GST to group-CBT for Social Anxiety with a co-morbid APD diagnosis (Baljé et al., 2016) and a small pilot study ($N = 8$) for mixed PDs (75% APD; Skewes et al., 2015). This pilot study found strong improvements in APD patients regarding PD severity, depression, and anxiety scores at follow-up 1 year after treatment completion. Moreover, at follow-up five out of six APD patients achieved a loss of PD diagnosis. While these preliminary results and clinical observations may be promising, an adequately sized study into the effectiveness of GST for

Key Practitioner Message

- Group schema therapy is a promising treatment for Cluster-C personality disorder.
- Medium to large treatment effects and a relatively low dropout rate were found.
- Severity before start of the treatment, change trajectories and effectiveness may differ between the specific Cluster-C personality disorders.
- More research is needed to further document the effectiveness of group schema therapy.

Cluster-C PD is needed to substantiate the legitimacy of its wide application in clinical practice beyond (assumed) cost-saving reasons.

The current investigation aimed to elucidate the effects of GST for Cluster-C PD with a multicentre open trial. Moreover, differences between the specific Cluster-C PDs were explored to help generate hypotheses for future research. The primary outcome was change in manifestations of the pertinent Cluster-C PD. Secondary outcomes included general psychopathological symptoms, quality of life, functional impairment, happiness, PD-related beliefs, self-esteem, self-ideal discrepancy, early maladaptive schemas, and functional and dysfunctional schema modes. Moreover, patient use of psychotropic medications and psychological treatments other than GST and serious adverse events were assessed. Based on previous research, we hypothesized that Cluster-C PD patients would improve on primary and secondary outcomes.

2 | METHODS

2.1 | Patients

Participants were eligible to participate in the study if they (1) were between 18 and 70 years old; (2) had a primary diagnosis of APD, DPD or OCPD; and (3) were proficient in Dutch. Patients were excluded if they (1) fulfilled the criteria for a current substance use disorder (participation was possible after 3 months of abstinence), bipolar disorder type I (current or past) and/or psychotic disorder; (2) met four or more criteria of BPD; (3) had acute suicide risk; (4) had an IQ below 80; (5) had received ST in the past year; and (6) were not able to attend the treatment sessions. In addition, patients could not start treatment unless their medication use was stable for at least 3 months or if patients discontinued medication use. Finally, patients were not allowed to start with any form of psychological treatment or medication between the screening period and end of the treatment. Patients were allowed to continue with non-PD-focused supportive treatment during the screening and waitlist period.

2.2 | Procedure

The study was preregistered at the Netherlands Trial Register, part of the Dutch Cochrane Centre (registration number NL5531), and approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences, University of Amsterdam (registration number 2017-CP-7563). Participants were recruited from April 2017 to June 2019 at nine Dutch mental healthcare centres, including Altrecht, Emergis, GGZ Oost Brabant Helmond, GGZ Oost Brabant Oss, IPGGZ, Pro Persona, PsyQ Amsterdam, PsyQ Haarlem, and the Viersprong Institute for Studies on Personality Disorders. After providing written informed consent, screening of potential participants involved the assessment of PDs with the Structural Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) or the Structural Clinical Interview for DSM-5 Personality Disorders (SCID-5-PD) and syndrome disorders with the Structural Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Structural Clinical Interview for DSM-5 Syndrome Disorders (SCID-5-S) or MINI-International Neuropsychiatric Interview (MINI). The choice of these specific instruments was based on site preference and availability. In addition, a screening interview was conducted to assess the participant's availability and motivation. Patients eligible for participation were invited for the baseline assessment, including interviews and computer-based self-report questionnaires. After completing the baseline assessment, patients started the treatment as soon as treatment slots were available. If patients could not start treatment within 2 months after completing the baseline assessment, the outcome instruments were reassessed (waitlist assessment).

After the treatment started, patients were reassessed after 15 sessions of GST (mid-treatment assessment; approximately after 4 months), after 30 sessions of GST (post-treatment assessment; approximately after 8 months), after the booster sessions (post-booster assessment; approximately after 12 months), and 1-year after the end of the treatment (follow-up assessment; approximately after 24 months). All assessments were a combination of computer-based self-report questionnaires and interviews, conducted by an independent research assistant. Therapists and participants were informed about the results.

2.3 | Treatment

2.3.1 | Group schema therapy

The treatment followed the GST for Cluster-C PD protocol developed by Tjoa and Muste (Tjoa & Muste, 2020). Although ST is a transdiagnostic approach, each schema mode requires specific treatment techniques. Therefore, specific ST protocols have been developed based on disorder-specific mode models (Jacob & Arntz, 2013). The protocol by Tjoa and Muste was an adapted version of the GST-protocol used in Baljé et al. (2016) with an emphasis on, but not limited to, schema modes relevant for APD. The treatment comprised 30 weekly 90-min GST sessions, along the Farrell and Shaw model (Farrell et al., 2015), a maximum of 180 min of individual ST by the GST therapist during the group treatment phase, and five optional monthly 1-h group booster sessions. Before start of the group treatment, patients received two individual preparation sessions. In the group treatment, schema modes were dealt with in repeating cycles of 10 sessions. Patients went through this cycle three times (total of 30 sessions), with consecutive levels of depth (i.e., phases). For each patient, the first phase (sessions 1–10) was aimed at creating awareness of different schema modes. In the second phase (sessions 11–20), the patient started practicing new emotional, behavioural and cognitive skills to reduce maladaptive schema modes and strengthening adaptive schema modes. The third phase (sessions 21–30) focused on using the new skills in everyday life. The group had a semi-open format with a maximum capacity of nine patients with mixed primary Cluster-C PD diagnoses and of varying ages. Every 10 sessions, patients who completed 30 sessions left the group and in the following session new patients,

usually three, entered the group. Consequently, a group contained patients of different phases. Techniques, assignments and homework assignments were therefore used in differential forms of depth, depending on the phase of the patient. Individual ST was used to address problems with group participation, such as specific fears, and to conduct individual imagery rescripting. After patients completed the 30 sessions, they had the option to attend five monthly booster sessions in small groups to sustain the treatment effect and enhance generalization. A clinical illustration of the GST protocol is presented in Bachrach and Arntz (2021).

The group was led by two GST therapists. Therapists in this study were licensed psychologists, psychotherapists or psychiatrists who completed basic training in individual ST and a four-day training in the GST model of Farrell and Shaw (Farrell et al., 2015). In addition, all therapists received a one-day training in the protocol developed by Tjoa and Muste (Tjoa & Muste, 2020). During the study, at each site, therapists met every week for 1 h for peer-supervision.

2.3.2 | Other treatment

During treatment, patients were not allowed to start with medication or engage in another form of psychological treatment in addition to the study treatment. However, in case of acute crisis, the crisis procedures of the participating centres were followed (e.g., crisis intervention and medication). Protocol deviations (i.e., another form of psychological treatment and/or starting with medication during the study treatment) were monitored, but, following the intention-to-treat principle, did not result in exclusion from the study treatment and the data-analysis. After the treatment, a 1-year treatment-free period was encouraged.

2.3.3 | Treatment integrity

To assess treatment integrity, 45 randomly selected GST sessions were assessed by four trained blind raters (undergraduate psychology students) using the Treatment Integrity Checklist for Group Schema Therapy (TIC-GST) (Greeven et al., 2017). Of the 45 recordings, 40 recordings were rated by two different raters. Interrater agreement (ICC) ranged between .34 (competence) and .68 (adherence), indicating poor to moderate reliability. An overall treatment integrity of 78% was found (competence: 66% and adherence: 90%), which can be interpreted as moderate to high treatment integrity (Perepletchikova & Kazdin, 2005).

2.4 | Measures

Screening measures included diagnostic instruments and a screening interview. The primary outcome was change in severity of the manifestations of the pertinent Cluster-C PD. For each disorder, a semi-structured severity interview exists, measuring the severity of the DSM-5 criteria of the pertinent PD in the last month. Secondary outcomes included general psychopathological symptoms, quality of life, functional impairment, happiness, PD-related beliefs, self-esteem, self-ideal discrepancy, schemas and schema modes. Moreover, demographics, including patient use of psychotropic medications and psychological treatments other than GST, and serious adverse events were assessed. Appendix A provides details on the instruments used.

2.5 | Statistical analysis

The statistical analysis was performed according to the intention-to-treat principle, using all available data from all patients that started with the treatment. Change in the primary and secondary outcomes and differences in change between the disorder categories (i.e., APD, DPD and OCPD) were analysed using generalized linear mixed regression models. Mixed regression was used to deal with missing data and the nested structure of the data, as the observations were nested within patients, which were nested within subgroups (i.e., patients that entered the treatment group at the same time), which were again nested within treatment groups, and finally, treatment groups were nested within institutions. Skewed distributions were analysed with gamma regression with a log link, using residual pseudo-likelihood (RSPL) as estimation method (Stroup & Claassen, 2020). Restricted maximum likelihood was used as estimation method for Gaussian data. The degrees of freedom were estimated with Satterthwaite's method. To enable gamma regression, when an outcome had zero scores, a small value was added to the scores to bring the minimum to just greater than 0. Moreover, a left skewed distribution was changed to a right skewed distribution by mirroring scores.

The fixed part of the model contained time, diagnosis and the interaction between time and diagnosis. APD was the reference category. The time model was either linear or piecewise. The choice for a linear, two-piecewise or three-piecewise model was based on visual inspection, significance of the interaction effects and likelihood ratio tests (not possible for skewed data). For the repeated part, the appropriate covariance structure (unstructured, toeplitz, heterogenous toeplitz, compound symmetry, heterogenous compound symmetry, and first-order autoregressive moving average) was determined based on the ratio of likelihoods (Gaussian data) or pseudo-likelihoods (skewed data; supported by COVTEST in SAS). If convergence allowed, a random intercept or slope was added for treatment group to account for variability in outcomes due to differences between treatment groups (Feaster et al., 2011). If several random parts could be added to the model, the preferred random part was determined based on fit indices (log likelihood, Akaike's information criterion and Bayesian information criteria). If possible, a random intercept or slope was added for subgroup. Since only one institution included more than one treatment group, no random effect was added for institution.

After the final model was determined, estimated means with 95% confidence intervals and overall time effects (0–24 months) were determined for each disorder (Huang, 2015). Confidence intervals for estimated means of outcomes analysed using gamma regression were obtained by inversely linking the confidence bounds on the linear

scale, and were, therefore, not symmetric (Schabenberger, 2005). Cohen's *d* effect sizes were based on the estimated coefficients, divided by the standard deviation of the baseline scores (Gaussian data) or the square root of the baseline variance in the transformed scale, estimated in a gamma regression model with a fixed intercept and no random parts (skewed data). Residual analyses were performed to check the model assumptions (i.e., normality, homoscedasticity and presence of outliers). Outliers were based on conditional studentized residuals (West et al., 2007), with an absolute value of three or more. In case outliers were present, sensitivity analyses were conducted by rerunning the mixed regression analysis without the outliers. Finally, for patients that completed the 2-year follow-up assessment, reliable change was calculated for the primary outcome based on Jacobson and Truax's formula (1991) by using the standard deviation at baseline and the internal consistency reported in the validation studies (Baljé et al., 2022; Tese, 2019; Verheul et al., 2020).

The analyses were conducted in SAS, version 9.4, using the MIXED procedure for Gaussian data and the GLIMMIX procedure for skewed data. Statistical significance was set at a two-tailed *p*-value of <.05.

3 | RESULTS

3.1 | Patient accrual

The Consolidated Standards of Reporting Trials (CONSORT) diagram of the study is presented in Figure 1. Out of 151 patients referred to the study, 14 patients were not eligible for participation ($n = 4$ did not meet the inclusion criteria, $n = 4$ met the exclusion criteria and $n = 6$ declined participation). Therefore, 137 patients participated in the study. Table 1 shows the baseline characteristics of the patients. Most patients had a primary diagnosis of APD (78.1%), were female (73.0%) and had a Dutch ethnicity (97.1%). In addition, almost half of the patients (49.6%) received benefits (i.e., disability or welfare) before start of treatment. Moreover, a diagnosis of at least one co-morbid syndrome disorder was common among patients (81.0%).

During treatment, 16 (11.7%) patients dropped out of treatment. Another 12 (8.8%) patients did not attend (all) booster sessions. These patients were not counted as 'treatment drop-outs' as attending the booster sessions was optional. During the study period, there was on average 24.1% of missing data on the reassessments (varying between 10.2% at the mid-treatment assessment to 41.6% at the follow-up assessment), because of different reasons (e.g., patient refused participation, patient could not be contacted, and research assistant was not available).

3.2 | Treatment outcomes

3.2.1 | Primary outcome

Results of the primary and secondary outcomes are presented in Table B1 and Table B2 in Appendix B. Figure 2 shows the estimated means of each disorder at each assessment for the primary outcome. A two-piecewise regression model fitted the data best, including time effects for during the treatment (0–8 months) and after the treatment (8–24 months). As severity scores were calculated for each disorder separately and transformed to *Z*-scores before the severity scores were merged, results of the intercept and main effect of diagnosis were meaningless (i.e., estimated intercepts vary around zero) and therefore not reported. During both time periods, the reduction in severity of PD manifestations was significant for APD patients, although the reduction after the treatment period was significantly smaller compared to the reduction during the treatment period. During the treatment period, the reduction in PD severity was smaller in DPD compared to APD, while after the treatment period the reduction was larger. In addition, the reduction after the treatment period was smaller for OCPD compared to APD. However, the effect became a trend effect after outliers were removed. Furthermore, change over time (0–24 months) was examined for all disorders. There was a significant reduction in severity of Cluster-C PD manifestations over time in all disorders, with medium (OCPD: $d = 0.59$) to large (APD: $d = 1.33$; DPD: $d = 1.53$) effect sizes. The reduction in severity was significantly smaller in OCPD compared to APD and DPD. Finally, reliable change was calculated for patients that completed the 2-year follow-up assessment. Among APD patients ($N = 60$), 68.3% improved reliably while 1.7% deteriorated. In the DPD sample ($N = 6$), 83.3% improved and no patients deteriorated. Among the OCPD patients ($N = 11$), 54.5% improved whereas 18.2% showed a deterioration.

3.2.2 | Secondary outcomes

Figure 3 shows the estimated means of each disorder at each assessment for all secondary outcomes. For general psychopathological symptoms, functional impairment, happiness, PD-related beliefs, self-ideal discrepancy, functional schema modes and dysfunctional schema modes, a two-piecewise regression model was fitted including time effects for during the treatment (0–8 months) and after the treatment (8–24 months). A three-piecewise regression model, including time effects for during the treatment (0–8 months), during the optional booster sessions (8–12 months) and after the booster sessions (12–24 months), was analysed for quality of life, self-esteem and early maladaptive schemas.

Baseline differences

At baseline, OCPD scored significantly more positive compared to APD on most secondary outcomes (general psychopathological symptoms, PD-related beliefs, self-ideal discrepancy, early maladaptive schemas, functional schema modes and dysfunctional schema modes; trend effect for functional impairment), suggesting that patients with OCPD showed less severe psychopathology on these outcome measures before start of the treatment.

Change during treatment (0–8 months)

During the treatment period (0–8 months), there was a significant improvement in all secondary outcomes in APD. The improvement

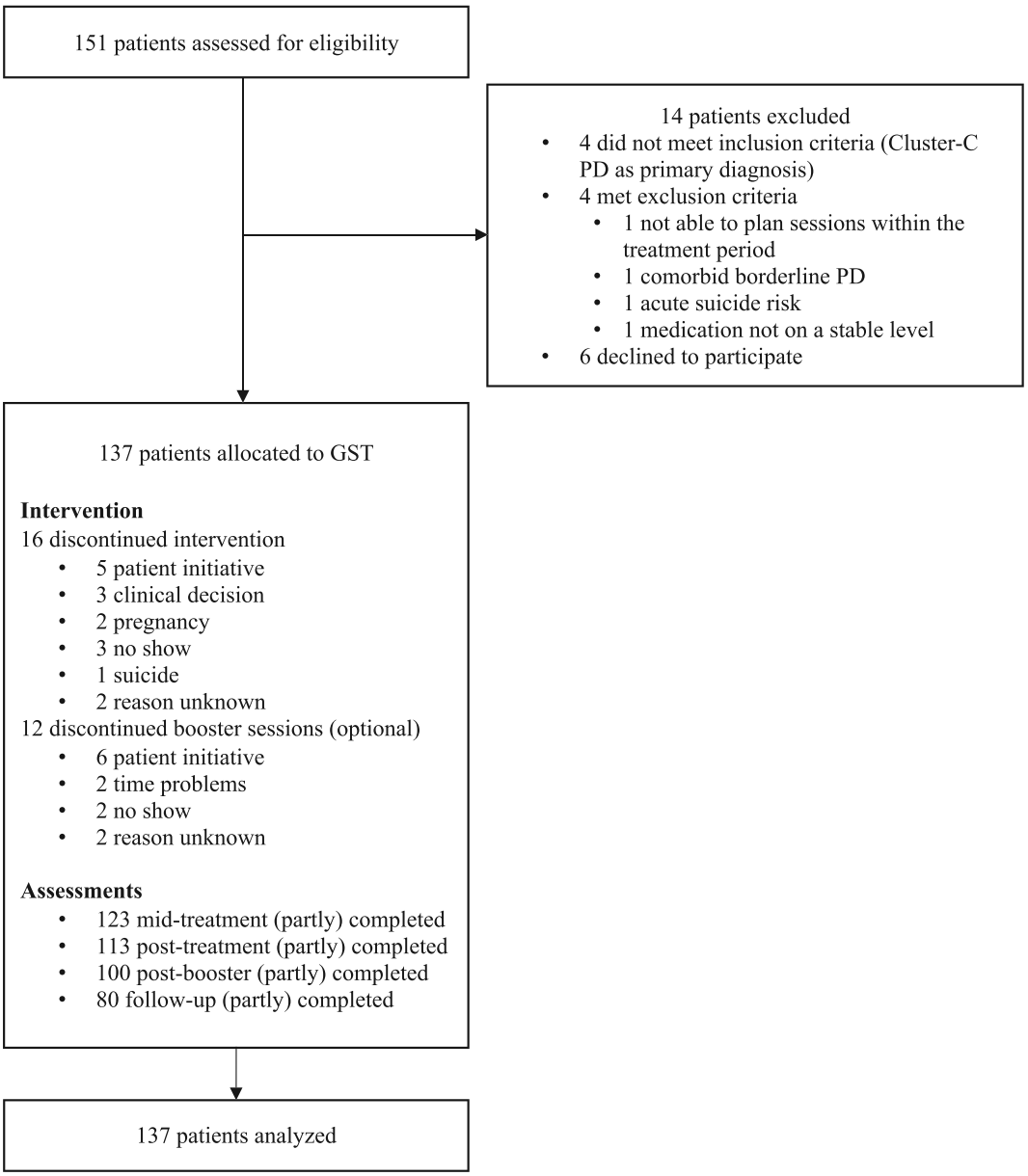

FIGURE 1 CONSORT flowchart of participants.

during the treatment period in DPD was significantly smaller compared to APD for dysfunctional schema modes, with a trend effect for general psychopathological symptoms (significant when outliers were removed). Moreover, the improvement during the treatment period in quality of life was significantly larger for OCPD compared to APD.

Change after treatment (8–24 months) in general psychopathological symptoms, functional impairment, happiness, PD-related beliefs, self-ideal discrepancy, functional schema modes and dysfunctional schema modes

After the treatment period (8–24 months), there was a significant improvement for APD patients in general psychopathological symptoms, functional impairment, PD-related beliefs, self-ideal discrepancy, functional schema modes and dysfunctional schema modes. No significant improvement was found for happiness. For all these outcome measures except one (functional impairment), the improvements after the treatment period were significantly smaller compared to the improvements during the treatment period. Moreover, the improvements after the treatment period were significantly larger for DPD compared to APD in functional impairment and dysfunctional schema modes. The smaller improvement in dysfunctional schema modes during the treatment period in DPD was therefore cancelled out by the larger improvement after the treatment period. Finally, OCPD showed less improvement in dysfunctional schema modes after the treatment period compared to APD.

TABLE 1 Demographic and clinical characteristics of patients at baseline ($N = 137$).

Characteristic	*M*	*SD*
Age	37.18	9.20
	N	*%*
Female	100	73.0
Dutch ethnicity	133	97.1
Employment status		
Housewife	3	2.2
Student	6	4.4
Disability	49	35.8
Welfare	19	13.9
Employed	54	39.4
Other	6	4.4
Education[a]		
Primary	2	1.5
Lower secondary	12	8.8
Upper secondary	61	44.5
Bachelor	32	23.4
Master	24	17.5
Other	6	4.4
Primary personality disorder		
Avoidant	107	78.1
Dependent	11	8.0
Obsessive-compulsive	19	13.9
Co-morbid personality disorder		
Cluster A	2	1.5
Cluster B	3	2.2
Cluster C	15	10.9
Not otherwise specified	3	2.2
None	117	85.4
Co-morbid syndrome disorder		
Anxiety disorder	76	55.1
Depressive disorder	59	42.8
Other	40	29.0
None	26	19.0
Psychotropic medication use	67	48.9
	M	*SD*
Number of previous treatments	2.31	1.25

[a]Based on the International Standard Classification of Education, 2011 version.

Change after treatment (8–12 and 12–24 months) in quality of life, self-esteem and early maladaptive schemas

During the booster sessions (8–12 months) and after the booster sessions (12–24 months), no significant improvements were found in quality of life, self-esteem, and early maladaptive schemas for APD patients. Moreover, the improvements during the booster sessions (for self-esteem and early maladaptive schemas) and after the booster sessions (for quality of life, self-esteem, and early maladaptive schemas) were significantly smaller compared to the improvements during the treatment period. No significant differences were found in changes during versus after the booster sessions. Remarkably, during the booster sessions, the improvements in self-esteem and early maladaptive schemas were significantly smaller in DPD compared to APD, while the improvements were significantly larger after the booster sessions. Consequently, the negative effects during the booster sessions (relapse) were therefore cancelled out by the positive effects after the booster sessions. In addition, the improvement during the booster sessions in quality of life was significantly smaller for OCPD compared to APD. However, as the improvement during the treatment period in quality of life was significantly larger for OCPD compared to APD, the negative effect during the booster sessions did not have a large impact on the final score after 24 months. Finally, for quality of life, the improvement after the booster sessions was significantly larger for DPD compared to APD.

Change during the study period (0–24 months)

For all secondary outcomes, there was a significant improvement over time (0–24 months) in all disorders, except for general psychopathological symptoms in DPD. The effect sizes for APD and DPD were mainly large ($d = 0.88$–1.48, respectively $d = 0.91$–1.99), with one exception for APD ($d = 0.58$ for quality of life) and one exception for DPD ($d = 0.52$ for general psychopathological symptoms). The effect sizes for OCPD were for the most part large ($d = 0.79$–1.27), although the effect sizes for self-esteem, self-ideal discrepancy, early maladaptive schemas and dysfunctional schema modes were medium or medium-to-large ($d = 0.52$–0.73). The improvement over 24 months in functional impairment was significantly larger in DPD ($d = 1.99$) compared to APD ($d = 0.88$) and OCPD ($d = 1.00$). In addition, the improvement in self-ideal discrepancy was significantly smaller in OCPD ($d = 0.52$) compared to APD ($d = 1.18$).

3.2.3 | Other

During the study period, two serious adverse events were reported. One patient died by suicide approximately 3 months after start of the treatment. Another patient was hospitalized because of medical reasons (psychosis) approximately 2 months after start of treatment. In retrospect, for both patients the screening was not performed properly. Suicidality and psychotic symptoms were already present during the referral, and therefore, both patients should not have been included in the study.

Patient use of psychotropic medications across time is shown in Figure 4. Based on visual inspection, the percentage of patients using psychotropic medications did not change substantially over time. In addition, patient use of psychological treatments other than GST after start of treatment is presented in Figure 5. More patients (13.9%) received individual treatment during the follow-up period compared to during the treatment period (4.1%–4.4%) or during the booster sessions (2.0%). Moreover, one fifth of the patients received a form of

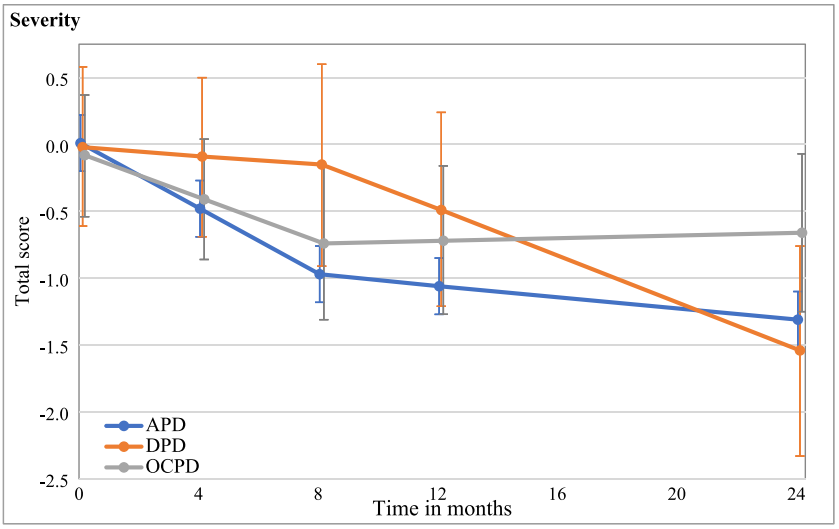

FIGURE 2 Change in primary outcome.

psychological treatment during the follow-up period. Exploratory post-hoc analyses examining the influences of psychotropic medication use and receiving psychological treatment during the follow-up period on change in the primary outcome are reported in Appendix C.

4 | DISCUSSION

This study aimed to investigate the effectiveness of GST for Cluster-C PD. Moreover, differences between the specific Cluster-C PDs (APD, DPD and OCPD) were explored. The findings demonstrate that 30 sessions of GST were effective for all Cluster-C PDs for severity of the manifestations of the pertinent PD as well as for the secondary outcomes, with effect sizes varying from medium to large after 2 years. A treatment dropout rate of 11.7% was observed. In general, there were no differences between the disorders in improvements over the 2-year study period, except for severity of Cluster-C PD manifestations (less improvement in OCPD compared to APD and DPD), self-ideal discrepancy (less improvement in OCPD compared to APD) and functional impairment (more improvement in DPD compared to APD and OCPD). Moreover, compared to APD, OCPD was associated with less severe psychopathology on most outcome measures before start of treatment.

The findings of our multicentre open study demonstrated that it is possible to treat Cluster-C PD patients effectively with GST. Dropout rates in GST (11.7%) were comparable with IST after 30 treatment sessions (11.0%). The effect sizes ranged from medium to large after 2 years, while effect sizes in the studies of Bamelis et al. (2014) and Skewes et al. (2015) were large. However, the effect sizes in the study of Bamelis et al. (2014) were based on a longer treatment and follow-up period and a mixed population (90% Cluster-C PD), and the study of Skewes et al. (2015) included a small sample size ($N = 6$) of only treatment completers, all with APD (whilst OCPD showed the smallest effects in the present study). The findings of the current study are promising, especially since in group treatment more patients can be treated in less therapist time compared to individual treatment. Consequently, this might increase cost-effectiveness and availability of evidence-based treatments for Cluster-C PD. However, it is important to note that one fifth of the patients received (mostly individual) psychological treatment during the follow-up period, suggesting that a relatively short group treatment with limited individual time might be insufficient for some patients. More insight into predictors of treatment response is therefore needed, which will be the focus of a separate paper into the identification of patient characteristics (e.g., childhood maltreatment, autistic traits and demographics) that predict response to GST for Cluster-C PD.

Based on the findings, GST may be most effective for APD and DPD, as OCPD patients showed less improvement in severity of PD manifestations and self-ideal discrepancy. Moreover, relatively less OCPD patients showed a reliable improvement in PD severity compared to APD and DPD patients. However, outcomes on Cluster-C PD severity cannot be compared directly between the Cluster-C PDs because different severity interviews were used. It is therefore uncertain whether the differences in findings on Cluster-C PD severity reflect differences between instruments or that OCPD patients actually benefited less from GST. On the other hand, on seven of the 11 outcome measures, OCPD has the lowest effect sizes, suggesting that the tendency of smaller effects in OCPD is general. One possible explanation for this might lie in the GST-protocol by Tjoa and Muste (Tjoa & Muste, 2020). This protocol was an adapted version of the GST-protocol used by Baljé et al. (2016), which was originally developed for APD. Specific issues related to OCPD might have therefore been insufficiently elaborated on in this version of the protocol. The present findings led to changes in the protocol, which now contains elements that specifically address OCPD-relevant schema modes (Tjoa & Muste, 2021). Another possible explanation could be that OCPD patients showed less severe psychopathology on most outcome measures before start of the treatment compared to APD, giving less room for improvement. In previous studies, OCPD was also found to be associated with less impairment, higher functioning and fewer co-morbid disorders (Cramer et al., 2006; Morey et al., 2002; Skodol et al., 2002). Moreover, OCPD is sometimes viewed as a

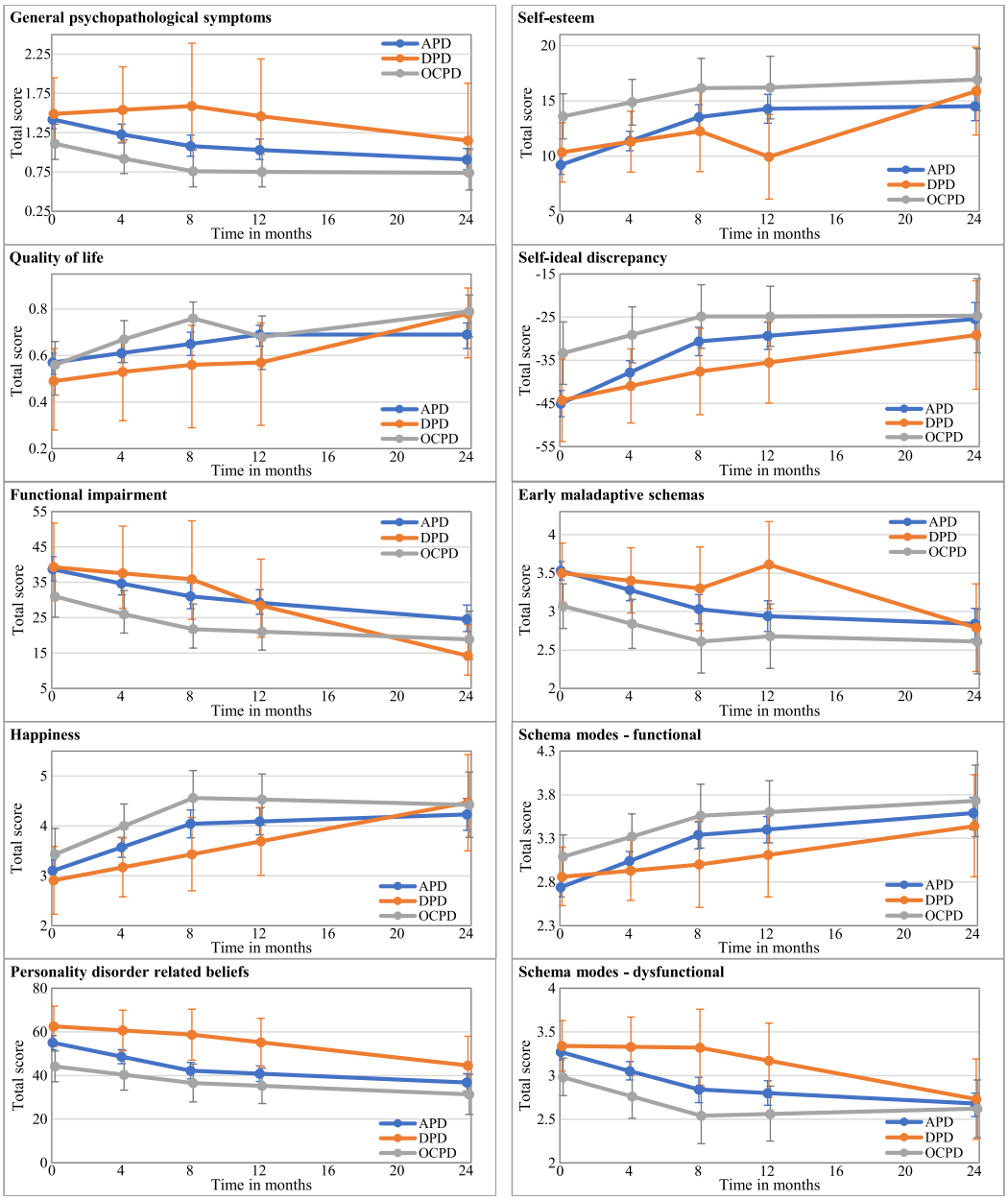

FIGURE 3 Change in secondary outcomes. [Corrections added on 16 September 2023, after first online publication: The top left panel of Figure 3 has been corrected in this version.]

disorder of mild or moderate severity (Wilberg et al., 2009) and some authors even question whether all these patients should receive a PD diagnosis (Skodol et al., 2002). However, in this study well-being (i.e., happiness and quality of life) was equally impaired, suggesting, in accordance with Skodol et al. (2002), that in specific areas OCPD patients are just as impaired as other Cluster-C PD patients.

The overall pattern of change in our study was similar to other treatment studies (Simon, 2009). Most improvements occurred during the treatment, while after the treatment (i.e., during the booster sessions and follow-up period) the improvements continued but at a slower speed. In addition, there were indications that the change pattern of DPD differed from APD. Compared to APD, DPD patients showed less improvements in self-esteem, early maladaptive schemas, dysfunctional schema modes, and general psychopathological symptoms (trend effect) during the first year, with even a relapse after treatment ended (i.e., during the booster sessions) in self-esteem and early maladaptive schemas, while the improvements were larger in the second year. This deviating change pattern in DPD might be related to difficulties with terminating the treatment, caused by their dependency (Bornstein, 2007; Disney, 2013). The fact that they improved after treatment termination, despite initially showing little effects, is however promising. This suggests that treatment should be time-limited for DPD, allowing patients to discover that (after a sufficient dose of therapy) they are capable of successfully functioning without therapy. However, as the current study merely explored differences in treatment effectiveness meant for hypothesis generation rather than

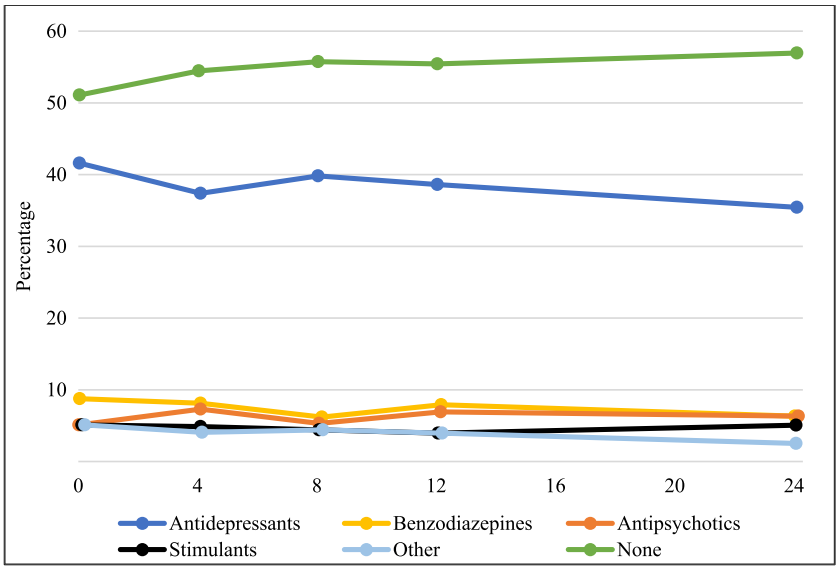

FIGURE 4 Patient use of psychotropic medications across time. *Note*: Other psychotropic medications include for example antiepileptics, antihistamine, lithium, and pregabalin.

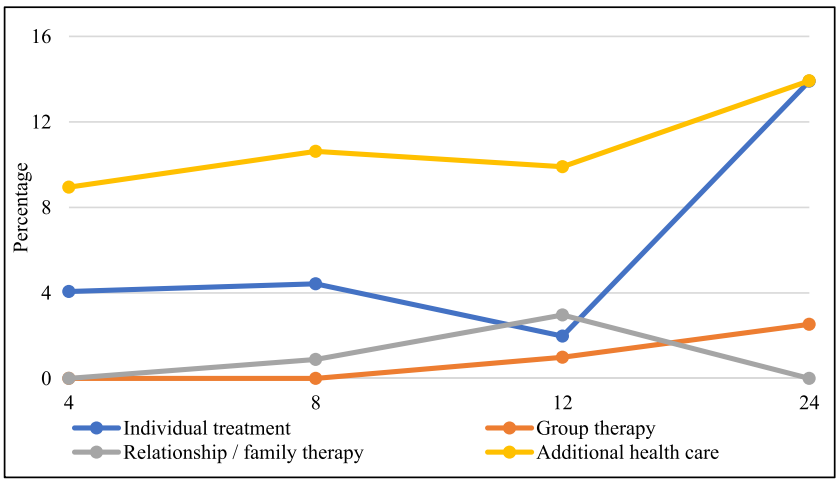

FIGURE 5 Patient use of psychological treatments other than GST. *Note*: Additional health care includes for example psychiatric consultation for medication management, haptotherapy, consultation with the company doctor, and coaching.

confirmation, these findings should be interpreted with caution. In accordance with the meta-analysis of Simon (2009) on psychotherapy outcomes of Cluster-C PD patients, we underline the necessity of more research into the effectiveness of treatments for the specific Cluster-C PDs.

This study has several limitations. First, this study was an open multicentre study without a control group. It was therefore not possible to determine the extent to which the improvements were due to treatment in general or GST specifically or to other factors (e.g., attention and maturation). Moreover, patients were not randomized to treatment, but referral to the treatment was based on indication by the clinician. It is therefore uncertain whether all eligible patients were referred to the study, which may jeopardize generalization of the results. Second, despite the known questionable validity of competence ratings by nonexpert (novice) judges (Waltz et al., 1993; Weck et al., 2011), competence as well as adherence was rated by undergraduate students. Indeed, a low interrater agreement was found for competence, suggesting that the competence rating should be interpreted with caution. Third, the overall sample size of Cluster-C PD patients was sufficient, however DPD and OCPD patients were underrepresented (respectively 8.0% and 13.9%). Although the overrepresentation of APD patients reflects clinical practice and previous studies (Bamelis et al., 2014; Torgersen, 2012), the small sample sizes of DPD and OCPD patients restrict conclusions that can be drawn related to the effectiveness of GST for DPD and OCPD given the limited power. Finally, there was a considerable amount of missing data (41.6%) on the follow-up assessment, which threatens the reliability of the findings at follow-up.

In conclusion, this study demonstrated that GST is a promising treatment for Cluster-C PD. The next step is a large RCT to further

document the (cost-)effectiveness of GST compared to IST and TAU as well as to gain more insight into optimal matching of Cluster-C PD patients to treatment. This RCT is currently being executed (Groot et al., 2022). More research into treatment outcomes of Cluster-C PD patients is warranted, especially since Cluster-C PD has received relatively limited attention in the scientific field (Hutsebaut et al., 2018). To our knowledge, the present study is the first multicentre study into the effectiveness of GST with a large sample of Cluster-C PD patients. As such, this study makes an important contribution in developing evidence-based treatments for Cluster-C PD.

AUTHOR CONTRIBUTIONS

Wrote the manuscript; involved in the implementation and coordination of the data collection: Carlijn J. M. Wibbelink. *Wrote the manuscript; involved in the coordination of the data collection*: Anne-Sophie S. M. Venhuizen. *Principal investigator; initial conception and design of the study*: Arnoud Arntz. *Statistical counselling*: Raoul P. P. P. Grasman. *Involved in the coordination of the data collection*: Charlotte van den Hengel. *Responsible for the recruitment of participants and data collection in their mental healthcare centre*: Nathan Bachrach, Sandy Hudepohl, Hinde de Lange, Mark A. Louter, Suzy J. M. A. Matthijssen, Arita Schaling, Simone Walhout, and Karen (Renske) Wichers. *Read, contributed to, and approved the final manuscript*: All authors.

ACKNOWLEDGMENTS

We would like to thank all patients, therapists, research assistants and students for their involvement in the study.

CONFLICT OF INTEREST STATEMENT

The authors declared no conflicts of interest.

DATA AVAILABILITY STATEMENT

As the participant-level dataset could contain information that compromises the anonymity of participants, this will not be made publicly available. The data that support the findings of this study are available from the corresponding author (CJMW) upon reasonable request.

ETHICS STATEMENT

The study was approved by the Ethics Review Board of the Faculty of Social and Behavioural Sciences, University of Amsterdam (registration number 2017-CP-7563). We obtained signed informed consents from all participants in the study. The study was preregistered at the Netherlands Trial Register, part of the Dutch Cochrane Center (registration number NL5531).

ORCID

Carlijn J. M. Wibbelink https://orcid.org/0000-0001-6002-6892
Nathan Bachrach https://orcid.org/0000-0002-0746-2692

REFERENCES

Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item scale. *Social Behaviour and Personality*, *34*(2), 139–150. https://doi.org/10.2224/sbp.2006.34.2.139

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief 'red flags' for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(2), 202–212.e7. https://doi.org/10.1016/j.jaac.2011.11.003

Arntz, A., Dreessen, L., Schouten, E., & Weertman, A. (2004). Beliefs in personality disorders: A test with the personality disorder belief questionnaire. *Behaviour Research and Therapy*, *42*(10), 1215–1225. https://doi.org/10.1016/j.brat.2003.08.004

Arntz, A., Jacob, G. A., Lee, C. W., Brand-de Wilde, O. M., Fassbinder, E., Harper, R. P., Lavender, A., Lockwood, G., Malogiannis, I. A., Ruths, F. A., Schweiger, U., Shaw, I. A., Zarbock, G., & Farrell, J. M. (2022). Effectiveness of predominantly group schema therapy and combined individual and group schema therapy for borderline personality disorder. *JAMA Psychiatry*, *79*(4), 287–299. https://doi.org/10.1001/jamapsychiatry.2022.0010

Arntz, A., Kamphuis, J. H., & Derks, J. L. (2018). *Gestructureerd klinisch interview voor DSM-5 Syndroomstoornissen Vragenlijst (SCID-5-SV)*.

Bachrach, N., & Arntz, A. (2021). Group schema therapy for patients with cluster-C personality disorders: A case study on avoidant personality disorder. *Journal of Clinical Psychology*, *77*(5), 1233–1248. https://doi.org/10.1002/jclp.23118

Baljé, A., Greeven, A., van Giezen, A., Korrelboom, K., Arntz, A., & Spinhoven, P. (2016). Group schema therapy versus group cognitive behavioral therapy for social anxiety disorder with comorbid avoidant personality disorder: Study protocol for a randomized controlled trial. *Trials*, *17*(1), 487. https://doi.org/10.1186/s13063-016-1605-9

Baljé, A. E., Karch, J. D., Greeven, A., van Giezen, A. E., Muste, E. H., Arntz, A., & Spinhoven, P. (2022). Avoidant personality disorder severity index: Dimensional structure and psychometric properties [Unpublished manuscript]. Department of Anxiety, PsyQ.

Bamelis, L. L. M., Arntz, A., Wetzelaer, P., Verdoorn, R., & Evers, S. M. A. A. (2015). Economic evaluation of schema therapy and clarification-oriented psychotherapy for personality disorders. *The Journal of Clinical Psychiatry*, *76*(11), e1432–e1440. https://doi.org/10.4088/JCP.14m09412

Bamelis, L. L. M., Evers, S. M. A. A., Spinhoven, P., & Arntz, A. (2014). Results of a multicenter randomized controlled trial of the clinical effectiveness of schema therapy for personality disorders. *American Journal of Psychiatry*, *171*(3), 305–322. https://doi.org/10.1176/appi.ajp.2013.12040518

Bamelis, L. L. M., Renner, F., Heidkamp, D., & Arntz, A. (2011). Extended schema mode conceptualizations for specific personality disorders: An empirical study. *Journal of Personality Disorders*, *25*(1), 41–58. https://doi.org/10.1521/pedi.2011.25.1.41

Barber, J. P., Morse, J. Q., Krakauer, I. D., Chittams, J., & Crits-Christoph, K. (1997). Change in obsessive-compulsive and avoidant personality disorders following time-limited supportive-expressive therapy. *Psychotherapy*, *34*(2), 133–143. https://doi.org/10.1037/h0087774

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Handelsman, L., Medrano, M., Desmond, D., & Zule, W. (2003). Development and validation of a brief screening version of the childhood trauma questionnaire. *Child Abuse & Neglect*, *27*(2), 169–190. https://doi.org/10.1016/S0145-2134(02)00541-0

Bornstein, R. F. (2007). Dependent Personality Disorder. In *Personality disorders: Toward the DSM-V* (pp. 307–324). SAGE Publications, Inc.. https://doi.org/10.4135/9781483328980.n11

Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry*, *188*(5), 423–431. https://doi.org/10.1192/bjp.188.5.423

Cramer, V., Torgersen, S., & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive Psychiatry*, *47*(3), 178–184. https://doi.org/10.1016/j.comppsych.2005.06.002

de Beurs, E. (2011). Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011 [Brief Symptom Inventory and BSI 18. Manual 2011]. PITS.

de Beurs, E., & Zitman, F. (2006). De brief symptom inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief [reliability and validity of a practical alternative to the SCL 90]. *Maandblad Geestelijke Volksgezondheid*, *61*, 120–141.

Derogatis, L. R., Lipman, R., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale--preliminary report. *Psychopharmacology Bulletin*, *9*(1), 13–28.

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: An introductory report. *Psychological Medicine*, *13*(3), 595–605. https://doi.org/10.1017/S0033291700048017

Dickhaut, V., & Arntz, A. (2014). Combined group and individual schema therapy for borderline personality disorder: A pilot study. *Journal of Behaviour Therapy and Experimental Psychiatry*, *45*(2), 242–251. https://doi.org/10.1016/j.jbtep.2013.11.004

Disney, K. L. (2013). Dependent personality disorder: A critical review. *Clinical Psychology Review*, *33*(8), 1184–1196. https://doi.org/10.1016/j.cpr.2013.10.001

Dreessen, L., & Arntz, A. (1995). The personality disorder beliefs questionnaire (short version).

Farrell, J., Reiss, N., & Shaw, I. (2016). *Schematherapie in de klinische praktijk. Een complete gids voor individuele, groeps-en geïntegreerde behandeling met schemamodi*. Uitgeverij Nieuwezijds.

Farrell, J. M., Reiss, N., & Shaw, I. A. (2015). *Schematherapie in de klinische praktijk*. Nieuwezijds.

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*(2), 317–328. https://doi.org/10.1016/j.jbtep.2009.01.002

Feaster, D. J., Mikulich-Gilbertson, S., & Brincks, A. M. (2011). Modeling site effects in the design and analysis of multi-site trials. *The American Journal of Drug and Alcohol Abuse*, *37*(5), 383–391. https://doi.org/10.3109/00952990.2011.600386

First, M. B., Gibbon, M., Spitzer, R. L., Benjamin, L. S., & Williams, J. B. W. (1997). *Structured clinical interview for DSM-IV Axis II personality disorders (SCID-II)*. American Psychiatric Press.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). The structured clinical interview for DSM-III-R personality disorders (SCID-II): I. description. *Journal of Personality Disorders*, *9*(2), 83–91. https://doi.org/10.1521/pedi.1995.9.2.83

First, M. B., Spitzer, R. L., Miriam, G., & Williams, J. B. (1996). *Structured clinical interview for DSM-IV Axis I disorders*. American Psychiatric Press, Inc.

First, M. B., Williams, J. B., Benjamin, L., & Spitzer, R. L. (2015). *User's guide for the SCID-5-PD (structured clinical interview for DSM-5 personality disorder)*. American Psychiatric Association.

First, M. B., Williams, J. B., Benjamin, L., & Spitzer, R. L. (2016). *Structured clinical interview for DSM-5 screening personality questionnaire (SCID-5-SPQ)*. American Psychiatric Association.

Franck, E., de Raedt, R., Barbez, C., & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg self-esteem scale. *Psychologica Belgica*, *48*(1), 25–35. https://doi.org/10.5334/pb-48-1-25

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., Kremers, I., Nadort, M., & Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder. *Archives of General Psychiatry*, *63*(6), 649. https://doi.org/10.1001/archpsyc.63.6.649

Greeven, A., Hijne, K., Baljé, A., & Tjoa, E. (2017). Treatment integrity checklist for group schematherapy (TIC-GST).

Groot, I. Z., Venhuizen, A.-S. S. M., Bachrach, N., Walhout, S., de Moor, B., Nikkels, K., Dalmeijer, S., Maarschalkerweerd, M., van Aalderen, J. R., de Lange, H., Wichers, R., Hollander, A. Ph., Evers, S. M. A. A., Grasman, R. P. P. P., & Arntz, A. (2022). Design of an RCT on cost-effectiveness of group schema therapy versus individual schema therapy for patients with Cluster-C personality disorder: The QUEST-CLC study protocol. *BMC Psychiatry*, *22*(1), 637. https://doi.org/10.1186/s12888-022-04248-9

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, *20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Huang, Q. (2015). Hands-on tutorial for piecewise linear mixed-effects models using SAS® PROC MIXED. https://api.semanticscholar.org/CorpusID:63445181

Hutsebaut, J., Willemsen, E. M. C., & Vani, H. L. (2018). Time for cluster C personality disorders: State of the art. *Tijdschrift voor Psychiatrie*, *60*(5), 306–314.

Jacob, G. A., & Arntz, A. (2013). Schema therapy for personality disorders-a review. *International Journal of Cognitive Therapy*, *6*(2), 171–185. https://doi.org/10.1521/ijct.2013.6.2.171

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Janssen, M. F., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P., & Busschbach, J. (2013). Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: A multi-country study. *Quality of Life Research*, *22*(7), 1717–1727. https://doi.org/10.1007/s11136-012-0322-4

Karos, K., Niederstrasser, N., Abidi, L., Bernstein, D. P., & Bader, K. (2014). Factor structure, reliability, and known groups validity of the German version of the childhood trauma questionnaire (short-form) in Swiss patients and nonpatients. *Journal of Child Sexual Abuse*, *23*(4), 418–430. https://doi.org/10.1080/10538712.2014.896840

Karterud, S., Pedersen, G., Bjordal, E., Brabrand, J., Friis, S., Haaseth, Ø., Haavaldsen, G., Irion, T., Leirvåg, H., Tørum, E., & Urnes, Ø. (2003). Day treatment of patients with personality disorders: Experiences from a Norwegian treatment research network. *Journal of Personality Disorders*, *17*(3), 243–262. https://doi.org/10.1521/pedi.17.3.243.22151

Lee, S. J., Choi, Y. H., Rim, H. D., Won, S. H., & Lee, D. W. (2015). Reliability and validity of the Korean young schema questionnaire-short form-3 in medical students. *Psychiatry Investigation*, *12*(3), 295–304. https://doi.org/10.4306/pi.2015.12.3.295

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). *Clinical Psychology and Psychotherapy*, *18*(1), 75–79. https://doi.org/10.1002/cpp.693

Lundin, A., Kosidou, K., & Dalman, C. (2019). Measuring autism traits in the adult general population with the brief autism-Spectrum quotient, AQ-10: Findings from the Stockholm public health cohort. *Journal of Autism and Developmental Disorders*, *49*(2), 773–780. https://doi.org/10.1007/s10803-018-3749-9

Maffei, C., Fossati, A., Agostoni, I., Barraco, A., Bagnato, M., Deborah, D., Namia, C., Novella, L., & Petrachi, M. (1997). Interrater reliability and internal consistency of the structured clinical interview for DSM-IV Axis II personality disorders (SCID-II), version 2.0. *Journal of Personality Disorders*, *11*(3), 279–284. https://doi.org/10.1521/pedi.1997.11.3.279

Martinussen, M., Friborg, O., Schmierer, P., Kaiser, S., Øvergård, K. T., Neunhoeffer, A.-L., Martinsen, E. W., & Rosenvinge, J. H. (2017). The comorbidity of personality disorders in eating disorders: A meta-analysis. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *22*(2), 201–209. https://doi.org/10.1007/s40519-016-0345-x

Miskimins, R. W., & Braucht, G. N. (1971). *Description of the self*. Rocky Mountain Behavioural Science Institute.

Mohammadkhani, P., Forouzan, A. S., Hooshyari, Z., & Abasi, I. (2020). Psychometric properties of Persian version of structured clinical interview for DSM-5-research version (SCID-5-RV): A diagnostic accuracy study.

Iranian Journal of Psychiatry and Behavioural Sciences, *14*(2), e100930. https://doi.org/10.5812/ijpbs.100930

Molendijk, M. L., Bamelis, L., van Emmerik, A. A. P., Arntz, A., Haringsma, R., & Spinhoven, P. (2010). Word use of outpatients with a personality disorder and concurrent or previous major depressive disorder. *Behaviour Research and Therapy*, *48*(1), 44–51. https://doi.org/10.1016/j.brat.2009.09.007

Morey, L. C., Gunderson, J. G., Quigley, B. D., Shea, M. T., Skodol, A. E., Mcglashan, T. H., Stout, R. L., & Zanarini, M. C. (2002). The representation of borderline, avoidant, obsessive-compulsive, and schizotypal personality disorders by the five-factor model. *Journal of Personality Disorders*, *16*(3), 215–234. https://doi.org/10.1521/pedi.16.3.215.22541

Nadort, M., Arntz, A., Smit, J. H., Giesen-Bloo, J., Eikelenboom, M., Spinhoven, P., van Asselt, T., Wensing, M., & van Dyck, R. (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*(11), 961–973. https://doi.org/10.1016/j.brat.2009.07.013

Osório, F. L., Loureiro, S. R., Hallak, J. E. C., Machado-de-Sousa, J. P., Ushirohira, J. M., Baes, C. V. W., Apolinario, T. D., Donadon, M. F., Bolsoni, L. M., Guimarães, T., Fracon, V. S., Silva-Rodrigues, A. P. C., Pizeta, F. A., Souza, R. M., Sanches, R. F., dos Santos, R. G., Martin-Santos, R., & Crippa, J. A. S. (2019). Clinical validity and intrarater and test–retest reliability of the structured clinical interview for DSM-5 – Clinician version (SCID-5-CV). *Psychiatry and Clinical Neurosciences*, *73*(12), 754–760. https://doi.org/10.1111/pcn.12931

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice*, *12*(4). https://doi.org/10.1093/clipsy/bpi045

Popolo, R., MacBeth, A., Lazzerini, L., Brunello, S., Venturelli, G., Rebecchi, D., Morales, M. F., & Dimaggio, G. (2022). Metacognitive interpersonal therapy in group versus TAU + waiting list for young adults with personality disorders: Randomized clinical trial. *Personality Disorders: Theory, Research, and Treatment*, *13*(6), 619–628. https://doi.org/10.1037/per0000497

Rijkeboer, M. M. (2013). *Young schema questionnaire - revised*.

Roelofs, J., Papageorgiou, C., Gerber, R. D., Huibers, M., Peeters, F., & Arntz, A. (2007). On the links between self-discrepancies, rumination, metacognitions, and symptoms of depression in undergraduates. *Behaviour Research and Therapy*, *45*(6), 1295–1305. https://doi.org/10.1016/j.brat.2006.10.005

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Schabenberger, O. (2005). Introducing the GLIMMIX procedure for generalized linear mixed models. Proceedings of the Thirtieth Annual SAS users Group International Conference.

Seivewright, H., Tyrer, P., & Johnson, T. (2002). Change in personality status in neurotic disorders. *Lancet*, *359*(9325), 2253–2254. https://doi.org/10.1016/S0140-6736(02)09266-8

Shankman, S. A., Funkhouser, C. J., Klein, D. N., Davila, J., Lerner, D., & Hee, D. (2018). Reliability and validity of severity dimensions of psychopathology assessed using the structured clinical interview for DSM-5 (SCID). *International Journal of Methods in Psychiatric Research*, *27*(1), 1–12. https://doi.org/10.1002/mpr.1590

Sheehan, D. V., Lecrubier, Y., Harnett Sheehan, K., Janavs, J., Weiller, E., Keskiner, A., Schinka, J., Knapp, E., Sheehan, M. F., & Dunbar, G. C. (1997). The validity of the MINI international neuropsychiatric interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*, *12*(5), 232–241. https://doi.org/10.1016/S0924-9338(97)83297-X

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The MINI-international neuropsychiatric interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*, 22–33.

Simon, W. (2009). Follow-up psychotherapy outcome of patients with dependent, avoidant and obsessive-compulsive personality disorders: A meta-analytic review. *International Journal of Psychiatry in Clinical Practice*, *13*(2), 153–165. https://doi.org/10.1080/13651500802570972

Simonsen, S., Popolo, R., Juul, S., Frandsen, F. W., Sørensen, P., & Dimaggio, G. (2022). Treating avoidant personality disorder with combined individual metacognitive interpersonal therapy and group mentalization-based treatment. *Journal of Nervous & Mental Disease*, *210*(3), 163–171. https://doi.org/10.1097/NMD.0000000000001432

Skewes, S. A., Samson, R. A., Simpson, S. G., & van Vreeswijk, M. (2015). Short-term group schema therapy for mixed personality disorders: A pilot study. *Frontiers in Psychology*, *5*, 1592. https://doi.org/10.3389/fpsyg.2014.01592

Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., Grilo, C. M., Shea, M. T., Zanarini, M. C., Morey, L. C., Sanislow, C. A., & Oldham, J. M. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry*, *159*(2), 276–283. https://doi.org/10.1176/appi.ajp.159.2.276

Soeteman, D. I., Verheul, R., & Busschbach, J. J. V. (2008). The burden of disease in personality disorders: Diagnosis-specific quality of life. *Journal of Personality Disorders*, *22*(3), 259–268. https://doi.org/10.1521/pedi.2008.22.3.259

Somma, A., Borroni, S., Maffei, C., Besson, E., Garbini, A., Granozio, S., Limuti, B., Perego, G., Pietrobon, A., Rugi, C., Turano, E., & Fossati, A. (2017). Inter-rater reliability of the Italian translation of the structured clinical interview for DSM-5 personality disorders (SCID-5-PD): A study on consecutively admitted clinical adult participants. *Journal of Psychopathology*, *23*(3), 105–111.

Stroup, W., & Claassen, E. (2020). Pseudo-likelihood or quadrature? What we thought we knew, what we think we know, and what we are still trying to Figure out. *Journal of Agricultural, Biological, and Environmental Statistics*, *25*(4), 639–656. https://doi.org/10.1007/s13253-020-00402-6

Tese, S. (2019). The reliability and validity of the dependent personality disorder severity index [Unpublished manuscript]. Rino Groep Utrecht.

Thombs, B. D., Bernstein, D. P., Lobbestael, J., & Arntz, A. (2009). A validation study of the Dutch childhood trauma questionnaire-short form: Factor structure, reliability, and known-groups validity. *Child Abuse and Neglect*, *33*(8), 518–523. https://doi.org/10.1016/j.chiabu.2009.03.001

Tjoa, E. E. M. L., & Muste, E. H. (2020). GST CL-C manual 30 session protocol: Therapist manual.

Tjoa, E. E. M. L., & Muste, E. H. (2021). *Handleiding groepsschematherapie cluster C-persoonlijkheidsstoornissen*. Bohn Staleu van Loghum.

Torgersen, S. (2012). Epidemiology. In T. A. Widiger (Ed.), *The Oxford handbook of personality disorders* (pp. 185–205). Oxford University Press. 10.1093/oxfordhb/9780199735013.013.0009

Üstün, T., Kostanjsek, N., Chatterji, S., & Rehm, J. (2010). *Measuring health and disability: Manual for WHO disability assessment schedule WHODAS 2.0*. World Health Organization.

Verheul, H., Arntz, A., van Empelen, C., Debets, C., & de Vries, E. (2020). Psychometric evaluation of the obsessive compulsive personality disorder severity index [Unpublished manuscript]. Rino Groep Utrecht.

Versteegh, M. M., Vermeulen, K. M., Evers, S. M., De Wit, G. A., Prenger, R., & Stolk, E. A. (2016). Dutch tariff for the five-level version of EQ-5D. *Value in Health*, *19*(4), 343–352. https://doi.org/10.1016/j.jval.2016.01.003

Volkert, J., Gablonski, T.-C., & Rabung, S. (2018). Prevalence of personality disorders in the general adult population in Western countries: Systematic review and meta-analysis. *The British Journal of Psychiatry*, *213*(6), 709–715. https://doi.org/10.1192/bjp.2018.202

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and

competence. *Journal of Consulting and Clinical Psychology*, *61*(4), 620–630. https://doi.org/10.1037/0022-006X.61.4.620

Weck, F., Hilling, C., Schermelleh-Engel, K., Rudari, V., & Stangier, U. (2011). Reliability of adherence and competence assessment in cognitive behavioral therapy: Influence of clinical experience. *Journal of Nervous and Mental Disease*, *199*(4), 276–279. https://doi.org/10.1097/NMD.0b013e3182124617

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. Chapman & Hall/CRC.

Wilberg, T., Karterud, S., Pedersen, G., & Urnes, Ø. (2009). The impact of avoidant personality disorder on psychosocial impairment is substantial. *Nordic Journal of Psychiatry*, *63*(5), 390–396. https://doi.org/10.1080/08039480902831322

Young, J. (1999). *Cognitive therapy for personality disorders: A schema-focused approach*. Professional Resource Press/Professional Resource Exchange.

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

Young, J. E. (2005). Young Schema Questionnaire – Short Form 3 (YSQ-S3). Schema Therapy Institute.

Young, J. E., Arntz, A., Atkinson, T., Lobbestael, J., Weishaar, M., van Vreeswijk, M. F., & Klokman, J. (2007). *The schema mode inventory (SMI)*. Schema Therapy Institute.

Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., Morey, L. C., Grilo, C. M., Shea, M. T., McGlashan, T. H., & Gunderson, J. G. (2000). The collaborative longitudinal personality disorders study: Reliability of axis I and II diagnoses. *Journal of Personality Disorders*, *14*(4), 291–299. https://doi.org/10.1521/pedi.2000.14.4.291

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry*, *162*(10), 1911–1918. https://doi.org/10.1176/appi.ajp.162.10.1911

How to cite this article: Wibbelink, C. J. M., Venhuizen, A. S. S. M., Grasman, R. P. P. P., Bachrach, N., van den Hengel, C., Hudepohl, S., Kunst, L., de Lange, H., Louter, M. A., Matthijssen, S. J. M. A., Schaling, A., Walhout, S., Wichers, K. (R.), & Arntz, A. (2023). Group schema therapy for cluster-C personality disorders: A multicentre open pilot study. *Clinical Psychology & Psychotherapy*, *30*(6), 1279–1302. https://doi.org/10.1002/cpp.2903

APPENDIX A

A.1 | MEASURES

A.1.1 | Screening and prognostic measures

A.1.1.1 | Mental disorders

Based on site preference and availability, personality disorders were assessed using the SCID-II (First et al., 1997) or SCID-5-PD (First et al., 2015) and syndrome disorders were assessed with the SCID-I (First et al., 1996), SCID-5-S (Arntz et al., 2018) or MINI (Sheehan et al., 1998). It was allowed to administer the self-report screening questionnaires (SCID-II-PQ; First et al., 1995, SCID-5-SPQ; First et al., 2016, and SCID-5-SV; Arntz et al., 2018) before conducting the SCID interviews. The interviews have demonstrated satisfactory psychometrics properties (e.g., Lobbestael et al., 2011; Maffei et al., 1997; Mohammadkhani et al., 2020; Osório et al., 2019; Shankman et al., 2018; Sheehan et al., 1997; Somma et al., 2017; Zanarini et al., 2000).

A.1.1.2 | Motivation and availability

A 12-item semi-structured screening interview was administered to assess several exclusion criteria (e.g., have received ST in the past year) and patient's availability and motivation for the study.

A.1.1.3 | Prognostic factors

Potential prognostic factors include autistic traits and childhood maltreatment. Autistics traits were assessed using an abbreviated version of the Adult Autism Spectrum Quotient (AQ-10; Allison et al., 2012). The AQ-10 contains 10 items rated on a 4-point Likert scale and has shown acceptable psychometric properties (Lundin et al., 2019). Childhood maltreatment was assessed with the Childhood Trauma Questionnaire-Short Form (CTQ-SF; Bernstein et al., 2003). The CTQ-SF consists of 28 5-point Likert scale items measuring five types of childhood maltreatment. The CTQ-SF has demonstrated satisfactory psychometric properties (Karos et al., 2014; Thombs et al., 2009). Predictive effects of the AQ-10 and CTQ-SF will be reported in a separate paper.

A.1.2 | Primary outcome

A.1.2.1 | Cluster C personality disorder severity

The primary outcome is change in severity of the manifestations of the pertinent Cluster-C PD. For each disorder, a semi-structured severity interview exists (APD: Avoidant Personality Disorder Severity Index (AVPDSI; Baljé et al., 2022), DPD: Dependent Personality Disorder Severity Index (DPDSI; Tese, 2019), and OCPD: Obsessive Compulsive Personality Disorder Severity Index (OCPDSI; Verheul et al., 2020)), measuring the severity of the DSM-5 criteria of the pertinent PD in the last month. The AVPDSI consists of 53 items divided over seven subscales related to the seven DSM-5 criteria. Each item includes two 6-point Likert scale questions (0 = *never* to 5 = *every day*) measuring avoidance and fear. A total score was calculated by averaging the means of the seven subscales. As the first subscale contains questions related to work or school, this subscale was only included for patients who worked or studied at least 8 h a week and of which the activities were sufficiently representative for a work/study situation (e.g., not only working at home and with personal contact with colleagues). The DPDSI and OCPDSI consist of respectively 63 and 67 items, divided over eight subscales reflecting the eight DSM-5 criteria. Each subscale includes several frequency questions to determine the frequency of the specific DSM-5 criterion on a 6-point Likert scale (0 = *never* to 5 = *daily*), as well as two burden and impact items to assess the subjective burden and impact of the specific DSM-5 criterion on a 10-point Likert scale (1 = *none* to 10 = *unbearable*). An overall frequency score was calculated by summing the mean frequency scores of the subscales. An overall burden and impact score was obtained by averaging the burden and impact items. Subsequently, a total score was calculated by adding the overall frequency score and three times the overall burden and impact score. The psychometric properties of the severity interviews were adequate (Baljé et al., 2022.; Tese, 2019; Verheul et al., 2020). In the current study, reliability was satisfactory on the different time points (AVPDSI: $\alpha = .85–.90$, DPDSI: $\alpha = .85–.93$, and OCPDSI: $\alpha = .76–.90$).

As the severity scores were calculated for each disorder separately, the scores were transformed to *Z*-scores (i.e., for each severity interview, the raw scores were subtracted by the mean baseline score and, subsequently, divided by the baseline standard deviation) before the severity scores were merged.

A.1.3 | Secondary outcomes

A.1.3.1 | Demographics

General patient characteristics (e.g., age, ethnicity, employment status and educational level) were assessed using a semi-structured demographic interview. Additional information, including medication use and use of psychological treatments other than GST, were also collected at each assessment.

A.1.3.2 | General psychopathological symptoms

General psychopathological symptoms were assessed using the Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983), a short version of the Symptom-Check-List (SCL-90-R; Derogatis et al., 1973). The BSI contains 53 5-point Likert scale items assessing nine symptom dimensions (i.e., somatization, obsession-compulsion,

interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoia and psychoticism) and three indices of distress (i.e., positive symptom distress index, positive symptom total, and global severity index). The BSI has shown satisfactory psychometric properties (de Beurs, 2011; de Beurs & Zitman, 2006). In this study, the global severity index was used as an index of general psychopathological symptoms. Cronbach's alpha was high for all time points (.95–.97).

A.1.3.3 | Quality of life

Quality of life was measured using the 5-level EuroQol 5D version (EQ-5D-5L; Herdman et al., 2011). The EQ-5D-5L assesses five health state dimensions (i.e., mobility, self-care, usual activities, pain/discomfort and anxiety/depression), whereby each dimension is divided into five severity levels. Previous research has demonstrated adequate psychometric properties (Janssen et al., 2013), while reliability was satisfactory in our sample ($\alpha = .72–.77$). Dutch social tariffs were applied to generate an index value (Versteegh et al., 2016).

A.1.3.4 | Functional impairment

The 36-item World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) interview version (Üstün et al., 2010) was used to assess global functioning and impairment. The WHODAS 2.0 measures disability in six life domains (i.e., cognition, mobility, self-care, getting along, life activities and participation). All items are assessed using a 5-point Likert scale. The WHODAS 2.0 has shown to be a reliable and valid measure (Üstün et al., 2010). A summary score was derived by using the complex scoring method (Üstün et al., 2010). Cronbach's alpha ranged from .91 to .94 on the different time points.

A.1.3.5 | Happiness

Happiness was measured using a single item (Abdel-Khalek, 2006). This item measures general happiness in the months prior to the assessment on a 7-point Likert scale. This item has satisfactory psychometric properties (Abdel-Khalek, 2006; Dickhaut & Arntz, 2014).

A.1.3.6 | Personality disorder related beliefs

PD-related beliefs were assessed with the Personality Disorder Beliefs Questionnaire (PDBQ; Arntz et al., 2004; Dreessen & Arntz, 1995). The PDBQ consists of 69 items measuring beliefs hypothesized to be typical for various PDs. In this study, only beliefs specific for APD (10 items), OCPD (9 items), and DPD (13 items) were assessed and rated on a 100 mm visual analog scale. The average of the three subscales was used as an indication of severity of PD-related beliefs. The PDBQ has demonstrated good reliability (Arntz et al., 2004). Cronbach's alpha varied between .92 and .96 in our sample.

A.1.3.7 | Self-esteem

Self-esteem was assessed with the Rosenberg's Self-Esteem Scale (RSES; Franck et al., 2008; Rosenberg, 1965). The RSES consists of ten 4-point Likert scale items. The RSES has shown good psychometric properties (Franck et al., 2008) and high reliability in our sample ($\alpha = .86–.92$).

A.1.3.8 | Self-ideal discrepancy

The Miskimins Self-Goal-Other Discrepancy Scale (MSGO; Miskimins & Braucht, 1971) was used to measure self-ideal discrepancy. The MSGO contains 15 adjective pair items describing a trait (e.g., 'alert and active' vs. 'dull and lifeless'). For each trait, participants indicate on a 100 mm visual analogue scale where they currently place themselves (actual self) and where they would ideally like to be (ideal self). Of interest were not the absolute ratings, but the differences between the actual self and ideal self ratings. A single discrepancy score was calculated by averaging all difference scores. Based on previous research, the MSGO has shown adequate validity and reliability (Miskimins & Braucht, 1971; Molendijk et al., 2010; Roelofs et al., 2007). In our study, Cronbach's alpha ranged from .85 to .88 on the different time points.

A.1.3.9 | Early maladaptive schemas

Early maladaptive schemas were measured by the 90-item Young Schema Questionnaire – Revised (YSQ-R; Rijkeboer, 2013; Young, 2005). The YSQ-R measures 18 early maladaptive schemas. All items are assessed using a 6-point Likert scale. The YSQ-R has shown adequate psychometric properties (Lee et al., 2015). A mean score was used in this study as an index of early maladaptive schemas. Cronbach's alpha ranged from .96 to .98 on the different time points.

A.1.3.10 | Schema modes

The Schema Mode Inventory-2 (SMI-2; Bamelis et al., 2011; Young et al., 2007) was used to assess the presence of functional and dysfunctional schema modes. The SMI-2 contains 192 6-point Likert scale items assessing 18 different schema modes (one functional and 17 dysfunctional). Previous research has demonstrated adequate reliability (Bamelis et al., 2011). For this study, an index of maladaptive schema mode activation was calculated by averaging the dysfunctional schema modes. The score on the functional schema mode (Healthy Adult) was used as an indication for functional schema mode activation. In the current study, high reliabilities were found for the functional ($\alpha = .76–.91$) and dysfunctional ($\alpha = .97–.99$) schema modes.

APPENDIX B

B.1 | RESULTS OF THE PRIMARY AND SECONDARY OUTCOMES

TABLE B1 Results of the mixed regression analyses.

Outcome, effect	*b*	*df*[1]	*t*	*p*
Primary outcome				
Severity[2]				
Time 0–8 months APD	−0.12[a]	11	−8.04	<.001
DPD vs. APD	0.11	122	2.51	.013
OCPD vs. APD	0.04	126	1.18	.241
Time 8–24 months APD	−0.02[b]	84	−4.16	<.001
DPD vs. APD	−0.07	82	−3.79	<.001
OCPD vs. APD	0.03	82	2.07	.041[3]
Δ 0–24 months APD	−1.32[x]	14	−10.24	<.001
Δ 0–24 months DPD	−1.51[x]	109	−4.25	<.001
Δ 0–24 months OCPD	−0.59[y]	84	−2.17	.033
Secondary outcomes				
General psychopathological symptoms				
Intercept APD	0.35	133	8.26	<.001
DPD vs. APD	0.05	133	0.37	.711
OCPD vs. APD	−0.24	132	−2.19	.030
Time 0–8 months APD	−0.03[a]	114	−5.61	<.001
DPD vs. APD	0.04	128	1.97	.051[4]
OCPD vs. APD	−0.01	119	−0.83	.408
Time 8–24 months APD	−0.01[b]	81	−3.27	.002
DPD vs. APD	−0.01	80	−0.8	.428
OCPD vs. APD	0.01	78	1.1	.274
Δ 0–24 months APD	−0.44[x]	102	−7.17	<.001
Δ 0–24 months DPD	−0.26[x]	105	−1.23	.223
Δ 0–24 months OCPD	−0.40[x]	97	−2.84	.006
Quality of Life[5]				
Intercept APD	−0.82	140	−15.29	<.001
DPD vs. APD	0.17	140	0.97	.333
OCPD vs. APD	0.02	140	0.14	.891
Time 0–8 months APD	−0.03[a]	227	−3.43	.001
DPD vs. APD	0.01	248	0.28	.777
OCPD vs. APD	−0.04	236	−2.17	.031
Time 8–12 months APD	−0.02[a,b]	249	−1.38	.169
DPD vs. APD	0.02	261	0.32	.746
OCPD vs. APD	0.09	255	2.08	.039
Time 12–24 months APD	−0.00[b]	165	−0.19	.852
DPD vs. APD	−0.05	159	−2.16	.032
OCPD vs. APD	−0.03	166	−1.87	.063
Δ 0–24 months APD	−0.33[x]	112	−4.18	<.001
Δ 0–24 months DPD	−0.83[x]	110	−3.08	.003
Δ 0–24 months OCPD	−0.71[x]	114	−3.98	<.001

(Continues)

TABLE B1 (Continued)

Outcome, effect	*b*	*df*[1]	*t*	*p*
Functional impairment				
Intercept APD	3.66	144	81.58	<.001
DPD vs. APD	0.01	144	0.09	.927
OCPD vs. APD	−0.22	144	−1.93	.056
Time 0–8 months APD	−0.03[a]	218	−5.47	<.001
DPD vs. APD	0.02	251	0.92	.357
OCPD vs. APD	−0.02	232	−1.28	.202
Time 8–24 months APD	−0.01[a]	7	−3.35	.011
DPD vs. APD	−0.04	116	−3.98	<.001
OCPD vs. APD	0.00	128	0.52	.605
Δ 0–24 months APD	−0.45[x]	12	−5.77	<.001
Δ 0–24 months DPD	−1.02[y]	100	−4.82	<.001
Δ 0–24 months OCPD	−0.51[x]	69	−3.26	.002
Happiness				
Intercept APD	3.10	270	28.01	<.001
DPD vs. APD	−0.20	270	−0.54	.591
OCPD vs. APD	0.33	268	1.14	.253
Time 0–8 months APD	0.12[a]	16	6.89	<.001
DPD vs. APD	−0.05	296	−1.00	.320
OCPD vs. APD	0.02	278	0.63	.530
Time 8–24 months APD	0.01[b]	348	1.35	.177
DPD vs. APD	0.05	352	1.65	.100
OCPD vs. APD	−0.02	342	−0.96	.339
Δ 0–24 months APD	1.13[x]	35	6.71	<.001
Δ 0–24 months DPD	1.56[x]	217	2.94	.004
Δ 0–24 months OCPD	0.99[x]	163	2.71	.008
Personality disorder related beliefs				
Intercept APD	55.07	14	36.68	<.001
DPD vs. APD	7.45	149	1.52	.131
OCPD vs. APD	−10.90	148	−2.84	.005
Time 0–8 months APD	−1.61[a]	223	−8.67	<.001
DPD vs. APD	1.14	251	1.76	.080
OCPD vs. APD	0.66	231	1.36	.175
Time 8–24 months APD	−0.34[b]	169	−3.10	.002
DPD vs. APD	−0.55	157	−1.39	.166
OCPD vs. APD	0.01	171	0.05	.963
Δ 0–24 months APD	−18.29[x]	107	−8.98	<.001
Δ 0–24 months DPD	−17.94[x]	113	−2.59	.011
Δ 0–24 months OCPD	−12.83[x]	103	−2.71	.008
Self-esteem				
Intercept APD	9.21	134	21.13	<.001
DPD vs. APD	1.14	134	0.80	.426
OCPD vs. APD	4.39	133	3.92	<.001
Time 0–8 months APD	0.54[a]	119	8.64	<.001
DPD vs. APD	−0.30	134	−1.38	.170
OCPD vs. APD	−0.22	124	−1.35	.179

TABLE B1 (Continued)

Outcome, effect	*b*	*df*[1]	*t*	*p*
Time 8–12 months APD	0.19[b]	11	1.74	.109
DPD vs. APD	−0.77	99	−2.46	.016
OCPD vs. APD	−0.18	96	−0.79	.433
Time 12–24 months APD	0.02[b]	86	0.50	.617
DPD vs. APD	0.48	82	3.42	.001
OCPD vs. APD	0.04	80	0.43	.672
Δ 0–24 months APD	5.31[x]	39	8.34	<.001
Δ 0–24 months DPD	5.53[x]	117	2.79	.006
Δ 0–24 months OCPD	3.33[x]	102	2.40	.018
Self-ideal discrepancy				
Intercept APD	−45.03	230	−29.00	<.001
DPD vs. APD	0.77	231	0.15	.881
OCPD vs. APD	11.69	229	2.93	.004
Time 0–8 months APD	1.80[a]	19	9.83	<.001
DPD vs. APD	−0.97	302	−1.55	.122
OCPD vs. APD	−0.74	283	−1.58	.115
Time 8–24 months APD	0.33[b]	342	3.04	.003
DPD vs. APD	0.20	350	0.53	.599
OCPD vs. APD	−0.31	345	−1.18	.237
Δ 0–24 months APD	19.59[x]	45	10.40	<.001
Δ 0–24 months DPD	15.12[x,y]	172	2.38	.018
Δ 0–24 months OCPD	8.67[y]	137	2.01	.047
Early maladaptive schemas				
Intercept APD	3.53	134	57.02	<.001
DPD vs. APD	−0.02	134	−0.11	.909
OCPD vs. APD	−0.46	134	−2.88	.005
Time 0–8 months APD	−0.06[a]	5	−7.46	<.001
DPD vs. APD	0.04	120	1.38	.169
OCPD vs. APD	0.01	119	0.26	.796
Time 8–12 months APD	−0.02[b]	93	−1.88	.064
DPD vs. APD	0.10	99	2.33	.022
OCPD vs. APD	0.04	91	1.33	.187
Time 12–24 months APD	−0.01[b]	14	−1.74	.104
DPD vs. APD	−0.06	79	−3.67	<.001
OCPD vs. APD	0.00	75	0.18	.860
Δ 0–24 months APD	−0.69[x]	6	−9.47	<.001
Δ 0–24 months DPD	−0.71[x]	96	−3.15	.002
Δ 0–24 months OCPD	−0.46[x]	53	−2.91	.005
Schema modes - functional				
Intercept APD	2.69	134	49.84	<.001
DPD vs. APD	0.13	135	0.72	.473
OCPD vs. APD	0.35	134	2.54	.012
Time 0–8 months APD	0.08[a]	119	8.02	<.001
DPD vs. APD	−0.06	131	−1.80	.074
OCPD vs. APD	−0.02	123	−0.68	.499
Time 8–24 months APD	0.02[b]	7	3.92	.005

(Continues)

TABLE B1 (Continued)

Outcome, effect	*b*	*df*[1]	*t*	*p*
DPD vs. APD	0.01	79	0.83	.408
OCPD vs. APD	−0.01	76	−0.52	.605
Δ 0–24 months APD	0.85[x]	23	10.11	<.001
Δ 0–24 months DPD	0.58[x]	110	2.06	.042
Δ 0–24 months OCPD	0.64[x]	89	3.27	.002
Schema modes - dysfunctional				
Intercept APD	3.28	132	70.17	<.001
DPD vs. APD	0.07	133	0.45	.652
OCPD vs. APD	−0.29	132	−2.40	.018
Time 0–8 months APD	−0.05[a]	6	−9.00	<.001
DPD vs. APD	0.05	125	2.55	.012
OCPD vs. APD	−0.00	121	−0.04	.970
Time 8–24 months APD	−0.01[b]	9	−3.98	.003
DPD vs. APD	−0.03	79	−2.94	.004
OCPD vs. APD	0.01	77	2.37	.020
Δ 0–24 months APD	−0.60[x]	9	−10.69	<.001
Δ 0–24 months DPD	−0.61[x]	105	−3.31	.001
Δ 0–24 months OCPD	−0.36[x]	66	−2.82	.006

Note: Slopes between 0–8 and 8–24 months (two-piecewise model) or 0–8, 8–12 and 12–24 months (three-piecewise model) with different letters (a, b, c) were significantly different from each other. Effects between 0 and 24 months with different letters (x, y, z) were significantly different from each other.
Abbreviations: APD, avoidant personality disorder; DPD, dependent personality disorder; OCPD, obsessive-compulsive personality disorder.
[1]Rounded values.
[2]Results of the intercept and main effect of diagnosis are meaningless and therefore not presented. Severity scores were calculated for each disorder separately and transformed to *Z*-scores before the severity scores were merged.
[3]Trend effect ($p = .055$) when outliers were removed.
[4]Significant effect ($p = .049$) when outliers were removed.
[5]Lower scores indicate a positive effect. To enable gamma regression, scores were transformed by multiplying the score with −1 and adding 1.01.

TABLE B2 Primary and secondary outcomes across all time points.

	APD			DPD			OCPD		
Outcome, time[a]	Estimated mean (95% CI)[b]		*d*[c]	Estimated mean (95% CI)[b]		*d*[c]	Estimated mean (95% CI)[b]		*d*[c]
Primary outcome									
Severity									
0	0.01	(−0.19;0.20)	NA	−0.02	(−0.62;0.57)	NA	−0.08	(−0.53;0.38)	NA
4	−0.48	(−0.69;-0.28)	0.49	−0.09	(−0.68;0.51)	0.06	−0.41	(−0.86;0.04)	0.34
8	−0.97	(−1.26;-0.69)	0.99	−0.15	(−0.90;0.61)	0.12	−0.74	(−1.31;-0.17)	0.67
12	−1.06	(−1.34;-0.78)	1.07	−0.49	(−1.22;0.23)	0.47	−0.72	(−1.28;-0.17)	0.65
24	−1.31	(−1.61;-1.02)	1.33	−1.54	(−2.32;-0.75)	1.53	−0.66	(−1.25;-0.07)	0.59
Secondary outcomes									
General psychopathological symptoms									
0	1.42	(1.30;1.54)	NA	1.49	(1.14;1.95)	NA	1.11	(0.91;1.36)	NA
4	1.23	(1.12;1.36)	0.27	1.54	(1.13;2.09)	−0.06	0.92	(0.73;1.16)	0.38
8	1.08	(0.95;1.22)	0.55	1.59	(1.05;2.39)	−0.12	0.76	(0.56;1.03)	0.76
12	1.03	(0.91;1.17)	0.63	1.46	(0.98;2.19)	0.04	0.75	(0.56;1.02)	0.77
24	0.91	(0.78;1.05)	0.89	1.15	(0.71;1.88)	0.52	0.74	(0.52;1.04)	0.81
Quality of life									
0	0.57	(0.52;0.61)	NA	0.49	(0.28;0.63)	NA	0.56	(0.43;0.66)	NA
4	0.61	(0.57;0.66)	0.19	0.53	(0.32;0.67)	0.14	0.67	(0.57;0.75)	0.51
8	0.65	(0.60;0.70)	0.39	0.56	(0.29;0.73)	0.27	0.76	(0.65;0.83)	1.02
12	0.69	(0.64;0.73)	0.55	0.57	(0.30;0.74)	0.30	0.68	(0.54;0.77)	0.54
24	0.69	(0.63;0.74)	0.58	0.78	(0.59;0.89)	1.49	0.79	(0.68;0.86)	1.27
Functional impairment									
0	38.67	(35.39;42.25)	NA	39.27	(29.81;51.74)	NA	30.99	(25.12;38.22)	NA
4	34.62	(31.42;38.15)	0.21	37.51	(27.65;50.88)	0.09	25.96	(20.63;32.67)	0.34
8	31.00	(27.52;34.92)	0.43	35.83	(24.50;52.39)	0.18	21.75	(16.39;28.85)	0.69
12	29.24	(25.96;32.94)	0.54	28.42	(19.44;41.56)	0.63	20.99	(15.82;27.83)	0.77
24	24.54	(21.07;28.57)	0.88	14.19	(8.74;23.03)	1.99	18.86	(13.23;26.87)	1.00
Happiness									
0	3.10	(2.89;3.32)	NA	2.91	(2.23;3.59)	NA	3.43	(2.91;3.95)	NA
4	3.57	(3.37;3.77)	0.42	3.17	(2.58;3.76)	0.24	4.00	(3.55;4.44)	0.51
8	4.04	(3.76;4.32)	0.84	3.43	(2.70;4.17)	0.47	4.56	(4.02;5.11)	1.02
12	4.09	(3.82;4.36)	0.89	3.69	(3.01;4.37)	0.70	4.53	(4.02;5.04)	0.99
24	4.23	(3.91;4.55)	1.02	4.47	(3.50;5.43)	1.40	4.42	(3.77;5.08)	0.89
Personality disorder related beliefs									
0	55.07	(51.84;58.30)	NA	62.52	(53.28;71.77)	NA	44.17	(37.14;51.20)	NA
4	48.63	(45.39;51.87)	0.40	60.63	(51.28;69.97)	0.12	40.36	(33.30;47.42)	0.24
8	42.19	(38.40;45.98)	0.80	58.73	(47.03;70.42)	0.23	36.55	(27.89;45.21)	0.47
12	40.84	(37.26;44.42)	0.88	55.19	(44.21;66.18)	0.45	35.25	(27.14;43.36)	0.55
24	36.78	(32.71;40.85)	1.13	44.59	(31.19;57.98)	1.11	31.35	(22.15;40.54)	0.79
Self-esteem									
0	9.21	(8.35;10.07)	NA	10.35	(7.66;13.04)	NA	13.60	(11.56;15.64)	NA
4	11.37	(10.49;12.25)	0.45	11.31	(8.55;14.06)	0.20	14.88	(12.81;16.95)	0.27
8	13.53	(12.40;14.66)	0.90	12.26	(8.60;15.93)	0.40	16.16	(13.47;18.85)	0.53
12	14.28	(12.98;15.59)	1.05	9.94	(6.11;13.77)	−0.08	16.21	(13.37;19.04)	0.54
24	14.52	(13.21;15.83)	1.10	15.88	(11.91;19.85)	1.15	16.93	(14.13;19.73)	0.69

(Continues)

TABLE B2 (Continued)

Outcome, time[a]	APD			DPD			OCPD		
	Estimated mean (95% CI)[b]		d[c]	Estimated mean (95% CI)[b]		d[c]	Estimated mean (95% CI)[b]		d[c]
Self-ideal discrepancy									
0	−45.03	(−48.09;-41.97)	NA	−44.26	(−53.81;-34.71)	NA	−33.34	(−40.59;-26.09)	NA
4	−37.83	(−40.59;-35.07)	0.43	−40.93	(−49.52;-32.34)	0.20	−29.10	(−35.57;-22.62)	0.25
8	−30.63	(−33.92;-27.35)	0.86	−37.59	(−47.60;-27.58)	0.40	−24.85	(−32.22;-17.48)	0.51
12	−29.33	(−32.48;-26.19)	0.94	−35.48	(−44.93;-26.03)	0.53	−24.81	(−31.77;-17.84)	0.51
24	−25.43	(−29.27;-21.60)	1.18	−29.14	(−41.71;-16.57)	0.91	−24.67	(−33.27;-16.08)	0.52
Early maladaptive schemas									
0	3.53	(3.41;3.65)	NA	3.50	(3.12;3.89)	NA	3.07	(2.78;3.36)	NA
4	3.28	(3.14;3.42)	0.38	3.40	(2.98;3.83)	0.16	2.84	(2.52;3.16)	0.35
8	3.03	(2.84;3.22)	0.76	3.30	(2.75;3.84)	0.31	2.61	(2.20;3.02)	0.70
12	2.94	(2.74;3.14)	0.90	3.61	(3.04;4.17)	−0.16	2.68	(2.26;3.10)	0.59
24	2.84	(2.64;3.04)	1.05	2.79	(2.22;3.36)	1.08	2.61	(2.19;3.03)	0.70
Schema modes – Functional									
0	2.69	(2.58;2.79)	NA	2.81	(2.48;3.15)	NA	3.04	(2.79;3.29)	NA
4	2.99	(2.88;3.10)	0.52	2.88	(2.54;3.23)	0.12	3.27	(3.01;3.53)	0.41
8	3.29	(3.13;3.44)	1.04	2.95	(2.46;3.45)	0.24	3.51	(3.14;3.87)	0.82
12	3.35	(3.20;3.50)	1.15	3.06	(2.58;3.55)	0.43	3.55	(3.20;3.91)	0.89
24	3.54	(3.35;3.72)	1.48	3.39	(2.81;3.98)	1.01	3.68	(3.27;4.09)	1.11
Schema modes – Dysfunctional									
0	3.27	(3.18;3.36)	NA	3.34	(3.05;3.63)	NA	2.98	(2.77;3.20)	NA
4	3.05	(2.95;3.16)	0.44	3.33	(3.00;3.67)	0.02	2.76	(2.51;3.02)	0.44
8	2.84	(2.69;2.98)	0.88	3.32	(2.89;3.76)	0.04	2.54	(2.22;2.87)	0.89
12	2.80	(2.66;2.94)	0.96	3.17	(2.75;3.60)	0.34	2.56	(2.25;2.88)	0.85
24	2.68	(2.53;2.83)	1.20	2.73	(2.27;3.19)	1.24	2.62	(2.29;2.95)	0.73

Abbreviations: APD, avoidant personality disorder; CI, confidence interval; DPD, dependent personality disorder; NA, not available; OCPD, obsessive-compulsive personality disorder.

[a]Time in months.

[b]Estimated means and 95% confidence intervals are in original scale.

[c]Effect sizes are positive when there is a positive effect.

APPENDIX C

C.1 | POST-HOC ANALYSES

C.1.1 | Statistical analysis

Four exploratory post hoc analyses were performed to examine the influences of psychotropic medication and receiving psychological treatment during the follow-up period on change in the primary outcome. Focusing on psychotropic medication use, we examined differences in change trajectories in the primary outcome between (1) patients using psychotropic medication before start of the treatment ('medication use' group) versus patients that did not use psychotropic medication before start of the treatment ('no medication use' group), and (2) patients of which the psychotropic medication use increased during the treatment period (i.e., patients that started with medication or increased the dosage of their medication; 'increased medication use' group) versus patients that did not use psychotropic medication, stopped with using psychotropic medication, or used the same dosage ('stable medication use' group). Focusing on psychological treatments during the follow-up period (i.e., after the study treatment), we assessed whether the change trajectories in the primary outcome differed between (3) patients receiving individual psychological treatment during the follow-up period ('individual treatment' group) versus patients that did not receive psychological treatment during the follow-up period ('no treatment' group), and (4) patients receiving any form of psychological treatment (i.e., individual treatment, group therapy or relationship/family therapy) during the follow-up period ('treatment' group) versus patients that did not receive psychological treatment during the follow-up period ('no treatment' group).

Change in the primary outcome was analysed using a two-piecewise mixed regression model with an unstructured covariance structure for the repeated part, a random slope for treatment group, and fixed effects for time, group (e.g., 'medication use' group vs. 'no medication use' group), and the interaction between time and group. The fixed effect of diagnosis could not be included in the analyses because of the low number of DPD and OCPD patients. In addition, the analyses included different numbers of patients as information of different assessments was necessary to be able to create the groups.

C.1.2 | Results

The results of the post-hoc analyses are presented in Table C1. No significant differences were found in the intercept and change trajectories (0–8 months and 8–24 months) between patients using psychotropic medication before start of the treatment compared to patients that did not use psychotropic medication before start of the treatment. Moreover, the intercept and change trajectories did not differ between patients that did not receive psychological treatment during the follow-up period compared to patients receiving individual psychological treatment or patients receiving any form of psychological treatment during the follow-up period. However, significant differences were found between patients of which the psychotropic medication use increased during the treatment period versus patients that did not use psychotropic medication, stopped with using psychotropic medication, or used the same dosage. Patients of which the psychotropic medication use increased during the treatment period scored significantly higher on PD severity at baseline compared to patients that did not use psychotropic medication, stopped with using psychotropic medication, or used the same dosage. In both groups, there was a significant decrease over time (0–8 months and 8–24 months) in severity of PD symptoms. However, the decrease over time during treatment (0–8 months) and after treatment (8–24 months) was significantly larger for patients that started with medication or increased the dosage of their medication during the treatment period versus patients that did not use medication, stopped with using medication or used the same dosage.

TABLE C1 Results of the exploratory mixed regression analyses.

Post hoc analysis	*b*	*df*[a]	*t*	*p*
1: Psychotropic medication use at baseline (*N* = *137*)				
Intercept no medication use	−0.04	133	−0.34	.734
Medication use vs. no medication use	0.07	133	0.42	.674
Time 0–8 months no medication use	−0.10	18	−5.64	<.001
Medication use vs. no medication use	−0.01	119	−0.61	.544
Time 8–24 months no medication use	−0.01	87	−2.06	.042
Medication use vs. no medication use	−0.02	86	−1.77	.081
2: Psychotropic medication use during the treatment (*N* = *112*)				
Intercept stable medication use	−0.17	108	−1.58	.117
Increased medication use vs. stable medication use	0.62	107	2.70	.008
Time 0–8 months stable medication use	−0.10	11	−7.00	<.001
Increased medication use vs. stable medication use	−0.06	108	−2.01	.047
Time 8–24 months stable medication use	−0.02	86	−2.90	.005
Increased medication use vs. stable medication use	−0.03	85	−2.37	.020
3: Individual psychological treatment during the follow-up period (*N* = *75*)				
Intercept no treatment	0.08	71	0.61	.547
Individual treatment vs. no treatment	−0.36	70	−1.10	.273
Time 0–8 months no treatment	−0.13	11	−5.72	<.001
Individual treatment vs. no treatment	0.05	70	1.11	.271
Time 8–24 months no treatment	−0.02	71	−4.04	<.001
Individual treatment vs. no treatment	0.00	74	0.20	.840
4: Any psychological treatment during the follow-up period (*N* = *81*)				
Intercept no treatment	0.08	77	0.62	.536
Psychological treatment vs. no treatment	−0.25	76	−0.90	.372
Time 0–8 months no treatment	−0.13	12	−6.08	<.001
Psychological treatment vs. no treatment	0.06	78	1.73	.088
Time 8–24 months no treatment	−0.02	75	−4.13	<.001
Psychological treatment vs. no treatment	0.01	78	0.56	.580

[a]Rounded values.