UvA-DARE (Digital Academic Repository)

Schema therapy for personality disorders in older adults: a multiple-baseline study

Videler, A.C.; van Alphen, S.P.J.; van Royen, R.J.J.; van der Feltz-Cornelis, C.M.; Rossi, G.; Arntz, A.

Routledge
Taylor & Francis Group

Check for updates

Schema therapy for personality disorders in older adults: a multiple-baseline study

Arjan C. Videler [ID][a], Sebastiaan P. J. van Alphen[b,c], Rita J. J. van Royen[d], Christina M. van der Feltz-Cornelis[e,f], Gina Rossi[b] and Arnoud Arntz [ID][g]

[a]Department of Geriatric Psychiatry, GGz Breburg, Tilburg, The Netherlands; [b]Department of Clinical and Life Span Psychology, Vrije Universiteit Brussel (VUB), Brussels, Belgium; [c]Department of Geriatric Psychiatry, Belgium and Mondriaan Hospital, Heerlen-Maastricht, The Netherlands; [d]Center for Psychotherapy, Beuningen, The Netherlands ; [e]Tranzo Department, Tilburg University, Tilburg, The Netherlands ; [f]GGz Breburg, Clinical Center of Excellence for Body, Mind and Health, Tilburg, The Netherlands.; [g]Department of Clinical Psychology, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

Objective: No studies have been conducted yet into the effectiveness of treatment of personality disorders in later life. This study is a first test of the effectiveness of schema therapy for personality disorders in older adults.

Method: Multiple-baseline design with eight cluster C personality disorder patients, with a mean age of 69. After a baseline phase with random length, schema therapy was given during the first year, followed by follow-up sessions during six months. Participants weekly rated the credibility of dysfunctional core beliefs. Symptomatic distress, early maladaptive schemas, quality of life and target complaints were assessed every six months and personality disorder diagnosis was assessed before baseline and after follow-up. Data were analyzed with mixed regression analyses.

Results: Results revealed significant linear trends during treatment phases, but not during baseline and follow-up. The scores during follow-up remained stable and were significantly lower compared to baseline, with high effect sizes. Seven participants remitted from their personality disorder diagnosis.

Conclusion: Schema therapy appears an effective treatment for cluster C personality disorders in older adults. This finding is highly innovative as this is the first study exploring the effectiveness of psychotherapy, in this case schema therapy, for personality disorders in older adults.

Introduction

The prevalence rate of personality disorders (PDs) among older adults over the age of 65 years old was 8% in US community samples (Schuster, Hoertel, Le Strat, Manetti, & Limosin, 2013). Moreover, PDs appeared strongly associated with disability, somatic problems and mental disorders (Schuster et al., 2013) as well as medical resource utilization (Powers, Strube, & Oltmanns, 2014). Notwithstanding, treatment of this complex patient group is highly underexplored (Van Alphen et al., 2015; Van Alphen, Derksen, Sadavoy, & Rosowsky, 2012). In a Delphi study on diagnosis and treatment of PDs in older adults, conducted in the Netherlands and Belgium (Van Alphen et al., 2012), which was recently cross-validated in the US (Rosowsky, Young, Malloy, Van Alphen, & Ellison, 2016), experts agreed that existing evidence-based therapies for PDs in younger age groups are also applicable to older adults over 60 years of age. However, still therapeutic nihilism prevails amongst clinicians and researchers concerning the feasibility and effectiveness of psychotherapy for PDs in later life. Some authors believe that the aim of changing pathological aspects of personality is not possible in older adults, because of the rigidity of lifelong dysfunctional patterns or the consequences of cognitive and physical decline (Segal, Coolidge, & Rosowsky, 2006; Van Alphen et al., 2012). This could explain why no effectiveness studies have been conducted yet into the treatment of PDs in later life as the main focus of therapy. There are only two studies into the treatment of depression with

comorbid PDs and PD features. One small randomized controlled trial (RCT) in 37 older adults explored the effectiveness of dialectical behavior therapy combined with pharmacotherapy for depression and comorbid PDs, compared to pharmacotherapy as stand-alone treatment (Lynch, Cheavens, Cukrowitz, Thorp, Bronner, & Beyer, 2007). The combined treatment did not improve depressive symptoms over medication alone, but only was superior with respect to improved interpersonal sensitivity and interpersonal aggression. The second study examined short group schema therapy (SCBT-g) in 31 older adults with depression and comorbid PDs or PD features with a pre-mid-post design (Videler, Rossi, Schoevaars, Van der Feltz-Cornelis, & Van Alphen, 2014). A medium effect size was found for reduction of depressive symptoms, early maladaptive schemas (EMS) and schema modes. However, treatment effect on the comorbid PD diagnosis was not assessed. Furthermore, this treatment did not involve experiential techniques, like imagery rescripting and chairwork, which are considered more powerful at achieving change at an emotional level than cognitive-behavioral techniques and thus at influencing EMS (Arntz & Van Genderen, 2012). In sum, these two studies do show the feasibility of psychotherapy for comorbid PDs in later life, but they shed no light on psychotherapy for PDs as the main focus of treatment in older adults.

Case studies indicate that schema therapy (ST) is applicable as a treatment of PDs in older adults (Videler, van Royen,

van Alphen, Rossi, & van der Feltz-Cornelis, in press; Videler et al., 2015). Videler et al. (2014) advocated that ST connects to the psychotherapy expectations of older adults, as it incorporates psychoeducation and is structured, skill-enhancing and problem-focused. ST is an integrative treatment, which combines cognitive behavior therapy (CBT), object relations theory, gestalt therapy and attachment theory into a systematic model for the treatment of PDs (Edwards & Arntz, 2012; Young, Klosko, & Weishaar, 2003). In this model, EMS are considered core elements of PDs. The goal of treatment is to decrease the impact of these EMS and to replace negative coping responses and schema modes with more healthy alternatives, so that patients succeed in getting their core emotional needs met (Rafaeli, Bernstein, & Young, 2011). In ST, besides CBT techniques, experiential techniques have a central place (Edwards & Arntz, 2012). There is accumulating evidence for the efficacy of ST in younger age groups, both in treating borderline (Giesen-Bloo et al., 2006; Nadort et al., 2009) and cluster C, paranoid, narcissistic and histrionic PDs (Bamelis, Evers, Spinhoven, & Arntz, 2014), but the effectiveness in older adults is unknown.

Therefore, the aim of the present study was to assess the effectiveness of individual ST as a treatment for PDs in older adults, using a multiple-baseline design (Kazdin, 2010) with eight patients with a PD diagnosis.

We chose the multiple-baseline design for several reasons. Often, as a first evaluation of a treatment, an open trial is done. Contrary to an open trial, a multiple-baseline design offers experimental control over time versus intervention effects. Furthermore, this design has some advantages over a RCT (Kazdin, 2010). The most important advantage of a multiple-baseline design over RCTs is that this design requires fewer participants, because participants act as their own controls, thus increasing power. Nevertheless, like a RCT, a multiple-baseline design can demonstrate significant change and also that this change is the result of the intervention and not of time (Hawkins, Sanson-Fisher, Shakeshaft, D'Este, & Green, 2007; Kazdin, 2010; Onghena, 2005). As treatment of PDs in older adults is a hitherto relatively neglected topic, there are few trained psychotherapists in the field, which complicates conducting a RCT. As our aim was to test the initial effectiveness of ST in later life, without comparing ST to another potentially powerful treatment, we considered the multiple-baseline design to be a good alternative for a RCT, considering the current phase of scientific research. Because the course of cluster B PDs appears to be much less stable throughout the life span than that of cluster C PDs, which especially complicates the diagnosis of cluster B PDs (Cooper, Balsis, & Oltmanns, 2014; Van Alphen et al., 2015), we decided to examine ST in cluster C PDs.

Multiple-baseline designs require dependent variables that are frequently assessed and are highly sensitive to change, to study the time and intervention effects (Kazdin, 2010). Such variables should represent a core aspect of the disorder that is addressed by the treatment (i.e. short-term treatment effects). The frequent assessments make it possible to distinguish time and treatment effects, and allow that each case is its own control. Thus, the high number of assessments of this central variable compensates for the relatively small number of participants. As central variable, we chose the strength of belief participants had in their personal core beliefs, which they viewed as central to their PD problems. This idiosyncratic measure represented the EMS that are assumed to underlie the patient's PD problems according to the ST model (Young et al., 2003).

Our hypothesis was that ST would lead to a decrease of dysfunctional core beliefs, symptomatic distress and EMS, and an increase of quality of life in cluster C PDs.

Methods

Participants

Participants were four patients from the Department of Geriatric Psychiatry of Breburg, and four patients from the Department of Geriatric Psychiatry of Mondriaan, both mental health institutes in the Netherlands. Inclusion criteria were: (1) primary diagnosis of a cluster C PD or PD not otherwise specified with cluster C traits, as assessed with the Dutch version of the Structured Clinical Interview for DSM-IV PDs (SCID-II; Weertman, Arntz, & Kerkhofs, 2000). PD not otherwise specified with cluster C traits was defined as meeting the general criteria for a PD while also meeting sub-threshold criteria for at least one specific cluster C PD; (2) age 60 years old or older; (3) willingness to participate in the study. Exclusion criteria were: (1) severe depression; (2) bipolar disorder; (3) psychotic disorder; (4) IQ under 80; (5) substance dependence; (6) cognitive disorder (Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975) under 25). During the course of the study, no other treatment was allowed and medication was kept constant. The eight participants were recruited from nine patients screened for participation; one declined participation in the study. Figure 1 presents the patient flow. Table 1 gives an overview of the characteristics and the treatments of the participants. Informed consent was obtained from all participants. The study was approved by the ethical committees of the Maastricht University Hospital, Breburg and Mondriaan.

Design

We used a non-concurrent multiple-baseline design (Kazdin, 2010), consisting of four phases. The first phase was a baseline phase varying in length from 3–8 weeks. The variation in baseline length offers the possibility to differentiate between time effects and experimental effects of the treatment. After baseline, weekly ST treatment sessions were given. In our analyses, we divided ST into two treatment phases in order to explore the effect of experiential techniques in older adults, as some authors – and many clinicians – assume that focusing on skills and symptoms is more attainable, like in CBT (for example, Segal et al., 2006). Therefore, we defined a CBT treatment phase, in which cognitive and behavioral techniques were used, and an experiential phase, where the latter was defined by the first introduction of experiential techniques. The two treatment phases differed in length between participants according to the ST methods described by Young et al. (2003): based upon each patient's case conceptualization, the therapist decided when to introduce experiential techniques, in order to match the individual aspects of the patient's problems. Maximum duration of baseline, CBT and experiential phases together was 40 sessions. Finally, a six months followup phase with 10 booster sessions followed to help maintain and assess the effects of ST. During all study phases, outcomes (described in assessments) were repeatedly assessed independently from the therapies.
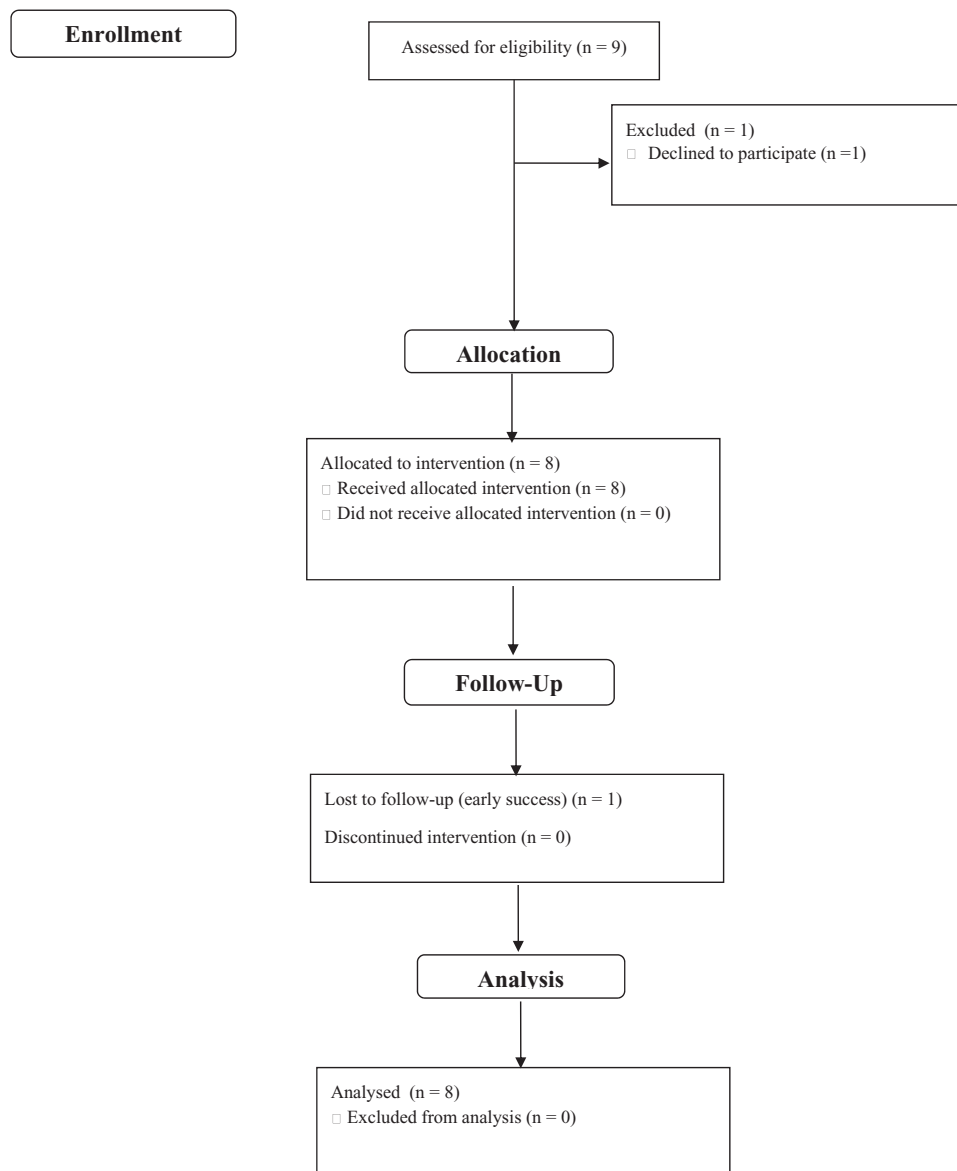
Figure 1. CONSORT flow diagram.

Table 1. Demographic data and treatment information of participants (*N* = 8).

Participants	1	2	3	4	5	6	7	8	Mean (SD)
Age	68	65	67	72	76	69	62	75	69.3/3.8
Gender[a]	F	M	M	F	F	F	F	F	
Educational level[b]	1	2	1	1	1	1	2	1	
PD diagnosis[c]	Avoid	Avoid	NOS	Avoid	OC	NOS	NOS	Dep	
Secondary diagnosis[d]	Depr	Adhd	Depr	Soc P	Panic	Depr	Depr	Depr	
Medication[e]	AD	None	AD	Benzo	AD	AD	None	AD	
Treatment duration prior to study (years)	0.8	6	0.8	0.3	2	4	0.5	0.6	1.9/2.1
Baseline sessions	6	6	5	4	8	6	8	6	6.1/1.4
CBT sessions	6	8	13	18	8	6	6	15	10 /5.1
Experiential sessions	28	28	22	–	26	25	32	17	25.4/5.3
Booster sessions	10	10	10	–	9	8	8	6	8.7/1.5
Total number of sessions	44	46	45	18	43	39	46	38	39.9/9.3

[a] M = male, F = female.
[b] 1 = elementary school with lower vocational training, 2 = higher secondary with vocational training.
[c] Avoid = avoidant PD, OC = obsessive–compulsive PD, Dep = dependent PD, NOS = PD cluster C Not Otherwise Specified.
[d] Depr = depression, Adhd = Attention Deficit Hyperactivity Disorder, Soc P = social phobia, Panic = panic disorder.
[e] AD = antidepressant, Benzo = benzodiazepine.
Note: medication was constant in all participants.

Assessments

As primary outcome, strength of idiosyncratic beliefs was assessed weekly. To formulate these beliefs, participants were interviewed with a semi-structured procedure to elicit 3–5 idiosyncratic dysfunctional beliefs they felt to be central to their PD problems. These dysfunctional core beliefs were then rated weekly by the participants on a visual analog scale (VAS) on 0%–100% credibility. The ratings of the participants were put in an envelope by them and given to the research team directly, so

these ratings were unknown to the therapists, in order to minimize demand effects. Core beliefs were chosen as the primary outcome as they can be frequently assessed, are sensitive for short-term change and are viewed in cognitive models as important representations of EMS deemed to underlie the PD problems (David & Freeman, 2014). The average of the ratings per assessment was taken as dependent variable (range 0–100).

The Dutch SCID-II (Weertman et al., 2000) was used to assess Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2013) PDs as secondary outcome, before baseline and after follow-up. Items are rated on a 3-point scale as absent, sub-threshold or threshold. Inter-rater agreement appeared excellent in adults with an average age of 35.5 years (range 18–61), with a mean value of Cohen's kappa of .84 (Lobbestael, Leurgans, & Arntz, 2011).

Symptomatic distress, another secondary outcome, was assessed with the Dutch version of the Symptom Checklist 90 (SCL-90; Arrindell & Ettema, 2003) four times, before baseline, after six months of treatment, at the end of treatment and after follow-up. The SCL-90 is a 90-item self-report measure of overall psychological distress. Items are scored on a 5-point Likert scale from 'not at all' to 'always.' The reliability of the Dutch SCL-90 is good, the convergent and divergent validity are satisfactory and no age-effect was found for older adults with an average age of 73.5 years (2003). It appeared sensitive to change in clinical settings.

Idiosyncratic target complaints, as secondary outcome, were discussed with all participants in the baseline phase by the therapists and assessed on a Likert scale of 1–9, ranging from 'not at all' to 'can't be worse.' Target complaints are the primary complaints of a patient and for which there is mutual consent between therapist and patient that these are the primary goals of treatment (Battle et al., 1966). Shorer (1970) reported considerable correlations (.71) between global assessments of improvement and improvement on target complaints. Test–retest reliability was .76 (Frey, Heckel, Salzberg, & Wackwitz, 1976). Target complaints were rated four times, in the baseline phase, after six months of treatment, at the end of treatment and after follow-up.

Quality of life (QOL), also a secondary outcome, was assessed with the Dutch World-Health-Organization-Quality-Of-Life, brief version (WHOQOL-BREF; Trompenaars, Masthoff, Van Heck, Hodiamont, & De Vries, 2005). The WHOQOL-BREF is a 26-item self-report measure, which is rated along a 5-point Likert scale. In two samples of older adults, with mean ages of 73 and 76 years, reliability was good and the construct validity satisfactory (Kalfoss, Low, & Molzahn, 2008). The WHOQOL-BREF was rated four times, before baseline, after six months of treatment, at the end of treatment and after follow-up.

EMS, as final secondary outcome, were measured using the Dutch Young Schema Questionnaire (YSQ; Sterk & Rijkeboer, 1997). The questionnaire consists of 205 items, which are phrased as negative core beliefs and rated along a 6-point Likert scale. The YSQ showed good reliability and convergent and discriminant validity in a clinical sample with a mean age of 33.9 years (range 18–74; Rijkeboer & Van den Berg, 2006), and was rated four times, before baseline, after six months of treatment, at the end of treatment and after follow-up.

Procedure

Patients with a primary multidisciplinary diagnosis of a cluster C PD, who met the inclusion criteria, were approached by the first author, until four participants were included at both sites. Potential participants were fully informed about the study and gave written consent to participate. One patient with an obsessive–compulsive PD decided not to participate; he preferred medication over ST. The SCID-II was applied to assess PD diagnosis. In the treatment phases, ST, according to the methods described by Young et al. (2003), was provided in weekly sessions by two therapists (with 6 and 15 years of experience). Treatment integrity was monitored by means of supervision by the third author, a certified ST supervisor. To provide feedback, the treatment of each participant was discussed in supervision at least 10 times and of each participant at least four therapy sessions were filmed and viewed by supervisor and psychotherapist together. In the CBT phase, underlying EMS were targeted by cognitive and behavioral techniques. The experiential phase started with the use of experiential techniques such as imagery rescripting and chairwork (Edwards & Arntz, 2012; Young et al., 2003). The number of treatment sessions was maximized at 40 sessions, although start of the booster sessions was allowed earlier if therapist's and participants' agreed treatment goals were reached; thus, mean length of treatment was somewhat shorter than 40 sessions (see Table 1). During the booster phase, in the last six months of treatment, a maximum of 10 sessions were dedicated to stabilize the progress the participants had made.

Statistical analysis

Core beliefs

Mixed regression analyses were used to assess the differences between the treatment and follow-up phases on the one hand, and baseline on the other hand, in average scores and linear change. The fixed model part consisted of (1) a general linear time effect, starting with time = 0 when the first assessment was taken for an individual, (2) dummy indicators for the CBT, experiential and follow-up phases (thus contrasting each to baseline) and (3) four centered time-within-condition covariates, one for every phase, to assess time-by-phase interaction, that is, changes in the time effect across phases (cf. Arntz, Sofi, & Van Breukelen, 2013; Vlaeyen, De Jong, Geilen, Heuts, & Van Breukelen, 2001). The random model part consisted of an AutoRegressive-Moving-Average model (ARMA11) for the within-subject covariance structure. Random slopes to allow inter-individual variation in time and condition effects led to reduced fit of the model or convergence problems, and were therefore not included.

The analytic strategy was to first test for a general time effect, next to assess the full model with all predictors entered, and then to delete in backward fashion the time-by-phase interactions that were non-significant. If the main time effect was non-significant, it was deleted at the last step. The time effect within baseline was also tested separately for the baseline assessments only. Cohen's d for the core beliefs were calculated as effect size of change at the end of a phase with respect to baseline: d = the mean outcome difference between baseline and current phase, derived from the fixed part of the mixed regression divided by the standard deviation of the residual outcome variance (the patient-specific outcome mean per phase has as variance random intercept (between-subject variance) + (residual (within-subject) variance/number of measurements per phase); the square root of this subject-specific variance is the denominator for d.

Other measures

For the analysis of symptoms, target complaints, QOL and EMS, an unstructured model fitted better for the within-subject covariance structure. For these measures, Cohen's *d* was similarly calculated, but only as effect size of the change between follow-up and baseline, with standard deviation of the baseline as denominator.

Results

Attrition

Participant 4 was considered recovered by herself and the therapist after the CBT phase and declined participation in the follow-up phase. She did fill out all measures four times, however, and although she did not participate in the final SCID-II interview, we did not exclude her from the analyses.

Core beliefs

The individual VAS-scores of the credibility of dysfunctional core beliefs during the different phases are shown in Figure 2. During baseline, the time effect was non-significant, $F(1, 4.83)$ = 1.75, p = .25. Visual inspection suggests decreases in credibility of dysfunctional core beliefs during the treatment phases in all eight participants, and lower scores during follow-up than during baseline in all but participant 4. Mixed regression revealed a significant linear effect of time when tested as single predictor, $t(37.67) = -7.37$, $p < 0.001$. With all predictors entered, the time-within-baseline and time-within-follow-up effects appeared to be non-significant, p's > .35. After stepwise deleting, the main effect of time appeared to be non-significant and was therefore also deleted. Table 2 presents the final results of the mixed regression analysis. The main effect of treatment (i.e. the change at the middle of both treatment phases compared to baseline) was significant, as was the main effect of follow-up (as compared to baseline). The time-within-treatment effect was significant, showing a steep decrease of credibility of core beliefs, both in the CBT phase and in the experiential phase. Effect sizes of treatment versus baseline, and follow-up versus baseline were very high; note that these represent not the middle point of phases but the end point of phases. Figure 3 depicts the predicted means from the analysis.

Symptomatic distress

The individual scores of the participants on the SCL-90 are given in Figure 4(a) (see Figure 4(a)). Visual inspection suggests that all scores decreased, except those of participants 2 and 4. All changes appeared significant including that of participant 2. Participant 4 left treatment after the CBT phase. Figure 4(b) shows the predicted means from the analysis (see Figure 4(b)). Effect size of treatment from baseline to follow-up was high (1.29; see Table 2).
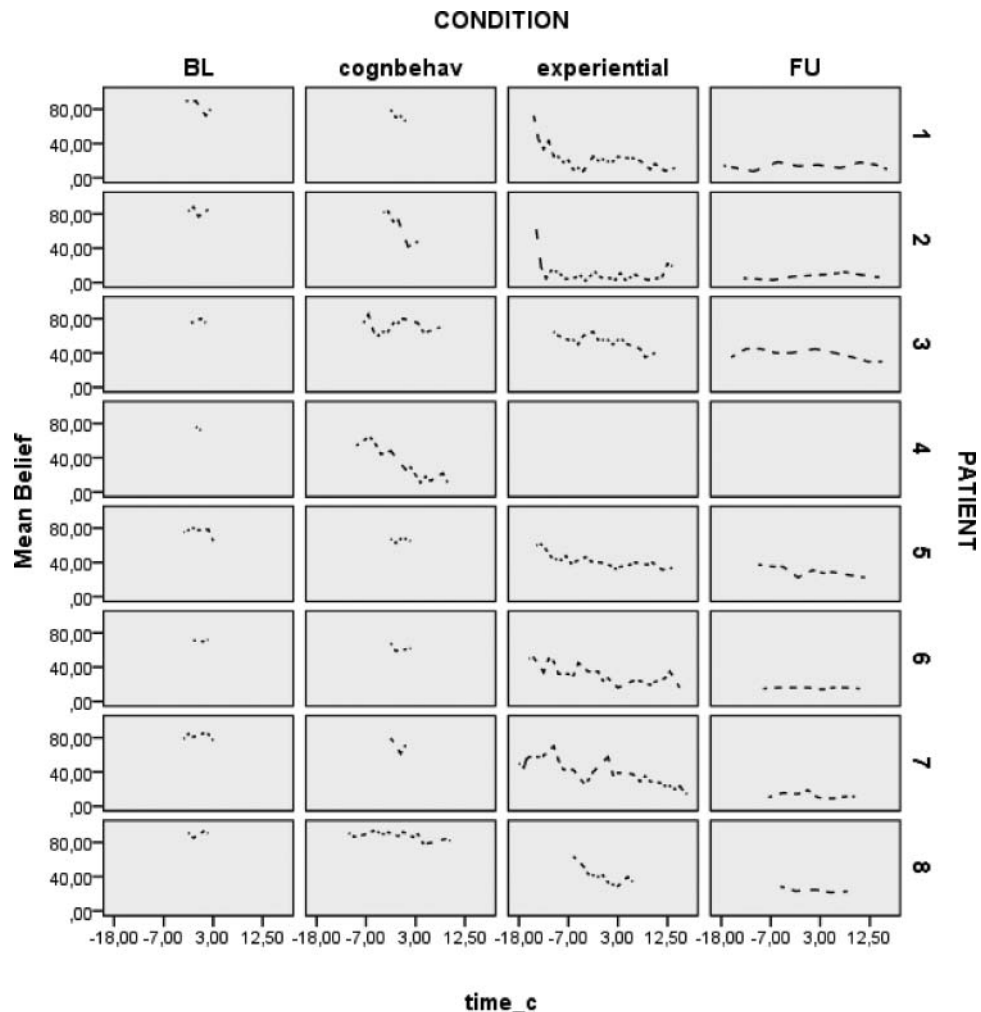

Figure 2. Individual averaged credibility ratings of core beliefs during time.

Table 2. Results of mixed regression analyses.

	Parameter	β	Std. error	df	t	p	Effect size[a] (Cohen's d)
Core beliefs	Intercept	74.678	4.934	020.397	15.135	<0.001	
	CBT phase	−10.455	2.531	171.282	−4.130	<0.001	2.136[b]
	Exp. phase	−39.515	4.275	070.895	−9.243	<0.001	5.028[b]/7.165[c]
	Follow-up	−56.146	6.031	041.700	−9.310	<0.001	0.411[b]/7.576[c]
	Time CBT	−1.117	0.348	152.619	−3.212	00.002	
	Time exp.	−1.109	0.215	074.991	−5.166	<0.001	
SCL-90	Intercept	203.00	14.294	7.000	14.202	<0.001	
	CBT phase	−26.625	13.749	7.000	−1.937	0.094	
	Exp. Phase	−47.000	12.944	7.000	−3.631	0.008	
	Follow-up	−52.125	8.961	7.000	−5.817	0.001	1.290
Target complaints	Intercept	7.519	.219	7.000	34.387	<0.001	
	CBT phase	−2.335	.644	7.000	−3.623	0.008	
	Exp. phase	−3.273	.630	7.000	−5.191	0.001	
	Follow-up	−3.626	.555	7.000	−6.535	<0.001	5.864
YSQ	Intercept	2.753	.191	7.000	14.393	<0.001	
	CBT phase	−.176	.118	7.000	−1.497	.178	
	Exp. Phase	−.526	.219	7.000	−2.391	.048	
	Follow-up	−.595	.232	7.000	−2.564	.037	1.010
WHOQOL	Intercept	81.125	6.460	7.000	12.558	<0.001	
	CBT phase	7.750	3.016	7.000	2.569	.037	
	Exp. phase	9.500	5.057	7.000	1.879	.102	
	Follow-up	11.500	3.470	7.000	3.315	0.013	.629

[a] Effect sizes calculated as change with respect to baseline, with baseline SD as denominator. Change based on the estimates from the mixed regression model.
[b] Effect size based on end of phase estimated value minus estimated baseline value, with SD based on mixed regression ARMA11 variance of the baseline values. Note that reported phase effects (CBT, experiential, follow-up) are mid-phase effects, not end-of-phase effects.
[c] Effect size based on end of phase estimated value minus end of previous phase estimated value, with SD based on mixed regression ARMA11 variance of the baseline values.

Target complaints

The individual target complaints are shown in Figure 4(c) (see Figure 4(c)). Again, visual inspection suggests a decrease in target complaints scores in all participants but one, participant 2. This participant had three target complaints and his mean target complaint's score remained high, probably because one of those was unaffected as it was determined by his comorbid Attention Deficit Hyperactivity Disorder (ADHD).

The predicted means from the analysis are shown in Figure 4 (d) (see Figure 4(d)). Effect size of treatment from baseline to follow-up was very high (5.864; see Table 2).

Quality of life

Individual scores on the WHOQOL-BREF are shown in Figure 5 (a) (see Figure 5(a)). All scores improved, except for participant
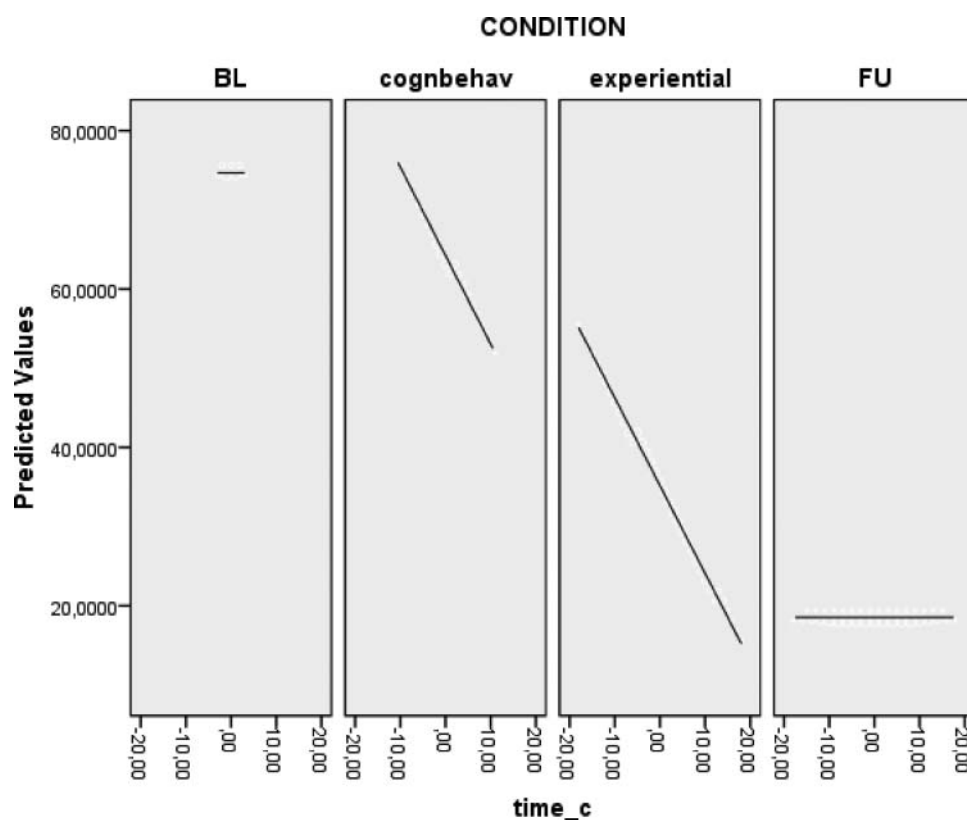

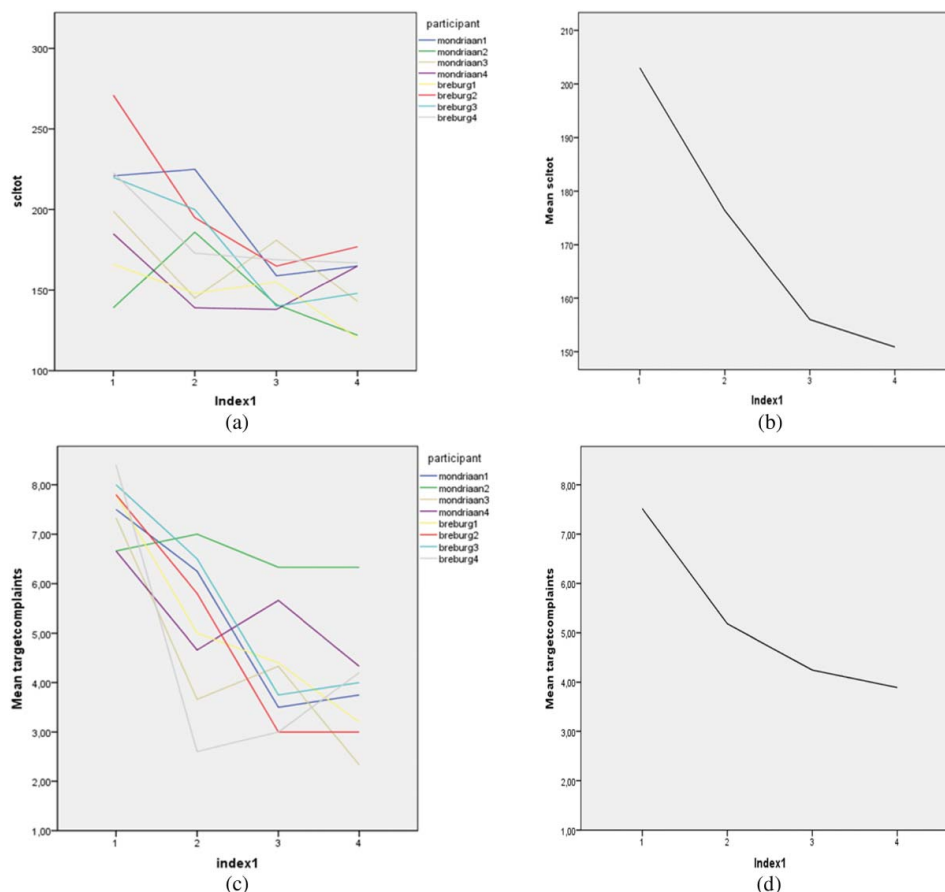**Figure 3.** Predicted means of the credibility of core beliefs.

Figure 4. Symptomatic distress and target complaints. (a) Individual SCL-90 scores during time. (b) Predicted means of the SCL-90. (c) Individual averaged target complaints during time. (d) Predicted means of averaged target complaints.

5. The predicted means from the analysis are shown in Figure 5(b) (see Figure 5(b)). Effect size of treatment from baseline to follow-up was medium (0.629; see Table 2).

Early maladaptive schemas

Scores on the YSQ are shown in Figure 5(c) (see Figure 5(c)). Visual inspection suggests that the scores of five participants improved, and the scores of three participants did not change (participants 1, 3 and 4). The predicted means from the analysis are shown in Figure 5(d) (see Figure 5(d)). Effect size of treatment from baseline to follow-up was very high (1.01; see Table 2).

Personality disorder diagnosis

All seven participants, whose PD diagnosis was assessed both at baseline and at follow-up, did not meet full criteria for a DSM PD diagnosis anymore at follow-up, again using both the general criteria for a PD and the cut-off for each specific PD for those patients who met the criteria of a specific Cluster C PD at baseline. The mean number of PD criteria decreased from baseline to follow-up from 13.71 with a SD of 2.69, to 4.57 with a SD of 2.44 ($t = 5.959$, df $= 6$, $p < 0.01$), with a very high effect size ($d = 3.56$).

Discussion

We investigated ST as a treatment for PDs in older adults, using a multiple-baseline design. We found strong effects of ST on the credibility of dysfunctional core beliefs, symptoms,

QOL and EMS. Mixed regression analyses revealed no evidence for significant time effects within baseline and follow-up phases, whereas the linear time effect during ST was strong, indicating that ST already had a positive impact on outcome during treatment. The general time effect disappeared after treatment conditions were entered into the model, indicating that it is highly unlikely that effects can be attributed to a time effect. Of the seven participants reassessed with the SCID-II at follow-up, all remitted from PD diagnosis. These results corroborated our hypotheses that ST would lead to a decrease of dysfunctional core beliefs, symptomatic distress and EMS, and an increase of QOL in cluster C PDs. Our finding that ST has a considerable positive effect on PDs in later life provides us with innovative results. This is the first study exploring the effectiveness of psychotherapy, in this case ST, for PDs as the main focus of treatment in older adults.

Participant 4, who was not reassessed concerning her PD diagnosis, did also improve concerning her core beliefs and target complaints, but she showed no improvement at follow-up on symptomatic distress, QOL and EMS. She was considered an early success by herself and the therapist and although she filled out all questionnaires, she stopped rating the credibility of her dysfunctional core beliefs, and did not cooperate in the second PD assessment. This patient was diagnosed with severe social phobia as well as avoidant PD. Possibly, the social phobia had improved, but not the underlying avoidant PD. In retrospect, she might be considered a drop-out.

Participants 1 and 3 improved on all measures, except their YSQ-scores. The initial YSQ-score of participant 1 was rather
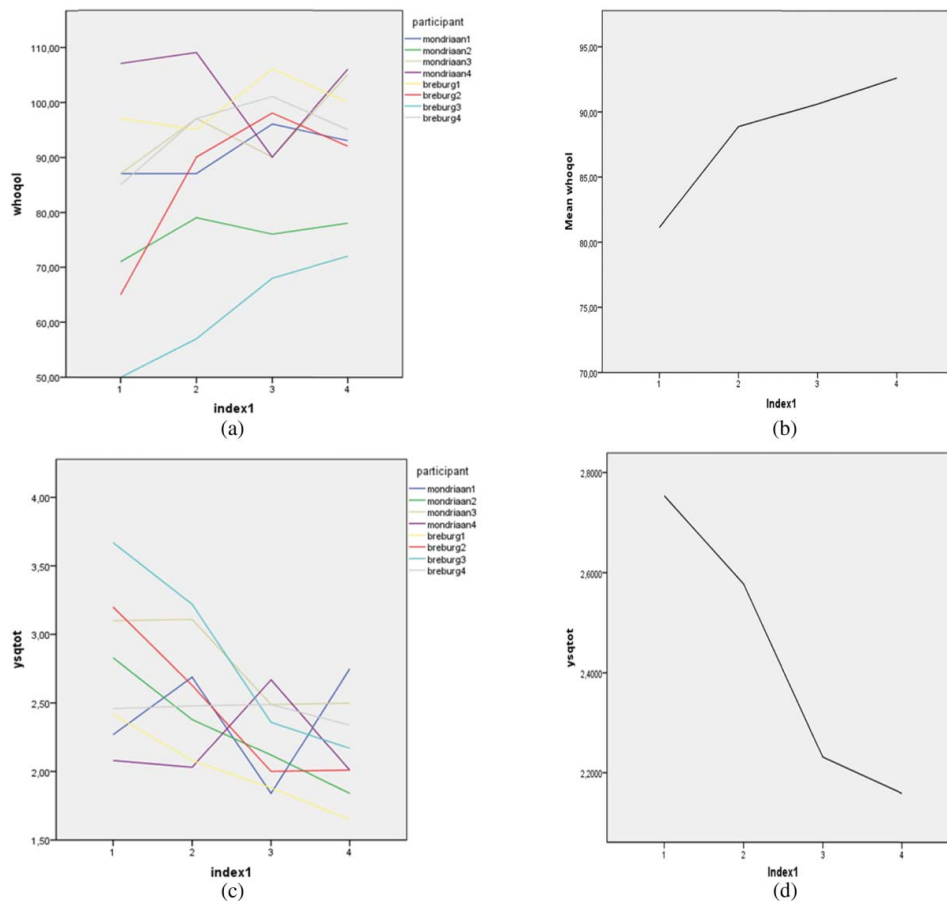
Figure 5. Quality of life and early maladaptive schemas. (a) Individual total WHOQOL-BREF scores. (b) Predicted means WHOQOL-BREF. (c) Individual total YSQ scores during time. (d) Predicted means of the YSQ.

low, possibly reflecting her avoidant coping style. In the course of treatment, she improved to a more active coping, and avoided negative feelings much less; in her own words: 'I learnt to feel my feelings more, which was hard at first.' Other studies found that treatment-related changes on the YSQ tend to be smaller than on other measures central to the patient's problems (Nadort et al., 2009). Possible explanations are that not all EMS are reported at the start of treatment, and second, some items are insensitive to change as they describe issues that cannot change (e.g. 'In my youth, …').

Some limitations of the present study should be mentioned. First, although the patients were recruited randomly from referred patients who met explicit inclusion and exclusion criteria, we cannot exclude some form of selective sampling. More research is necessary to test whether others can replicate the effects. Second, there were individual differences between responses to, and length of, the two treatment phases, the CBT and the experiential phases. In our analyses, we divided ST into two treatment phases in order to explore the effect of experiential techniques in older adults. The present data suggested that both sets of techniques contributed to the effectiveness of the full treatment. We believe this to be an advantage of the ST treatment model, as there are different techniques available to match individual aspects of a patient's problems. There was no evidence for superiority of CBT or experiential techniques, but as in younger cohorts, individual patients might differ in how much they change with these techniques (Weertman & Arntz, 2007). Third, all

participants had cluster C PDs, and we do not know whether similar effects would have been found in other PDs. As said, we chose to include cluster C as they are more stable, whereas cluster B PDs have a different expression in later life (Cooper et al., 2014; Van Alphen et al., 2015). Furthermore, three of the participants were diagnosed with PD NOS, which are usually less severe than 'pure' cluster PDs. The three participants with PD NOS in this study, however, met more PD criteria than the other participants did, so their PDs were possibly even more severe than those of the other five participants. A fourth limitation concerns the omission of assessing change in schema modes. In an earlier study into short group ST for older adults with depression and PD features, we assessed schema modes as well (Videler et al., 2014) and found a small effect size. It would have been interesting to explore whether effect sizes on change in schema modes would have been larger in this study on individual ST. We suggest to assess schema modes in future research into ST in later life, especially when studying ST in cluster B PDs. A final limitation concerns the randomization of the length of the baseline phase; it would have been better if randomization to baseline length was determined by an independent person, as we cannot exclude that characteristics of participants' presentations at the screening (e.g. severity, motivation for treatment, etc.) could have impacted the moment of the introduction of the treatment, thereby introducing bias into the data. However, start of treatment was determined by coincidence, mainly the agendas of the therapists, and checked for variance by the first author before the start of

treatment. So we do believe length of baseline was actually determined by random factors.

A possible criticism concerning this study could be its sample size of $N = 8$ which seems small to those who are not familiar with multiple-baseline designs, compared to RCTs which tend to become larger and larger because of power considerations. However, it was these same power considerations, but applied to multiple-baseline designs, which led us to choose for this sample size in the first place. In the statistical literature, it has been estimated that samples as small as $N = 4$ are sufficient to demonstrate treatment effects in multiple-baseline designs (Kazdin, 2010; Onghena, 2005). The reason for this is that the frequent assessment of the primary outcome (in this study core beliefs) and the use of each participant as his/her own control, compensates for the smaller sample size.

Some strengths of the current study also deserve to be acknowledged. All questionnaires were taken by an independent psychologist at Mondriaan and by the first author at Breburg, thus minimizing a demand effect of the participants towards their therapists. For the same reason, the ratings of the core beliefs were blind to the therapists. The findings on the self-report measures were validated by an independent assessment of the PDs, using the SCID-II. Finally, medication was constant in all participants throughout treatment, and no other treatment was allowed, so neither of them interfered with the outcomes. Of course, we cannot fully rule out external factors for change; a concurrent multiple-baseline design would be somewhat better as this further minimizes the effect of time. Of course, replication of our findings in a RCT is warranted.

Conclusions

Our study provides the first evidence for therapeutic optimism concerning the effectiveness of ST in the treatment of PDs in later life. This study also replicated a previous finding that in adult populations, on average there is no evidence for superiority of CBT or experiential techniques, but that individual patients might differ in how much they change with these techniques (Weertman & Arntz, 2007). Also in older PD patients, the effectiveness of ST was supported by using cognitive, behavioral and experiential channels to bring about change.

Disclosure statement

ORCID

Arjan C. Videler 🔾 http://orcid.org/0000-0002-2175-3453
Arnoud Arntz 🔾 http://orcid.org/0000-0002-7992-2272

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arntz, A., Sofi, D., & Van Breukelen, G. (2013). Imagery rescripting as treatment for complicated PTSD in refugees: A multiple baseline case series study. *Behaviour Research and Therapy, 51*(6), 274–283. doi:10.1016/j.brat.2013.02.009

Arntz, A., & Van Genderen, H. (2012). *Schema therapy for borderline personality disorder*. Oxford: Wiley-Blackwell.

Arrindell, W.A., & Ettema, J.H.M. (2003). *Klachtenlijst* [Symptom Checklist] (SCL-90). Lisse: Swets Test Publishers.

Bamelis, 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*, 305–322. doi:10.1176/appi.ajp.2013.12040518

Battle, C.C., Imber, S.D., Hoehn-Saric, R., Stone, A.R., Nash, E.R., & Frank, J.D. (1966). Target complaints as criteria of improvement. *American Journal of Psychotherapy, 20*, 184–192.

Cooper, L.D., Balsis, S., & Oltmanns, T.F. (2014). A longitudinal analysis of personality disorder dimensions and personality traits in a community sample of older adults: Perspectives from selves and informants. *Journal of Personality Disorders, 28*, 151–165. doi:10.1521/pedi.2014.28.1.151

David, D.O., & Freeman, A. (2014). Overview of cognitive-behavioral therapy of personality disorders. In A.T. Beck, D.D., Davis, & A. Freeman (Eds.), *Cognitive therapy of personality disorders* (3rd ed.). (pp. 3–18). New York, NY: The Guilford Press.

Edwards, D., & Arntz, A. (2012). Schema therapy in historical perspective. In M. Van Vreeswijk, J. Broersen, & M. Nadort (Eds.), *The Wiley-Blackwell handbook of schema therapy* (pp. 3–26). Chichester: Wiley & Sons.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198. doi:10.1016/0022-3956(75)90026-6

Frey, J., Heckel, R.V., Salzberg, H.C., & Wackwitz, J. (1976). Demographic variables as predictors of outcome in psychotherapy with children. *Journal of Clinical Psychology, 32*, 713–721. doi:10.1002/1097-4679 (197607)32:3<713::AID-JCLP2270320348>3.0.CO;2-W

Giesen-Bloo, J., Van Dyck, R., Spinhoven, P., Van Tilburg, W., Dirksen, C., Van Asselt, T., ... Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: Randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry, 63*, 649–658. doi:10.1001/archpsyc.63.6.649

Hawkins, N.G., Sanson-Fisher, R.W., Shakeshaft, A., D'Este, C., & Green, L.W. (2007). The multiple baseline design for evaluating population-based research. *American Journal of Preventive Medicine, 33*, 162–168. doi:10.1016/j.amepre.2007.03.020

Kalfoss, M.H., Low, G., & Molzahn, A.E. (2008). The suitability of the WHOQOL–BREF for Canadian and Norwegian older adults. *European Journal of Ageing, 5*, 77–89. doi:10.1007/s10433-008-0070-z

Kazdin, A.E. (2010). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.

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. doi:10.1002/cpp.693

Lynch, T.R., Cheavens, J.S., Cukrowitz, K.C., Thorp, S.R., Bronner, L., & Beyer, J. (2007). Treatment of older adults with co-morbid personality disorder and depression: A dialectical behavior therapy approach. *International Journal of Geriatric Psychiatry, 22*, 131–143. doi:10.1002/gps.1703

Nadort, M., Arntz, A., Smit, J.H., Giesen-Bloo, J., Eikelenboom, M., Spinhoven, P., ... 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. doi:10.1016/j.brat.2009.07.013

Onghena, P. (2005). Single-case designs. In D.C. Howell & B.S. Everitt (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 4, pp. 1850–1854). Chichester: Wiley.

Powers, A., Strube, M.J., & Oltmanns, T.F. (2014). Personality pathology and increased use of medical resources in later adulthood. *The American Journal of Geriatric Psychiatry : Official Journal of the American Association for Geriatric Psychiatry, 22*(12), 1478–1486. Retrieved from http://doi.org/10.1016/j.jagp.2013.10.009

Rafaeli, E., Bernstein, D., & Young, J. (2011). *Schema therapy*. Hove, East Sussex: Routledge.

Rijkeboer, M.M., & Van den Berg, H. (2006). Multiple group confirmatory factor analysis of the young schema-questionnaire in a Dutch clinical versus non-clinical population. *Cognitive Therapy and Research, 30*, 263–278. doi:10.1007/s10608-006-9051-8

Rosowsky, E., Young, A.S., Malloy, M.C., Van Alphen, S.P.J., & Ellison, J.M. (2016). A cross-validation Delphi method approach to the diagnosis and treatment of personality disorders in older adults. *Aging and Mental Health*, advance online publication. doi:10.1080/13607863.2016.1261796

Schuster, J.P., Hoertel, N., Le Strat, Y., Manetti, A., & Limosin, F. (2013). Personality disorders in older adults: Findings from the national epidemiologic survey on alcohol and related conditions. *American Journal of Geriatric Psychiatry, 21*, 757–768. doi:10.1016/j.jagp.2013.01.055

Segal, D.L., Coolidge, F.L., & Rosowsky, E. (2006). *Personality disorders and older adults: Diagnosis, assessment, and treatment*. New York, NY: Guilford Press.

Shorer, C. (1970). Improvement with and without psychotherapy. *Diseases of the Nervous System, 31*, 155–160.

Sterk, F., & Rijkeboer, M. (1997). *Schema-vragenlijst* [Schema Questionnaire]. Utrecht: Ambulatorium Universiteit Utrecht.

Trompenaars, F.J., Masthoff, E.D., Van Heck, G.L., Hodiamont, P.P., & De Vries, J. (2005). Content validity, construct validity, and reliability of the WHOQOL-Bref in a population of Dutch adult psychiatric outpatients. *Quality of Life Research, 14*, 151–160. doi:10.1007/s11136-004-0787-x

Van Alphen, S.P.J., Bolwerk, N., Videler, A.C., Tummers, J.H.A., Van Royen, R.J.J., Barendse, H.P.J., … Rosowsky, E. (2012). Age-related aspects and clinical implications of diagnosis and treatment of personality disorders in older adults. *Clinical Gerontologist, 25*, 27–41. doi:10.1080/07317115.2011.628368

Van Alphen, S.P.J., Van Dijk, S.D.M., Videler, A.C., Rossi, G., Dierckx, E., Bouckaert, F., & Oude Voshaar, R.C. (2015). Personality disorders in older adults: Emerging research issues. *Current Psychiatry Reports, 17*, 538–545. doi:10.1007/s11920-014-0538-9

Van Alphen, S.P.J., Derksen, J.J.L., Sadavoy, J., & Rosowsky, E. (2012). Features and challenges of personality disorders in late life. *Aging and Mental Health, 16*, 805–810. doi:10.1080/13607863.2012.667781

Videler, A.C., Rossi, G., Schoevaars, M., Van der Feltz-Cornelis, C.M., & Van Alphen, S.P.J. (2014). Effects of schema group therapy in older outpatients: A proof of concept study. *International Psychogeriatrics, 26*, 1709–1717. doi:10.1017/S1041610214001264

Videler, A.C., Van der Feltz-Cornelis, C.M., Rossi, G., Van Royen, R.J.J., Rosowsky, E., & Van Alphen, S.P.J. (2015). Psychotherapeutic treatment levels of personality disorders in older adults. *Clinical Gerontologist, 38*, 325–341. doi:10.1080/07317115.2015.1032464

Videler, A.C., van Royen, R.J.J., van Alphen, S.P.J., Rossi, G., & van der Feltz-Cornelis, C.M., (2017). Adapting schema therapy for personality disorders in older adults. *International Journal of Cognitive Therapy*, *10*, 62–78 doi: 10.1521/ijct.2017.10.1.62

Vlaeyen, J., De Jong, J., Geilen, M., Heuts, P., & Van Breukelen, G. (2001). Graded exposure in vivo in the treatment of pain-related fear: A replicated single-case experimental design in four patients with chronic low back pain. *Behaviour Research and Therapy, 39*, 151–166. doi:10.1016/S0005-7967(99)00174-6

Weertman, A., & Arntz, A. (2007). Effectiveness of treatment of childhood memories in cognitive therapy for personality disorders: A controlled study contrasting methods focusing on the present and methods focusing on childhood memories. *Behaviour Research and Therapy, 45*, 2133–2143. doi:10.1016/j.brat.2007.02.013

Weertman, A., Arntz, A., & Kerkhofs, M.L.M. (2000). *Gestructureerd diagnostisch interview voor DSM-IV persoonlijkheidsstoornissen (SCID-II)* [Structured diagnostic interview for DSM-IV personality disorders]. Lisse: Swets Test Publishers.

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