

Tilburg University

Systems training for emotional predictability and problem solving for borderline personality disorder

Ekiz, Erol; Van Alphen, Sebastiaan P. J.; Ouwens, Machteld A.; Van De Paar, Jamie; Videler, Arjan C.

Published in:
Personality and Mental Health

DOI:
[10.1002/pmh.1558](https://doi.org/10.1002/pmh.1558)

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Ekiz, E., Van Alphen, S. P. J., Ouwens, M. A., Van De Paar, J., & Videler, A. C. (2022). Systems training for emotional predictability and problem solving for borderline personality disorder: A systematic review. *Personality and Mental Health*. <https://doi.org/10.1002/pmh.1558>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Systems Training for Emotional Predictability and Problem Solving for borderline personality disorder: A systematic review

Erol Ekiz^{1,2}  | Sebastiaan P. J. van Alphen^{2,3,4,5} | Machteld A. Ouwens^{1,2} |
Jamie Van de Paar¹ | Arjan C. Videler^{1,2}

¹PersonaCura, Clinical Center of Excellence for Personality Disorders and Autism in Older Adults, GGz Breburg, Tilburg, The Netherlands

²Tranzo, Scientific Centre for Care and Wellbeing of the Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

³Clinical Center of Excellence for Personality Disorders in Older Adults, Mondriaan Mental Health Center, Heerlen-Maastricht, The Netherlands

⁴Personality and Psychopathology Research Group (PEPS), Department of Psychology (PE), Vrije Universiteit Brussel (VUB), Brussels, Belgium

⁵Department of Medical and Clinical Psychology, Tilburg University, Tilburg, The Netherlands

Correspondence

Erol Ekiz, PersonaCura, Clinical Center of Excellence for Personality Disorders and Autism in Older Adults, GGz Breburg, Postbus 770, 5000 AT, Tilburg, The Netherlands.

Email: e.ekiz@ggzbreburg.nl

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Abstract

Systems Training for Emotional Predictability and Problem Solving (STEPPS) is a group treatment program for patients with borderline personality disorder (BPD). The program was intended to be highly accessible, both for patients and therapists. During STEPPS, patients are taught emotion regulation and behavior management skills. This systematic review synthesizes the current empirical status of STEPPS, focusing on research designs, quality of studies, target groups, protocols, and outcome. We selected 20 studies, with three randomized controlled trials. Patients with BPD, subthreshold BPD, and patients with BPD and comorbid antisocial personality disorder were investigated. One study was conducted in adolescents. There were no studies in older adults. Results demonstrated STEPPS to be associated with reduced BPD symptoms, improved quality of life, decreased depressive symptoms, and decreased negative affectivity. Mixed results were found for impulsivity and suicidal behaviors. STEPPS has both been studied as an add-on therapy to patients' ongoing treatment, and, with the addition of individual STEPPS sessions, as a stand-alone treatment. High attrition rates were found in patients attending STEPPS, complicating the generalizability of the results. Although the evidence for STEPPS is promising, further research is needed before firm conclusions can be drawn. Recommendations for future research are discussed.

INTRODUCTION

The term borderline personality was introduced by American psychoanalyst Adolph Stern in 1938 to refer to a group of patients who did not seem to fit into the personality classifications described until then (National Collaborating Centre for Mental Health, 2009). This group of patients constituted a “borderline” case between psychotic and psychoneurotic patients. Over the years, the diagnostic label evolved into the classification of a chronic psychiatric disorder, first described in the third edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1980). Ellison et al. (2018) report the prevalence of borderline personality disorder (BPD) in the general population is about 1%, rising up to 12% in outpatients and 22% in inpatients. BPD is mainly characterized by a pattern of unstable interpersonal relationships, self-image and affects, and marked impulsivity (American Psychiatric Association, 2013). Because BPD is associated with severe functional impairment (Skodol et al., 2005) and lower quality of life (Cramer et al., 2006), and because managing the symptoms of BPD is challenging, both for patients and their environment (Bailey & Grenyer, 2013), as well as for therapists (Cleary et al., 2002; Ring & Lawn, 2019), much more research has been conducted into BPD than into other personality disorders (PDs). Various forms of psychotherapy and pharmacotherapy have been found to be beneficial in treating (core symptoms of) BPD (Cristea et al., 2017; Stoffers-Winterling et al., 2020; Storebø et al., 2020). One of those treatments is the Systems Training for Emotional Predictability and Problem Solving (STEPPS; Blum et al., 2012).

STEPPS is a group treatment program that was developed in 1995 at the University of Iowa (Blum et al., 2012). STEPPS was based on a skills training program for patients with BPD by Bartels and Crotty (1998). As traditional modes of therapy at the time had not been helpful in reducing deliberate self-harm and acting out behaviors, nor in reducing high hospitalization rates in BPD patients, the main objectives of STEPPS were to reduce hospitalization (in terms of length and rate) and self-harm acts (Blum et al., 2002). In order to achieve these objectives, patients' ongoing (individual) therapy was supplemented with STEPPS. Therefore, unlike other forms of evidence-based psychotherapies for BPD, such as dialectical behavior therapy (DBT; Linehan et al., 1991), mentalization-based treatment (Bateman & Fonagy, 1999), and schema therapy (Young et al., 2003), STEPPS was not developed as a stand-alone treatment. STEPPS was intended to be a highly accessible and easily implementable treatment program for professionals, easily fitted into the lives of the patients, and in which new skills could easily be learned by the patient and taught by

the therapist each session (Blum et al., 2002). Despite its main focus on improving emotion regulation skills, at the time, STEPPS was not based on an emotion regulation theory (e.g., the process model of emotion regulation by Gross, 1999). Instead, it was built on learning theory and working mechanisms of cognitive behavioral therapy for the target population (Blum et al., 2002; Bos et al., 2010). As STEPPS was designed to be cost effective and accessible, the STEPPS program aimed to provide a pragmatic answer to the objectives and to the specific setting and needs of patients and professionals (Blum et al., 2002).

The STEPPS program comprises 20 weekly group meetings of 2 h in a classroom format. The group consists of six to 10 participants and is facilitated by two trained therapists. Similarly to DBT, in STEPPS, it is assumed that patients lack skills to deal with their mounting emotions, and therefore, the training is highly structured and focuses on acquiring skills. The weekly agenda is incorporated in a manual, which forms the basis of the program, and participants are encouraged to do their weekly homework assignments. STEPPS consists of three distinct components (Black et al., 2009; Blum et al., 2002). The first component is psychoeducation about BPD, in which participants are taught information about BPD and emotion regulation. Because patients tend to regard the term BPD as pejorative and stigmatizing, and they tend to resist the diagnosis, BPD is reframed as an emotional intensity disorder (Blum et al., 2008). Furthermore, participants become aware of their capacity to learn new skills. The second component comprises an emotion regulation skills training. Participants learn five skills (i.e., distancing, communication, challenging thoughts, distracting, and managing problems) to cope with intense emotions and manage their emotional intensity disorder. The last component includes a behavior management skills training. Participants are motivated to master different skill areas and lifestyle behaviors (i.e., goal setting, healthy eating, sleep hygiene, physical exercise, leisure behaviors, health monitoring, interpersonal effectiveness, and avoiding abusive behaviors), as they provide directions to life and help prevent emotion regulation problems to arise. Furthermore, participants are asked to identify key family members, professionals, and friends, as part of their “reinforcement team”; these individuals are invited to participate in separate educational sessions and to help reinforce and support newly acquired skills in daily life.

Over the years, STEPPS has been introduced in different countries, and multiple studies have investigated its efficacy. However, to date, there is no systematic review of the available literature covering the current evidence for the effectiveness of STEPPS. The aim of the present review is to fill this gap and provide answers to the following questions:

- a. What (types of) empirical studies on STEPPS are described in the international literature?
- b. What is the quality of these studies?
- c. Which patient groups (age, gender, comorbid diagnoses) receiving STEPPS have been studied?
- d. What are main differences and similarities in the international STEPPS protocols and manuals?
- e. Which outcome variables are investigated, and what are the effect sizes?

METHOD

Search strategy

The protocol for this systematic review was preregistered with PROSPERO (ID = CRD42021253254). The review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). As search strategy and study selection for this review, we searched the electronic databases of PsycInfo, Embase, and Medline within the OVID interface. We used the following search term for the title, abstract, and keyword heading: “Systems Training for Emotional Predictability and Problem Solving.” The search was completed in May 2021, and an updated search was conducted in October 2021. Titles and abstracts of the studies were screened for eligibility by the first and fourth author independently. Discrepancies were resolved through discussion and by reading the full text of the study. Studies failing to reach consensus were submitted to the second and last author to reach a final verdict.

Inclusion and exclusion criteria

We considered published studies concerning STEPPS. Studies were found eligible for this review if they met the following criteria: (1) The study is published in an international peer-reviewed journal. (2) The study contains an original research question or design and original empirical outcome (reviews, meta-analyses, conference abstracts, comments, replies, and editorials were excluded). (3) Study participants were treated with STEPPS. We did not exclude studies based on date of publication.

Quality assessment tool

There is no international consensus on which rating scale to use in assessing the quality of a heterogeneous group of (non)experimental study designs. As there is no general quality assessment tool for different designs (i.e., the

tool should award extra points to a randomized controlled trial [RCT] vs. a cohort study, while also assessing the quality of one RCT above the other), we chose to formulate 10 items based on two existing quality assessment tools (i.e., five items from Barnicot et al., 2012 and five items from the Quality Assessment Tool for Quantitative Studies of Effective Public Healthcare Panacea Project [EPHPP]; Thomas et al., 2004). Items of our quality assessment tool are presented in Table 1. Studies are awarded 0, 1, or 2 points on these items. The total score is divided by the number of items applicable, resulting in a quality indication score between 0 and 1. The level of quality is qualified as follows: poor (0–0.25), fair (0.25–0.5), good (0.5–0.75), and excellent (0.75–1). The studies included were independently assessed by the first and fourth author. Discrepancies were supervised by the last author to reach a final verdict.

RESULTS

Study characteristics

See Figure 1 for the PRISMA flow diagram. The search strategy yielded 86 potential inclusions. After screening titles and abstracts, the first and fourth authors agreed on the inclusion of 18 studies. After reading the full text, there was no consensus on two other studies. These two studies were included after review by the second and last author. Thus, the search yielded a total of 20 included studies. Study characteristics are reported in Table 2.

All included studies were from Western countries: United States ($N = 9$), Spain ($N = 4$), the United Kingdom ($N = 3$), Italy ($N = 2$), and the Netherlands ($N = 2$). All studies from the United States were conducted in the state Iowa. Three RCTs have been performed on STEPPS (Blum et al., 2008; Bos et al., 2010; Bos et al., 2011), with all three reporting follow-up data. Eleven (uncontrolled) pretreatment versus posttreatment research designs were found (Alesiani et al., 2014; Black et al., 2008, 2013; Blum et al., 2002; Boccalon et al., 2017; González-González et al., 2021; Guillén et al., 2021; Harvey et al., 2010; Hezelyova et al., 2021; Hill et al., 2016; Llorens Ruiz et al., 2020). One study compared STEPPS with another form of psychotherapy (i.e., DBT; Guillén Botella et al., 2021). Finally, there were five secondary data analyses reported (Black et al., 2009, 2011, 2018a, 2018b, 2016). In these five studies, data from two original studies (i.e., Black et al., 2013; Blum et al., 2008) were reused to answer newly described research questions.

Five of the 20 included studies investigated pre- versus post-STEPPS measures of forensic offenders from

TABLE 1 Quality assessment tool

Item number	Item category	Item description	Ratings	Based on
1	Sample	Was the sample size representative?	0 = $N < 30$ 1 = $30 \leq N < 100$ 2 = $N \geq 100$	Barnicot et al. (2012)
2	Measurements	Were reliable structured interviews used to diagnose BPD?	0 = unknown or not used 1 = used, assessors were not blinded to research question, or blinding unknown 2 = used and assessors were blinded to research question	Barnicot et al. (2012)
3	Measurements	Were validated and reliable outcome measures used?	0 = mostly not validated and reliable (0% to 24%) 1 = some measures validated and reliable (25% to 74%) 2 = most measures validated and reliable (75% to 100%)	Barnicot et al. (2012)
4	Design	Was the outcome assessor blinded to the treatment arm?	n/a = no control condition 1 = not blinded/only self-report 2 = blinded	Barnicot et al. (2012)
5	Results	Were outcome distribution checks performed and appropriate analyses used?	0 = distribution not checked and inappropriate model used 1 = distribution not checked/unknown or inappropriate model used 2 = distribution checked and appropriate model used	Barnicot et al. (2012)
6	Sample	Are the individuals selected to participate in the study likely to be representative of the target population?	0 = not likely 1 = somewhat likely 2 = very likely	EPHPP
7	Design	What was the study design?	0 = other methods 1 = interrupted time series, cohort (one group pre + post) (before and after), case-control, cohort analytic (two groups pre + post) 2 = controlled clinical trial, randomized controlled trial	EPHPP
8	Design	Was the study appropriately randomized?	n/a = single group study 1 = not appropriately randomized 2 = yes	EPHPP
9	Results	Were withdrawals and dropouts reported in terms of numbers and reasons per group?	0 = no 1 = only numbers were reported 2 = no dropouts, or dropout numbers and reasons were reported	EPHPP
10	Results	What was the percentage of participants completing the study?	0 = unknown, or less than 60% 1 = 60% to 79% 2 = 80% to 100%	EPHPP

Abbreviation: EPHPP, Effective Public Healthcare Panacea Project.

correctional settings in Iowa (Black et al., 2018a, 2018b, 2008, 2013, 2016). One study reported differences between treatment responders and nonresponders (Black et al., 2009). Furthermore, one study investigated

differences between a group of participants who attended the STEPPS program and participants who declined participation to STEPPS or attended only one group session (González-González et al., 2021).

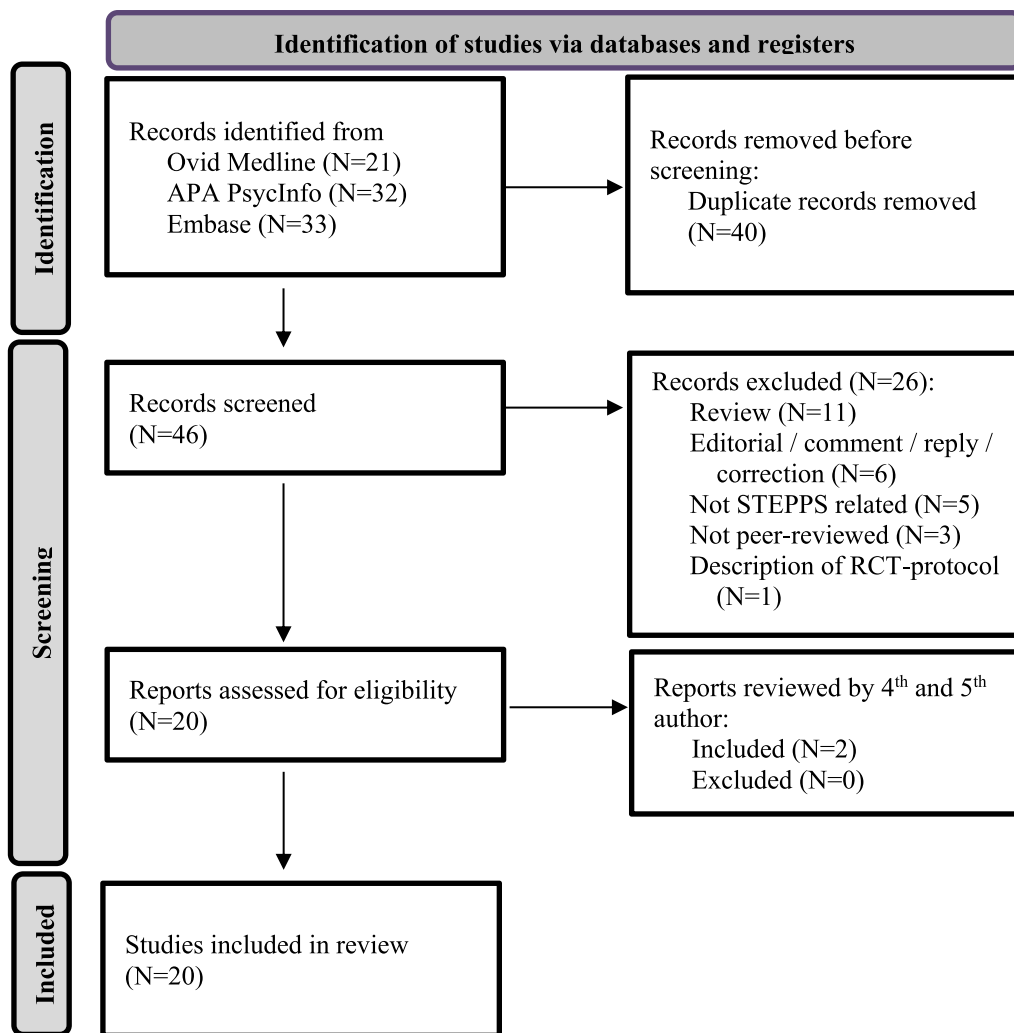


FIGURE 1 PRISMA flow diagram

Quality assessment

Table 2 presents the quality assessment scores of the studies that were included. The average quality rating for the studies was good (mean = 0.55, SD = 0.12). An overall methodological strength was that most used measurements were validated and reliable. In 14 studies, the diagnoses of BPD were established using validated diagnostic interviews: Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II; First et al., 1997), Structured Interview for DSM-IV Personality Disorders (SIDP-IV; Pfohl et al., 1997), and Diagnostic Interview for Borderline Disorder-Revised (DIB-r; Zanarini et al., 1989). In the remaining studies, BPD diagnoses were assessed through unspecified methods ($N = 4$) or using Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD; Zanarini, 2003; $N = 2$). In contrast to SCID-II, SIDP-IV, and DIB-r, which are developed to diagnose BPD, the ZAN-BPD is developed to assess severity and

change of BPD symptoms. Furthermore, most of the outcome measurements used were proven to be reliable and validated for the target population (Blum et al., 2002; Crawford & Henry, 2004; Steer et al., 2001). Most commonly used questionnaires were Borderline Evaluation of Severity Over Time (BEST; Pfohl et al., 2009; $N = 11$), Beck Depression Inventory (BDI; Beck, 1978) or Beck Depression Inventory II (BDI-II; Beck et al., 1996; $N = 10$), and Positive and Negative Affect Schedule (PANAS; Watson & Clark, 1994; $N = 7$).

The most common methodological problem, however, was related to study completion. The cumulative number of participants intended to receive treatment in the included studies was 1981. However, in addition to the five secondary data analyses, preliminary data from Black et al. (2008) were also included in Black et al. (2013). Furthermore, the two Dutch studies (Bos et al., 2010, 2011) had overlapping study participants. Considering this, there were 1162 unique participants, of which

TABLE 2 Treatment of BPD with STEPPS (study characteristics)

Study	Study design	N, total (N, STEPPS)	Gender, % female	Age, M in years (range)	Patient diagnoses	Comparison	Attrition, %	Protocol	Quality assessment
Alesiani et al. (2014)	Cohort	32	81%	44.4 (26–63)	PD	Pretreatment vs. posttreatment	47%	STEPPS (Italy) Changes in protocol	0.438 Fair
Black et al. (2008)	Cohort	12	100%	34.8 (19–49)	BPD (N = 8) BPD and ASPD (N = 4)	Pretreatment vs. posttreatment	17%	STEPPS (USA)	0.688 Good
Black et al. (2009)	Secondary data analysis	164	85%	Unknown	BPD	Treatment responders vs. nonresponders	Unknown	STEPPS (USA)	0.438 Fair
Black et al. (2011)	Secondary data analysis	165	85%	30.9	BPD	Pretreatment vs. posttreatment	45%	STEPPS (USA)	0.375 Fair
Black et al. (2013)	Cohort	77	82%	31.4 (19–50)	BPD (N = 20) BPD and ASPD (N = 44)	Pretreatment vs. posttreatment	47%	STEPPS (USA)	0.5 Fair
Black et al. (2016)—substudy I	Secondary data analysis	65	80%	31.4 (19–56)	BPD (N = 49) BPD and ASPD (N = 16)	BPD vs. BPD and ASPD	Not mentioned	STEPPS (USA)	0.556 Good
Black et al. (2016)—substudy II	Secondary data analysis	64	78%	30.9 (19–50)	BPD (N = 20) BPD and ASPD (N = 44)	BPD vs. BPD and ASPD	48%	STEPPS (USA)	0.556 Good
Black et al. (2018a)	Secondary data analysis	77	82%	31.4 (19–50)	BPD	Pretreatment vs. posttreatment	47%	STEPPS (USA)	0.563 Good
Black et al. (2018b)	Secondary data analysis	193	82%	Unknown	BPD	Pre- vs. post-STEPPS/ TAU/STEPPS + TAU	36%	STEPPS (USA)	0.625 Good
Blum et al. (2002)	Cohort	52	94%	33 (18–51)	BPD	Pretreatment vs. posttreatment	Unknown	STEPPS (USA)	0.375 Fair
Blum et al. (2008)	RCT	165 (93)	80% ^a	31.4 ^a	BPD	STEPPS + TAU vs. TAU	52% ^a	STEPPS (USA)	0.65 Good
Boccalon et al. (2017)	Cohort	24	83%	41.0	PD	Pretreatment vs. posttreatment	29%	STEPPS (Italy)	0.563 Good
Bos et al. (2010)	RCT	79 (42)	83% ^a	32.9 ^a	BPD	STEPPS vs. TAU	21% ^a	STEPPS (Netherlands) changes in protocol	0.8 Excellent
Bos et al. (2011)	RCT	168 (84)	88% ^a	33.5 ^a	BPD (N = 83)	STEPPS vs. TAU	26% ^a	STEPPS (Netherlands) changes in protocol	0.75 Good

TABLE 2 (Continued)

Study	Study design	N, total (N, STEPPS)	Gender, % female	Age, M in years (range)	Patient diagnoses	Comparison	Attrition, %	Protocol	Quality assessment
González-González et al. (2021)	Cohort	118	84% ^a	33.6 ^c (18–58)	BPD	Pretreatment/control vs. posttreatment/control and STEPPS vs. control group	56%	STEPPS (Spain) changes in protocol	0.45 Fair
Guillén et al. (2021)	Cohort	202 (32)	81% ^d	28.9 ^d	BPD	Pretreatment vs. posttreatment	28% ^a	STEPPS (Spain)	0.6 Good
Guillén Botella et al. (2021)	Nonrandomized clinical trial	72 (27)	100% ^a	32.1 ^b (14–60)	BPD	Pretreatment vs. posttreatment and STEPPS vs. DBT	31% ^a	STEPPS (Spain) changes in protocol	0.55 Good
Harvey et al. (2010)	Cohort	38	84%	37 (20–58)	BPD	Pretreatment vs. posttreatment	18%	STEPPS (USA) changes in protocol	0.625 Good
Hezelyova et al. (2021)	Cohort	148	80%	31.0 (18–73)	BPD traits	Pretreatment vs. posttreatment	45%	STEPPS EI Changes in protocol	0.438 Fair
Hill et al. (2016)	Cohort	45	84%	34.4 (19–59)	BPD	Pretreatment vs. posttreatment	33%	STEPPS (USA)	0.438 Fair
Llorens Ruiz et al. (2020)	Cohort	21	90%	15.5 (13–17)	BPD traits	Pretreatment vs. posttreatment	19%	STEPPS (Spain) changes in protocol	0.563 Good

Abbreviations: ASPD, antisocial personality disorder; BPD, borderline personality disorder; DBT, dialectical behavior therapy; PD, personality disorder; STEPPS, Systems Training for Emotional Predictability and Problem Solving; STEPPS EI, Systems Training for Emotional Predictability and Problem Solving Early Intervention; TAU, treatment as usual.

^aIn STEPPS group.

^bIn STEPPS and DBT groups.

^cIn STEPPS and control groups.

^dIn STEPPS, DBT, and TAU groups.

771 participants were allocated to STEPPS. The other participants were allocated to DBT ($N = 166$), treatment as usual (TAU; $N = 199$), a control group formed by participants who declined participation to STEPPS or attended only one group session ($N = 20$), or did not receive any treatment ($N = 6$). The cumulative dropout (i.e., participants who agreed to participate in the study but did not complete the study for any reason) was $N = 416$ (35.8%). Mean dropout rate of participants allocated to STEPPS was 42.0%. Three studies reported dropout reasons (Black et al., 2008; Bos et al., 2010, 2011). These reasons were dissatisfaction ($N = 8$), attended less than 12 sessions ($N = 5$), problems with group ($N = 3$), wrong treatment ($N = 3$), released from prison (in the prison sample; $N = 2$), removal ($N = 2$), and job-related reasons ($N = 1$). There were six studies in which patients received individual STEPPS sessions besides the original group sessions. In these studies, the attrition rate for patients attending STEPPS was 36.1%.

Patient characteristics

The majority of the study participants were female (range: 78% to 100%). Accordingly, the prison sample reached a high percentage of female participants (82%). The age of the participants ranged between 13 and 73 years, with an average of approximately 33 years. In one study, the mean age of the participants was 15.5 years (Llorens Ruiz et al., 2020). This is the only study conducted in a sample of adolescents. There were no studies conducted in older adults.

Out of the 1162 unique study participants, 594 (51.1%) were diagnosed with a PD with the use of a reliable structured interview (i.e., SCID-II, SIDP-IV, or DIB-r). Five hundred fifty-nine (94.1%) of them had BPD with 44 participants having a comorbid antisocial personality disorder (ASPD), and 35 (5.9%) of them had another PD. Ninety-five (8.2%) participants, who were assessed with the use of mentioned interviews, had no BPD. However, the outcome of the assessments were not disclosed, so they might have had other PDs. Out of the 1162 unique participants, another 325 (28.0%) participants were diagnosed with BPD without the use of a reliable structured interview designed to diagnose the presence of a PD. Finally, the remaining 148 (12.7%) out of the 1162 unique participants had threshold BPD or BPD traits, as assessed with an unidentified screening measure. The ratio, however, was not disclosed.

Changes in the STEPPS protocol

The original STEPPS protocol was investigated in 12 studies, most of them in the United States. The original protocol

consists of 20 weekly 2-h group sessions, and at least one session for the reinforcement team, added to individual TAU.

Major changes in the STEPPS protocol were made in the studies of Alesiani et al. (2014), Bos et al. (2010), Bos et al. (2011), and González-González et al. (2021). In Alesiani et al. (2014), the length of group sessions was 45 min twice weekly, the length of treatment was 6 to 8 months, and the group setting was an open group. Bos et al. (2010, 2011) used the Dutch STEPPS, abbreviated to VERS (*Vaardigheidstraining Emotie Regulatie Stoornis*). In VERS, the length of treatment is adjusted to 18 weekly sessions and one follow-up session after 3 to 6 months. Most importantly, structured individual sessions once every fortnight were added to the protocol. These individual sessions followed the VERS protocol and were intended to reinforce newly acquired skills. So, in the Netherlands, VERS is offered in an individual, group, and systems package, making it more of a stand-alone treatment for BPD. Finally, in the study of González-González et al. (2021), authors used adapted STEPPS content, participants received monthly individual psychotherapy sessions for 18 months, and five group psychotherapy sessions for the client's reinforcement team were added.

Other changes to the STEPPS protocol were made in the study of Harvey et al. (2010). In this study, participants had one-on-one skills reinforcement sessions on a weekly basis. Guillén Botella et al. (2021) made a similar adjustment. In this study, participants received weekly meetings with a clinician. Llorens Ruiz et al. (2020) used a STEPPS model for adolescents. Participants maintained their TAU, which consisted of individual therapy every 3 weeks and pharmacological treatment.

Furthermore, Hezelyova et al. (2021) studied the shortened, less-intensive 13-week STEPPS Early Intervention for participants with emotional intensity difficulties in primary care facilities in the United Kingdom.

Outcome in STEPPS studies

Baseline demographics

Baseline characteristics, such as gender, age, and ethnicity, were not associated with greater improvement after treatment (Black et al., 2018a). Nonetheless, higher educational level, better patient collaboration, higher baseline symptom severity, and absence of a bipolar disorder did predict better outcome (González-González et al., 2021; Hezelyova et al., 2021). Analyses of dropouts demonstrate no significant differences in sex, school level, and DSM diagnoses between dropouts and completers (Alesiani et al., 2014). Higher impulsivity, however, was associated with early discontinuation (Black et al., 2009).

TABLE 3 Treatment of BPD with STEPPS (outcome)

Study	Comparison	BPD symptoms (ES)	Global severity (ES)	Depression and anxiety (ES)	Others (ES)
Alesiani et al. (2014)	Pretreatment vs. posttreatment	- EIC* - FQ - Number of hospitalizations* - Number of suicide attempts*	-	-	-
Black et al. (2008)	Pretreatment vs. posttreatment	- BEST (0.73)*	-	- BDI (0.86)*	- PANAS PA (-0.05) - PANAS NA (0.41)*
Black et al. (2009)	Treatment responders vs. nonresponders	- ZAN-BPD (-0.33) - BEST (-0.02) - BIS (0.08)	- CGI Sev. (-0.13) - GAS (-0.18) - SCL-90 (0.12) - SAS (0.01)	- BDI (d = 0.06)	- Number of sessions attended (0.67)* - Treatment group (phi = 0.45)*
Black et al. (2011)	Pretreatment vs. posttreatment	-	- Medication usage - Number of outpatient health services received - Number of other treatments received	-	-
Black et al. (2013)	Pretreatment vs. posttreatment	- BEST (1.30)* - Suicidal behaviors*	-	- BDI (1.08)*	- PANAS PA (0.26) - PANAS NA (0.69)* - Disciplinary infractions*
Black et al. (2016)—substudy I	BPD vs. BPD and ASPD ^a	- ZAN-BPD (-0.71)* - BEST (-0.20) - BIS (-0.88)*	- CGI Sev. (-0.16) - CGI Impr. (-1.99)* - GAS (0.33) - SCL-90-R (-0.03) - SAS (-0.20) - PGSR (-0.91)	- BDI (-0.37)	- PANAS PA (0.02) - PANAS NA (-0.17)
Black et al. (2016)—substudy II	BPD vs. BPD and ASPD ^a	- BEST Total (-0.74)*	-	- BDI (-0.45)	- PANAS PA (0.89)* - PANAS NA (-0.29)

(Continues)

TABLE 3 (Continued)

Study	Comparison	BPD symptoms (ES)	Global severity (ES)	Depression and anxiety (ES)	Others (ES)
Black et al. (2018a)	Pretreatment vs. posttreatment	<ul style="list-style-type: none"> - BEST all groups (1.31)* - BEST female (1.30)* - BEST male (1.65)* - BEST minority (1.16)* - BEST Caucasian (1.36)* - BEST age ≥40 (1.32)* - BEST age <40 (1.32)* 	-	<ul style="list-style-type: none"> - BDI all groups (1.08)* - BDI female (1.13)* - BDI male (0.69) - BDI minority (0.86)* - BDI Caucasian (1.12)* - BDI age ≥40 (1.21)* - BDI age <40 (1.05)* 	<ul style="list-style-type: none"> - PANAS PA all groups (0.29) - PANAS PA female (0.39)* - PANAS PA male (-0.40) - PANAS PA minority (-0.16) - PANAS PA (- PANAS PA Caucasian (0.32) - PANAS PA age ≥40 (0.78) - PANAS PA age <40 (0.21) - PANAS NA all groups (0.69)* - PANAS NA female (0.81)* - PANAS NA male (0.33) - PANAS NA minority (1.09)* - PANAS NA Caucasian (0.66)* - PANAS NA age ≥40 (0.77)* - PANAS NA age <40 (0.67)*

TABLE 3 (Continued)

Study	Comparison	BPD symptoms (ES)	Global severity (ES)	Depression and anxiety (ES)	Others (ES)
Black et al. (2018b)	1. Pre- vs. post-STEPPS + TAU 2. Pre- vs. post-TAU 3. Pre- vs. post-STEPPS in correctional sample 4. Pre- vs. post-STEPPS + TAU and the correctional sample 5. STEPPS + TAU vs. TAU	- BEST. Median <i>d</i> comparison 1: 0.48* - BEST. Median <i>d</i> comparison 2: 0.28* - BEST. Median <i>d</i> comparison 3: 0.86* - BEST. Median <i>d</i> comparison 4: 0.66* - BEST. Median <i>d</i> comparison 5: 0.21 - ZAN-BPD. Median <i>d</i> comparison 1: 0.73* - ZAN-BPD. Median <i>d</i> comparison 2: 0.22 - ZAN-BPD. Median <i>d</i> comparison 3: N/A - ZAN-BPD. Median <i>d</i> comparison 4: 0.73* - ZAN-BPD. Median <i>d</i> comparison 5: 0.48	-	-	-
Blum et al. (2002)	Pretreatment vs. posttreatment	- BEST (-1.35)*	-	- BDI (-0.78)*	- PANAS PA - PANAS NA (-0.90)*
Blum et al. (2008)	STEPPS + TAU vs. TAU (baseline vs. Week 20)	- ZAN-BPD (0.84)* - BEST (0.47) - BIS (0.54)* - Number of hospitalizations - Number of emergency department visits - Number of crisis calls - Suicide attempts and self-harm	- CGI Improv. (1.09)* - CGI Sev. (0.75)* - CGI self (0.9)* - GAS (0.65)* - SCL-90 (0.44)* - SAS (0.43)	- BDI (0.5)*	- PANAS PA (0.15) - PANAS NA (0.43)*
Boccalon et al. (2017)	Pretreatment vs. posttreatment	- DERS* - Number of hospitalizations* - Number of suicide attempts*	-	-	-
Bos et al. (2010)	STEPPS vs. TAU	- BPD-40 (T2: 0.68; T3: 0.53)* - BPSI-IV subscale impulsivity - BPSI-IV subscale Parasuicide	- SCL-90 (T2: 0.68; T3: 0.56)* - WHOQOL-BREF; overall quality of life and general health (T3: 0.61)*	-	-
Bos et al. (2011)	STEPPS vs. TAU	- BPD-40 (T2: 0.57; T3: 0.42)*	- SCL-90 (T2: 0.68; T3: 0.51)* - WHOQOL-BREF (T2: 0.63; T3: 0.69)*	-	-

(Continues)

TABLE 3 (Continued)

Study	Comparison	BPD symptoms (ES)	Global severity (ES)	Depression and anxiety (ES)	Others (ES)
González-González et al. (2021)	1. Pretreatment vs. posttreatment 2. Pre- vs. post-control 3. STEPPS vs. control group	Pre- vs. post-STEPPS: - BEST* Pre- vs. post-control: - BEST STEPPS vs. control: - BEST* - Emergency visits - Hospitalizations - Psychological or psychiatric interventions during after post-test due to borderline symptoms	-	-	-
Guillén et al. (2021)	Pretreatment vs. posttreatment	-	- QLI* (0.35) - RS-15* (0.39)	- BDI-II* (0.41)	-
Guillén Botella et al. (2021)	STEPPS vs. DBT	- BSL-23 ($\eta^2 = 0.30$)* - BIS ($\eta^2 = 0.04$) - DES-II ($\eta^2 = 0.01$) - SRS ($\eta^2 = 0.01$) - STAXI-II ($\eta^2 = 0.03$) - DERS ($\eta^2 = 0.22$)	- QoL ($\eta^2 = 0.05$) - RFL ($\eta^2 = 0.52$) - RS-15 ($\eta^2 = 0.01$)	- BDI-II ($\eta^2 = 0.01$) - OASIS ($\eta^2 = 0.08$)	-
Harvey et al. (2010)	Pretreatment vs. posttreatment	- ZAN-BPD* - BEST*	- CORE-OM*	- BDI-II*	- PANAS PA* - PANAS NA
Hezlyova et al. (2021)	Pretreatment vs. posttreatment	- QuEST (0.95)*	-	- GAD-7 (0.61)* - PHQ-9 (0.55)*	-
Hill et al. (2016)	Pretreatment vs. posttreatment	- ZAN-BPD (0.63)* - FQ (average ES: 0.56)*	- QoLS (0.73)*	-	-
Llorens Ruiz et al. (2020)	Pretreatment vs. posttreatment	- DIB-R ($\eta^2 = 0.37$)* - BEST - BIS	- CGI Sev. - CGI Impr.	-	- PAI-A

Abbreviations: ASPD, antisocial personality disorder; BDI(-II), Beck Depression Inventory (II); BEST, Borderline Evaluation of Severity Over Time; BIS, Barratt Impulsiveness Scale; BPD, borderline personality disorder; BPD-40, Borderline Personality Disorder checklist-40; BPDST-IV, Borderline Personality Disorder Severity Index-IV; BSL-23, Borderline Symptom List 23; CGI Impr., Clinical Global Impression Improvement scale; CGI Sev., Clinical Global Impression Severity scale; CGI Self, Clinical Global Impression Patient self-rating; CORE-OM, Clinical Outcomes in Routine Evaluation-OM; DBT, dialectical behavior therapy; DERS, Difficulties in Emotional Regulation Scale; DES-II, Dissociative Experiences Scale-II; DIB-R, Diagnostic Interview for Borderline Disorder-Revised; EIC, Emotional Intensity Continuum; ES, effect size; FQ, Filter Questionnaire; GAD-7, General Anxiety Disorder-7; GAS, Global Assessment Scale; OASIS, Overall Anxiety Severity and Impairment Scale; PAI-A, Personality Inventory for Adolescents; PANAS PA, Positive and Negative Affect Schedule—Positive Affectivity; PANAS NA, Positive and Negative Affect Schedule—Negative Affectivity; PGSR, Patient Global Self-Rating; PHQ-9, Patient Health Questionnaire; QLI, Quality of Life Index; QoL, quality of life; QoLS, Quality of Life Scale; QuEST, Quick Evaluation of Severity Over Time; RFL, Reasons for Living; RS-15, Resilience Scale-15; SCL-90, Symptom Checklist-90; SAS, Social Adjustment Scale; SRS, Suicide Risk Scale; STAXI-II, State-Trait Anger Expression Inventory—II; STEPPS, Systems Training for Emotional Predictability and Problem Solving; T2, end of treatment; T3, 6-month follow-up; TAU, treatment as usual; WHOQOL-BREF, World Health Organization Quality of Life Assessment-Bref; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

*Negative effect sizes suggesting greater level of improvement for BPD + ASPD subjects, except for PANAS positive affectivity.
* $P \leq 0.05$.

Borderline symptoms

As can be seen in Table 3, the three RCTs (Blum et al., 2008; Bos et al., 2010, 2011) demonstrated STEPPS to be more effective than TAU in reducing BPD symptoms, as measured with Borderline Personality Disorder checklist-40 (BPD-40) and ZAN-BPD. Bos et al. (2010) describe a decrease in borderline severity, both in participants with threshold BPD, as in participants with sub-threshold BPD (Bos et al., 2011). This decrease was statistically significant at the end of treatment and at the 6-month follow-up, compared with the pretreatment score. The clinical relevance is demonstrated with small to medium effect sizes. This finding indicates STEPPS was equally effective in participants with and without threshold BPD. Furthermore, Blum et al. (2008) found large effect sizes for borderline severity, as assessed with ZAN-BPD, with the decrease maintained at 1-year follow-up. However, participants receiving STEPPS compared with participants receiving TAU did not improve significantly better on borderline symptom severity as assessed with the BEST. Most original pretreatment versus posttreatment studies investigating BPD symptoms with the use of BEST and the Quick Evaluation of Severity Over Time (i.e., a slightly modified version of the BEST) resulted in significant decreases (except for Llorens Ruiz et al., 2020), with effect sizes ranging from Cohen's $d = 0.73$ to $d = 1.35$. Moreover, measures of borderline severity with ZAN-BPD resulted in significant differences in premeasurements versus postmeasurements (Harvey et al., 2010; Hill et al., 2016), with Hill et al. (2016), reporting a medium effect size of $d = 0.63$.

Considering specific DSM-5 symptoms of BPD, mixed results were found. While one RCT (Bos et al., 2010) and a study in adolescents (Llorens Ruiz et al., 2020) found no significant decrease in impulsivity, another RCT (Blum et al., 2008) did find a significant and clinically relevant ($d = 0.54$) decrease in impulsivity with gains maintained at 1-year follow-up. Similarly, mixed results were found in recurrent suicidal behaviors. These behaviors were predominantly measured with observable variables (e.g., number of hospitalizations and number of suicide attempts) and with the use of a validated questionnaire (i.e., Borderline Personality Disorder Severity Index [BPDSI-IV]). In general, two RCTs provided no effect of STEPPS on this criterion. The Parasuicide subscale of the BPDSI-IV yielded no significant improvement (Bos et al., 2010). Likewise, there were no significant differences between STEPPS and TAU participants in the number of hospitalizations, number of crisis calls, and (time to first) suicide attempts and self-harm acts (Blum et al., 2008). Participants attending STEPPS, however, reported less emergency department visits during

treatment and follow-up. Again, most original uncontrolled studies (except for González-González et al., 2021) did find significant decreases in pretreatment versus post-treatment number of hospitalizations, suicide attempts, and suicidal behaviors. Finally, two studies reported significant decreases on affective instability, as measured with the Difficulties in Emotion Regulation Scale (DERS) by Boccalon et al. (2017) and Emotional Intensity Continuum by Alesiani et al. (2014).

Global severity

Throughout the studies, global severity was assessed with questionnaires measuring quality of life, psychological resilience, severity of illness, global improvement, and general psychological problems. Measures were both self-report and rater administered. Except for the study conducted in adolescents (Llorens Ruiz et al., 2020), all studies, including follow-up data in the RCTs, yielded significant improvements for these questionnaires, with small to large effect sizes. In the study of Guillén et al. (2021), participants received either STEPPS, DBT, or TAU. All participants demonstrated improved quality of life, with a small effect size and no significant differences between the groups. Although statistically significant, participants showed no clinically significant improvements in their quality of life (i.e., treatment did not cause a reliable change, and quality of life did not raise to normal levels found in nonclinical populations).

Depression and anxiety

Because major depression is common in patients with BPD (Zanarini et al., 1998), several studies also investigated changes in depressive symptoms with the use of the Patient Health Questionnaire and BDI(-II). STEPPS seemed to be correlated with a decrease of depressive symptoms; seven studies yielded less symptomatology, with small to large effect sizes. In their RCT, Blum et al. (2008) found a significant difference between participants attending STEPPS and participants receiving TAU, with a medium effect size of $d = 0.5$. Finally, symptoms of anxiety also declined after treatment, yielding a medium effect size.

Affectivity

Six studies investigated changes in positive and negative affectivity (i.e., the tendency to experience negative and positive emotions) for patients with BPD, after attending the STEPPS program. Results were fairly consistent. All

studies reported significant decreases in negative affectivity with small to large effect sizes. One study (Harvey et al., 2010) reported increased positive affectivity. This finding, however, could not be replicated in other studies.

ASPD

One pilot study (Black et al., 2008), one uncontrolled pretreatment versus posttreatment study (Black et al., 2013), and three secondary data analyses (Black et al., 2018a, 2018b, 2016) were performed on offenders with BPD. Black et al. (2016) were especially interested in whether participants with BPD and comorbid ASPD could benefit from STEPPS. Higher level of improvement was found for comorbid ASPD participants in impulsiveness, BPD symptoms, illness severity, and positive affectivity, than for participants with only BPD. In none of the investigated variables, participants with BPD alone had better outcome than participants with comorbid ASPD.

DBT

Guillén Botella et al. (2021) explored the effectiveness of DBT and STEPPS for participants with BPD in a naturalistic study without randomization of participants across treatment arms. In this study, the DBT and STEPPS protocols were adjusted to balance treatment dosage for both groups. In short, the DBT program was reduced to 24 group sessions plus weekly individual sessions, while the STEPPS program was supplemented with weekly meetings with a clinician. Pretreatment analyses revealed participants attending DBT had higher illness severity and a higher number of psychiatric comorbidities beforehand. Results indicate both groups had statistically significant improvements in BPD symptoms, emotion regulation, dissociation, anxiety, fear of suicide, suicidal risk, and anger as a trait, with medium to large effect sizes. There were no changes in depression, impulsiveness, quality of life, psychological resilience, and anger as state. Moreover, participants attending DBT demonstrated significantly more improvement in BPD symptoms (e.g., suicide attempts and self-harm acts; $\eta^2 = 0.30$) and higher fear of suicide ($\eta^2 = 0.30$), than participants attending STEPPS.

DISCUSSION

The current review summarized the results of 20 empirical studies on STEPPS, a group therapy program

developed for the treatment of (core symptoms of) BPD. The aim was to (a) identify (types of) studies on STEPPS, (b) assess the quality of these studies, (c) describe patient groups receiving STEPPS, (d) examine differences and similarities in the international STEPPS protocols and manuals, and (e) summarize outcome variables investigated and their effect sizes. We found empirical evidence that STEPPS is an effective treatment. However, this evidence arises from a modest amount of high quality studies. There is a diversity of study designs, with a clear methodological weakness concerning study completion. STEPPS has mainly been studied in adult BPD patients and in forensic offenders. Data on different age groups, especially in older adults, are missing. Below, we summarize the findings of this systematic review's aims.

Types of empirical studies on STEPPS

Three RCTs were performed to investigate the effectiveness of STEPPS. A recent meta-analysis by Storebø et al. (2020) yielded 75 psychological treatments RCTs for BPD. Most trials ($N = 24$) were performed on DBT. STEPPS, however, had a comparable amount of RCTs as schema therapy ($N = 4$), psychodynamic psychotherapy ($N = 3$), and client-centered therapy ($N = 2$).

Most conducted studies on STEPPS were uncontrolled trials ($N = 11$). A limitation of these type of studies is that the observed changes in outcome cannot be attributed to the intervention applied. Nevertheless, pretreatment versus posttreatment research designs serve several important purposes, such as identifying the most suitable patients for the treatment, and determining whether there is a clinical effect that warrants further research (White & Ernst, 2001).

Finally, all studies were conducted in Western countries, with nine studies originating from the USA. Five of these nine studies were secondary data analyses, thus containing no new study participants. Moreover, all studies from the USA were conducted in Iowa, a rural state in the Midwest, with high percentages of white Americans. This raises questions about the generalizability of study results to patients with other ethnicities, patients in more urban states of the USA and patients in non-Western countries.

Quality

To assess study quality, 10 items were formulated based on two existing quality assessment tools. The overall quality of the studies was good. The RCTs demonstrated good to excellent quality scores. A general

methodological strength was that most used outcome measurements were proven to be valid and reliable. Furthermore, in a large number of studies (14 out of 20), proper diagnostic instruments were used to diagnose PDs. About half of the unique study participants were properly diagnosed with BPD with the use of a reliable structured interview designed to diagnose the presence of a PD. Approximately 40% of the participants, however, were not assessed by a suitable diagnostic interview.

A methodological weakness was related to study completion and dropout. Quality analyses on participants allocated to STEPPS demonstrated a mean dropout rate of 42.0%. For example, a recent meta-analysis on dropout rates in DBT yielded an average dropout of 28% (Dixon & Linardon, 2019). Another review on treatment completion demonstrated an overall completion rate of 75% for participants with BPD, following interventions of <12-month duration (Barnicot et al., 2010). So participants attending STEPPS had high attrition rates. Moreover, multiple studies described that participants finishing the treatment did not attend all group sessions. For example, a mean of 12.9 group sessions were attended by STEPPS participants in the study of Blum et al. (2008). Because in most studies dropout and absence reasons are missing, a clear rationale for this high attrition rate is lacking. However, there are differences between STEPPS and other psychotherapies for BPD, causing a potential risk for dropout. STEPPS was intended to be highly accessible for patients. Therefore, STEPPS is supplemented on an ongoing therapy and, in its original protocol, contains no individual sessions. Furthermore, in contrast to DBT, there is no assessment and pretreatment for motivation, therapy commitment, and therapist–patient collaboration. In DBT, patients who are not sufficiently motivated or committed to the program are not admitted to the treatment. In STEPPS, there is no stringent indication procedure. As a negative consequence of the accessibility of the treatment, it is conceivable that suboptimal motivated patients are admitted to the program, causing high attrition rates. This raises the question whether there should be more attention to therapy motivation, commitment, and therapist–patient collaboration in STEPPS. Further research should address this issue.

Patients

Regarding patient characteristics, some notable conclusions can be drawn. First, most study participants were female (range: 78% to 100%). This was also true for the prison sample, despite data indicating higher percentages of male prisoners (i.e., 93.1%; United States Federal Bureau of Prisons, 2021), as well as higher percentage of

male prisoners with a PD (i.e., 22.4% of males and 14.5% of females had a PD diagnosis; Rotter et al., 2002). Results have shown gender, age, and ethnicity not to be associated with greater improvement after treatment.

Second, mean age of study participants was 33 years. No studies were conducted in older adults. Emotional dysregulation, however, does not seem to decline with age (Frias et al., 2017), and PDs are highly prevalent in old age (Penders et al., 2020). So older adults may benefit equally from STEPPS. This is line with previous findings, treating older adults with comorbid PDs and depression with DBT (Lynch et al., 2007) and treating PDs in older adults with schema therapy (Videler et al., 2018). Moreover, only one study was conducted in a sample of adolescents. This is surprising, because problems with self-regulation, relational aggression, affective, and impulsive symptoms are the early signs of BPD (Hutsebaut et al., 2019). These signs affect school and social functioning. Early intervention programs for these symptoms might positively affect the course of illness and are likely to be cost effective. Hutsebaut et al. (2019) advocate early intervention programs to include school-based prevention, psychoeducation, and a program for parents. Given its classroom format, emphasis on psychoeducation and the involvement of the reinforcement team, the STEPPS program seems to fit well in these requirements.

Finally, in accordance with the target group for STEPPS, most patients referred to STEPPS were thought to have BPD. Both participants with threshold BPD and subthreshold BPD were included in various studies, with similar effects. In addition, in several studies, (forensic) participants with comorbid ASPD have also been treated with STEPPS, with similar efficacy.

Protocols

The original STEPPS protocol consists of 20 weekly 2-h group sessions and at least one session for the reinforcement team, added to patient's TAU. Approximately half of the included studies have investigated the effects of the original STEPPS. In addition, several modified protocols have also been studied. One study investigated the effects of STEPPS Early Intervention, a shortened and less-intensive variant of STEPPS. In contrast to STEPPS, the target population of STEPPS Early Intervention is patients with emotional intensity difficulties, without a threshold BPD per se. The most common modification to the protocol was the addition of individual sessions to STEPPS. These individual sessions followed the STEPPS protocol. In our view, this modification alters STEPPS from an add-on treatment to a stand-alone psychotherapy.

Outcome

Four categories of outcome could be identified in the results of the included studies: BPD symptoms, global severity, depression and anxiety, and affectivity. It can be concluded that treatment with STEPPS seems to affect all these variables positively to some extent. Participants attending STEPPS had significantly less borderline severity, improved quality of life, less depressive and anxiety symptoms, and reduced tendency to experience negative emotions, after treatment. The original main objectives of STEPPS were to minimize hospitalization and self-harm acts. Surprisingly, the RCTs did not find differences in hospitalizations, suicide attempts, and self-harm acts between STEPPS and TAU. Blum et al. (2008) suggest that these behavioral changes may occur after a more active follow-up and longer treatment period than the 20-week STEPPS program. In addition, authors suggest that the low base rate of these behaviors in their sample, could have led to a false negative error. Uncontrolled studies, on the other hand, did find significant decreases in hospitalizations, suicide attempts, and self-harm acts after STEPPS. Mixed results were found for impulsivity and recurrent suicidal behaviors. Furthermore, despite STEPPS's focus on acquiring emotion regulation skills, only two studies included emotion regulation measurements (i.e., DERS). Although effect sizes were lacking, both studies demonstrated significant results in pretreatment versus posttreatment scores.

Both participants with threshold and subthreshold BPD were included in studies. In general, mentioned results did not differ for these patient groups. This suggests STEPPS is an equally effective treatment program for participants with threshold and subthreshold BPD. Interestingly, participants with BPD and comorbid ASPD seemed to benefit significantly more from STEPPS than participants with BPD alone. A clear rationale for this finding is lacking. Authors describe that participants with ASPD had higher symptom levels and therefore probably had more to gain from the program (Black et al., 2016). We hypothesize participants referred to STEPPS all experience problems with emotion regulation (a core feature of cluster B PDs; Neacsiu & Tkachuck, 2016), with most participants experiencing intense emotions and deficient skills to cope with these emotions. Some participants meet the DSM-5 criteria of BPD, while others do not. The underlying difficulties in emotion regulation might be the key feature on which STEPPS is effective, as emotional dysregulation can be seen as a transdiagnostic treatment construct, central to the development and maintenance of severe (PD) psychopathology (Sloan et al., 2017). This could implicate that other patient groups (e.g., patients with ASPD, bipolar disorder, or

autism spectrum disorder) experiencing emotion regulation deficiencies may benefit from STEPPS. Riemann et al. (2014) present a RCT protocol for patients with bipolar disorder and comorbid BPD traits. Results on these and other patient groups, however, are not available.

As also reported by Guillén Botella et al. (2021), in our clinical experience, patients with more severe BPD symptoms are often referred to the DBT program, while STEPPS is recommended for patients with less BPD symptomatology. In this study, participants attending DBT demonstrated significantly better outcome than participants attending STEPPS. However, participants were not randomized between DBT and STEPPS. Furthermore, as in the study of Black et al. (2016), through regression to the mean, higher illness severity of the DBT group might have influenced group differences found. Indeed, analyses of baseline characteristics reveal higher baseline symptom severity to be associated with better outcome in participants attending STEPPS. In addition, better patient collaboration also predicts better outcome. These findings suggest patients experiencing emotion regulation problems, regardless of illness severity, could be admitted to STEPPS. This fits the accessible nature of STEPPS. Strengthening patient–therapist collaboration, for example, through the addition of individual sessions, might result in better outcome. As noted earlier, this might also positively affect the attrition rate. When STEPPS is offered as an add-on treatment without adjunctive individual sessions, other allocation criteria should be used. Considering the results of the present review, we advocate to select patients, who are able to benefit from a classroom approach, and with intrinsic motivation, to be assigned to the STEPPS program. Possibly, patients who exhibit more therapy interfering behaviors and problems concerning therapy commitment should be considered to be assigned to DBT.

CONCLUSION

This is the first systematic review covering the current evidence for the effectiveness of STEPPS. In sum, there is empirical evidence that STEPPS is a promising treatment for BPD symptoms, global severity, depressive symptoms, and negative affectivity. The core symptoms treated seem to be difficulties in emotion regulation, also present in patients with threshold BPD and patients with BPD traits. Moreover, STEPPS seems to be a suitable treatment for patients with BPD and comorbid ASPD. Whether it is suitable and effective for other patient groups and different age categories, like adolescents and older adults, needs to be investigated. A major limitation of STEPPS is

the high attrition rate, which seems to be a negative consequence of its accessibility. Some authors have investigated an adjusted STEPPS protocol, in which patients also received individual sessions following the STEPPS protocol. More emphasis on commitment and therapist-patient collaboration, for example, through structured individual sessions, might decrease dropout rates. However, this might also affect the accessible character of the treatment. Furthermore, originally STEPPS is an add-on treatment. Supplementing STEPPS with structured individual sessions alters the program to a stand-alone psychotherapy. Future research should address these issues.

CONFLICT OF INTEREST

All authors declare that they have no conflicts of interest.

ETHICS STATEMENT

There was no ethical approval required as this concerns a systematic review.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author (EE) upon reasonable request.

AUTHOR CONTRIBUTIONS

EE, SPJvA, and ACV were responsible for the conception of the review, designed the study, wrote the protocol, and designed the search strategy. EE and JVdP conducted database searches and initial screening of articles and completed the full-text review of articles and quality assessments. SPJvA and ACV supervised the study selection and quality assessment. EE synthesized results of included studies and wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

CLINICAL TRIAL REGISTRATION

The protocol for this systematic review was preregistered with PROSPERO (ID = CRD42021253254).

ORCID

Erol Ekiz  <https://orcid.org/0000-0001-7143-5157>

REFERENCES

- Alesiani, R., Boccalon, S., Giarolli, L., Blum, N., & Fossati, A. (2014). Systems Training for Emotional Predictability and Problem Solving (STEPPS): Program efficacy and personality features as predictors of drop-out—An Italian study. *Comprehensive Psychiatry*, *55*(4), 920–927. <https://doi.org/10.1016/j.comppsy.2014.01.003>
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- Bailey, R. C., & Grenyer, B. F. S. (2013). Burden and support needs of carers of persons with borderline personality disorder. *Harvard Review of Psychiatry*, *21*(5), 248–258. <https://doi.org/10.1097/hrp.0b013e3182a75c2c>
- Barnicot, K., Katsakou, C., Bhatti, N., Savill, M., Fearn, N., & Priebe, S. (2012). Factors predicting the outcome of psychotherapy for borderline personality disorder: A systematic review. *Clinical Psychology Review*, *32*(5), 400–412. <https://doi.org/10.1016/j.cpr.2012.04.004>
- Barnicot, K., Katsakou, C., Marougka, S., & Priebe, S. (2010). Treatment completion in psychotherapy for borderline personality disorder—A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *123*(5), 327–338. <https://doi.org/10.1111/j.1600-0447.2010.01652.x>
- Bartels, N., & Crotty, T. (1998). *A systems approach to treatment: The borderline personality disorder skill training manual*. EID Treatment Systems.
- Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *American Journal of Psychiatry*, *156*(10), 1563–1569. <https://doi.org/10.1176/ajp.156.10.1563>
- Beck, A. T. (1978). *Depression inventory*. Philadelphia Center for Cognitive Therapy.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II* (2nd ed.). Psychological Corporation.
- Black, D. W., Allen, J., St. John, D., Pfohl, B., McCormick, B., & Blum, N. (2009). Predictors of response to Systems Training for Emotional Predictability and Problem Solving (STEPPS) for borderline personality disorder: An exploratory study. *Acta Psychiatrica Scandinavica*, *120*(1), 53–61. <https://doi.org/10.1111/j.1600-0447.2008.01340.x>
- Black, D. W., Allen, J., McCormick, B., & Blum, N. (2011). Treatment received by persons with BPD participating in a randomized clinical trial of the Systems Training for Emotional Predictability and Problem Solving programme. *Personality and Mental Health*, *5*(3), 159–168. <https://doi.org/10.1002/pmh.167>
- Black, D. W., Blum, N., & Allen, J. (2018a). Does response to the STEPPS program differ by sex, age, or race in offenders with borderline personality disorder? *Comprehensive Psychiatry*, *87*, 134–137. <https://doi.org/10.1016/j.comppsy.2018.10.005>
- Black, D. W., Blum, N., & Allen, J. (2018b). STEPPS treatment programme for borderline personality disorder: Which scale items improve? An item-level analysis. *Personality and Mental Health*, *12*(4), 345–354. <https://doi.org/10.1002/pmh.1431>
- Black, D. W., Blum, N., Eichinger, L., McCormick, B., Allen, J., & Sieleni, B. (2008). STEPPS: Systems Training for Emotional Predictability and Problem Solving in women offenders with borderline personality disorder in prison—A pilot study. *CNS Spectrums*, *13*(10), 881–886. <https://doi.org/10.1017/s1092852900016989>
- Black, D. W., Blum, N., McCormick, B., & Allen, J. (2013). Systems Training for Emotional Predictability and Problem Solving (STEPPS) group treatment for offenders with borderline personality disorder. *Journal of Nervous & Mental Disease*, *201*(2), 124–129. <https://doi.org/10.1097/nmd.0b013e31827f6435>

- Black, D. W., Simsek-Duran, F., Blum, N., McCormick, B., & Allen, J. (2016). Do people with borderline personality disorder complicated by antisocial personality disorder benefit from the STEPPS treatment program? *Personality and Mental Health, 10*(3), 205–215. <https://doi.org/10.1002/pmh.1326>
- Blum, N., St. John, D., Pfohl, B., Stuart, S., McCormick, B., Allen, J., Arndt, S., & Black, D. W. (2008). Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: A randomized controlled trial and 1-year follow-up. *American Journal of Psychiatry, 165*(4), 468–478. <https://doi.org/10.1176/appi.ajp.2007.07071079>
- Blum, N., Pfohl, B., St. John, D., Monahan, P., & Black, D. W. (2002). STEPPS: A cognitive-behavioral systems-based group treatment for outpatients with borderline personality disorder—A preliminary report. *Comprehensive Psychiatry, 43*(4), 301–310. <https://doi.org/10.1053/comp.2002.33497>
- Blum, N. S., Bartels, N. E., St. John, D., & Pfohl, B. (2012). *STEPPS: Group treatment program for borderline personality disorder* (2nd ed.). Level One Publishing.
- Boccalon, S., Alesiani, R., Giarolli, L., & Fossati, A. (2017). Systems Training for Emotional Predictability and Problem Solving program and emotion dysregulation. *Journal of Nervous & Mental Disease, 205*(3), 213–216. <https://doi.org/10.1097/nmd.0000000000000640>
- Bos, E. H., Van Wel, E. B., Appelo, M. T., & Verbraak, M. J. P. M. (2010). A randomized controlled trial of a Dutch version of Systems Training for Emotional Predictability and Problem Solving for borderline personality disorder. *Journal of Nervous & Mental Disease, 198*(4), 299–304. <https://doi.org/10.1097/nmd.0b013e3181d619cf>
- Bos, E. H., Van Wel, E. B., Appelo, M. T., & Verbraak, M. J. P. M. (2011). Effectiveness of Systems Training for Emotional Predictability and Problem Solving (STEPPS) for borderline personality problems in a ‘real-world’ sample: Moderation by diagnosis or severity. *Psychotherapy and Psychosomatics, 80*(3), 173–181. <https://doi.org/10.1159/000321793>
- Cleary, M., Siegfried, N., & Walter, G. (2002). Experience, knowledge and attitudes of mental health staff regarding clients with a borderline personality disorder. *International Journal of Mental Health Nursing, 11*(3), 186–191. <https://doi.org/10.1046/j.1440-0979.2002.00246.x>
- 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.comppsy.2005.06.002>
- Crawford, J. R., & Henry, J. D. (2004). The Positive And Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*(3), 245–265. <https://doi.org/10.1348/0144665031752934>
- Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder. *JAMA Psychiatry, 74*(4), 319–328. <https://doi.org/10.1001/jamapsychiatry.2016.4287>
- Dixon, L. J., & Linardon, J. (2019). A systematic review and meta-analysis of dropout rates from dialectical behaviour therapy in randomized controlled trials. *Cognitive Behaviour Therapy, 49*(3), 181–196. <https://doi.org/10.1080/16506073.2019.1620324>
- Ellison, W. D., Rosenstein, L. K., Morgan, T. A., & Zimmerman, M. (2018). Community and clinical epidemiology of borderline personality disorder. *Psychiatric Clinics of North America, 41*(4), 561–573. <https://doi.org/10.1016/j.psc.2018.07.008>
- First, M. B., Spitzer, R. L., Gibbon, M., Williams, G. B. W., & Benjamin, L. (1997). *Structured clinical interview for DSM-IV axis II personality disorders (SCID-II)*. American Psychiatric Press.
- Friás, L., Palma, C., Solves, L., Martínez, B., & Salvador, A. (2017). Differential symptomatology and functioning in borderline personality disorder across age groups. *Psychiatry Research, 258*, 44–50. <https://doi.org/10.1016/j.psychres.2017.09.081>
- González-González, S., Marañón-González, R., Hoyuela-Zatón, F., Gómez-Carazo, N., Hernández-Abellán, A., Pérez-Poo, T., Umaran-Alfageme, O., Cordero-Andrés, P., López-Sánchez, V., Black, D. W., Blum, N. S., Artal-Simón, J., & Ayesa-Arriola, R. (2021). STEPPS for borderline personality disorder: A pragmatic trial and naturalistic comparison with noncompleters. *Journal of Personality Disorders, 35*(6), 841–856. https://doi.org/10.1521/pedi_2021_35_512
- Gross, J. J. (1999). Emotion regulation: Past, present, future. *Cognition & Emotion, 13*(5), 551–573. <https://doi.org/10.1080/026999399379186>
- Guillén Botella, V., García-Palacios, A., Bolo Miñana, S., Baños, R., Botella, C., & Marco, J. H. (2021). Exploring the effectiveness of dialectical behavior therapy versus Systems Training for Emotional Predictability and Problem Solving in a sample of patients with borderline personality disorder. *Journal of Personality Disorders, 35*, 21–38. https://doi.org/10.1521/pedi_2020_34_477
- Guillén, V., Tormo, M. E., Fonseca-Baeza, S., Botella, C., Baños, R., García-Palacios, A., & Marco, J. H. (2021). Resilience as a predictor of quality of life in participants with borderline personality disorder before and after treatment. *BMC Psychiatry, 21*(1), 305. <https://doi.org/10.1186/s12888-021-03312-0>
- Harvey, R., Black, D. W., & Blum, N. (2010). Systems Training for Emotional Predictability and Problem Solving (STEPPS) in the United Kingdom: A preliminary report. *Journal of Contemporary Psychotherapy, 40*(4), 225–232. <https://doi.org/10.1007/s10879-010-9150-4>
- Hezelyova, I., Cribben, H., Melunsky, N., Moalypour, S., Goodwin, H., Serrat, C. M., Rahmanian, H., & Duncko, R. (2021). Investigating effectiveness and predicting outcome and dropout from Systems Training for Emotional Predictability and Problem Solving for emotional intensity difficulties (STEPPS EI) in UK primary care. *Journal of Personality Disorders, 35*(6), 902–916. https://doi.org/10.1521/pedi_2021_35_519
- Hill, N., Geoghegan, M., & Shawe-Taylor, M. (2016). Evaluating the outcomes of the STEPPS programme in a UK community-based population; implications for the multidisciplinary treatment of borderline personality disorder. *Journal of Psychiatric and Mental Health Nursing, 23*(6–7), 347–356. <https://doi.org/10.1111/jpm.12315>
- Hutsebaut, J., Videler, A. C., Verheul, R., & Van Alphen, S. P. J. (2019). Managing borderline personality disorder from a life course perspective: Clinical staging and health management. *Personality Disorders: Theory, Research, and Treatment, 10*(4), 309–316. <https://doi.org/10.1037/per0000341>

- Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry*, 48, 1060–1064. <https://doi.org/10.1001/archpsyc.1991.01810360024003>
- Llorens Ruiz, I., Lopez Seco, F., Masana Marín, A., San Martín Águila, E., Martínez Pinar, D., Martín Alonso, S., Vilella, E., & Gutierrez-Zotes, A. (2020). Emotional regulation for adolescents: A group-based treatment pilot study through the STEPPS programme. *Behavioural and Cognitive Psychotherapy*, 48(6), 734–738. <https://doi.org/10.1017/s1352465820000454>
- Lynch, T. R., Cheavens, J. S., Cukrowicz, 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(2), 131–143. <https://doi.org/10.1002/gps.1703>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- National Collaborating Centre for Mental Health. (2009). *Borderline personality disorder: Treatment and management NICE clinical guideline*. National Institute for Health and Clinical Excellence.
- Neacsiu, A. D., & Tkachuck, M. A. (2016). Dialectical behavior therapy skills use and emotion dysregulation in personality disorders and psychopathy: A community self-report study. *Borderline Personality Disorder and Emotion Dysregulation*, 3(1), 6. <https://doi.org/10.1186/s40479-016-0041-5>
- Penders, K. A., Peeters, I. G., Metsemakers, J. F., & Van Alphen, S. P. (2020). Personality disorders in older adults: A review of epidemiology, assessment, and treatment. *Current Psychiatry Reports*, 22(3), 1–14. <https://doi.org/10.1007/s11920-020-1133-x>
- Pfohl, B., Blum, N., St. John, D., McCormick, B., Allen, J., & Black, D. W. (2009). Reliability and validity of the Borderline Evaluation of Severity Over Time (BEST): A self-rated scale to measure severity and change in persons with borderline personality disorder. *Journal of Personality Disorders*, 23(3), 281–293. <https://doi.org/10.1521/pedi.2009.23.3.281>
- Pfohl, B., Blum, N., & Zimmerman, M. (1997). *Structured interview for DSM-IV personality disorders (SIDP-IV)*. American Psychiatric Press.
- Riemann, G., Weisscher, N., Goossens, P. J., Draijer, N., Apenhorst-Hol, M., & Kupka, R. W. (2014). The addition of STEPPS in the treatment of patients with bipolar disorder and comorbid borderline personality features: A protocol for a randomized controlled trial. *BMC Psychiatry*, 14(1), 172. <https://doi.org/10.1186/1471-244x-14-172>
- Ring, D., & Lawn, S. (2019). Stigma perpetuation at the interface of mental health care: A review to compare patient and clinician perspectives of stigma and borderline personality disorder. *Journal of Mental Health*, 1–21. <https://doi.org/10.1080/09638237.2019.1581337>
- Rotter, M., Way, B., Steinbacher, M., Sawyer, D., & Smith, H. (2002). Personality disorders in prison: Aren't they all antisocial? *Psychiatric Quarterly*, 73(4), 337–349. <https://doi.org/10.1023/a:1020468117930>
- Skodol, A. E., Gunderson, J. G., Shea, M. T., McGlashan, T. H., Morey, L. C., Sanislow, C. A., Bender, D. S., Grilo, C. M., Zanarini, M. C., Yen, S., Pagano, M. E., & Stout, R. L. (2005). The collaborative longitudinal personality disorders study (CLPS): Overview and implications. *Journal of Personality Disorders*, 19(5), 487–504. <https://doi.org/10.1521/pedi.2005.19.5.487>
- Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017). Emotion regulation as a transdiagnostic treatment construct across anxiety, depression, substance, eating and borderline personality disorders: A systematic review. *Clinical Psychology Review*, 57, 141–163. <https://doi.org/10.1016/j.cpr.2017.09.002>
- Steer, R. A., Brown, G. K., Beck, A. T., & Sanderson, W. C. (2001). Mean Beck Depression Inventory–II scores by severity of major depressive episode. *Psychological Reports*, 88(3_suppl), 1075–1076. <https://doi.org/10.2466/pr0.2001.88.3c.1075>
- Stoffers-Winterling, J., Storebø, O. J., & Lieb, K. (2020). Pharmacotherapy for borderline personality disorder: An update of published, unpublished and ongoing studies. *Current Psychiatry Reports*, 22(8), 37. <https://doi.org/10.1007/s11920-020-01164-1>
- Storebø, O. J., Stoffers-Winterling, J. M., Völlm, B. A., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., Faltinsen, E., Todorovac, A., Sales, C. P., Callesen, H. E., Lieb, K., & Simonsen, E. (2020). Psychological therapies for people with borderline personality disorder. *Cochrane Database of Systematic Reviews*, 2020(11), CD012955. <https://doi.org/10.1002/14651858.cd012955.pub2>
- Thomas, B., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, 1(3), 176–184. <https://doi.org/10.1111/j.1524-475x.2004.04006.x>
- United States Federal Bureau of Prisons. (2021). Inmate gender statistics. https://www.bop.gov/about/statistics/statistics_inmate_gender.jsp
- Videler, A. C., van Alphen, S. P. J., van Royen, R. J., van der Feltz-Cornelis, C. M., Rossi, G., & Arntz, A. (2018). Schema therapy for personality disorders in older adults: A multiple-baseline study. *Aging & Mental Health*, 22, 738–747. <https://doi.org/10.1080/13607863.2017.1318260>
- Watson, D., & Clark, L. A. (1994). *The PANAS-Manual for the Positive and Negative Affect Schedule—expanded form*. The University of Iowa.
- White, A., & Ernst, E. (2001). The case for uncontrolled clinical trials: A starting point for the evidence base for CAM. *Complementary Therapies in Medicine*, 9(2), 111–116. <https://doi.org/10.1054/ctim.2001.0441>
- Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. Guilford Press.
- Zanarini, M. C. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): A continuous measure of DSM-IV borderline psychopathology. *Journal of Personality Disorders*, 17(3), 233–242. <https://doi.org/10.1521/pedi.17.3.233.22147>

- Zanarini, M. C., Frankenburg, F. R., Dubo, E. D., Sickel, A. E., Trikha, A., Levin, A., & Reynolds, V. (1998). Axis I comorbidity of borderline personality disorder. *American Journal of Psychiatry*, 155(12), 1733–1739. <https://doi.org/10.1176/ajp.155.12.1733>
- Zanarini, M. C., Gunderson, J. G., Frankenburg, F. R., & Chauncey, D. L. (1989). The revised diagnostic interview for borderlines: Discriminating BPD from other axis II disorders. *Journal of Personality Disorders*, 3(1), 10–18. <https://doi.org/10.1521/pedi.1989.3.1.10>

How to cite this article: Ekiz, E., van Alphen, S. P. J., Ouwens, M. A., Van de Paar, J., & Videler, A. C. (2022). Systems Training for Emotional Predictability and Problem Solving for borderline personality disorder: A systematic review. *Personality and Mental Health*, 1–20. <https://doi.org/10.1002/pmh.1558>