Effectiveness of the Unified Protocol for transdiagnostic treatment of emotional disorders in group format in Spain: Results from a randomized controlled trial with 6-Months Follow-up

Osma, J.[1], [2]*, Peris-Baquero, O.[1], [2], Suso-Ribera, C.[3], Farchione, T.J.[4], & Barlow, D.H.[4]

[1]Universidad de Zaragoza (Zaragoza, Spain)

[2]Instituto de Investigación Sanitaria de Aragón (Zaragoza, Spain)

[3]Universitat Jaume I (Castellón, Spain)

[4]Center for Anxiety and Related Disorders, Boston University (EEUU)

* Jorge Osma. Universidad de Zaragoza. Departamento de Psicología y Sociología. Facultad de Ciencias Sociales y Humanas. C/Ciudad Escolar s/n. 44003. Teruel (Spain). Tel.: (+34) 978645390 (ext. 861390). E-mail: osma@unizar.es

Acknowledgements

We would like to thank all collaborating centers, all healthcare professionals, and all volunteers and participants that made this study possible.

Abstract

Background: The present study aims to investigate the effectiveness of the Unified Protocol (UP), a transdiagnostic treatment of emotional disorders (EDs), when applied in a group format in the public mental health system in Spain.

Methods: 488 participants with a primary diagnosis of ED were randomized to the UP group or to the treatment as usual (TAU; individual, disorder-specific cognitive behavioral therapy). Personality, depression and anxiety symptoms, affect, and quality of life were assessed at pre-treatment, 3 months after treatment onset (coinciding with the end of the UP treatment), and 6 and 9 months after treatment onset (follow-ups). The moderating effect of the treatment condition and the number of sessions received in the evolution of study outcomes was investigated with a linear mixed model analysis.

Results: A significant improvement in outcomes occurred in both conditions, except for extraversion in the TAU. Improvements in depression, anxiety and quality of life were larger in the UP condition. After the treatment, improvements were maintained at follow-ups in all study outcomes. An interaction between Time*Condition*Sessions was found for depression.

Conclusion: The results add to the existing evidence on the effectiveness of the UP and may be important for implementation purposes in the Spanish or other similar public mental health systems. Trial registration number NCT03064477 (March 10, 2017).

Keywords*: Emotional disorders, Unified Protocol, transdiagnostic, group format, public mental health.*

Introduction

The Unified Protocol for the transdiagnostic treatment of emotional disorders (UP) is a cognitive-behavioral intervention developed to address emotional disorders (EDs; anxiety, depression and related disorders; Barlow et al., 2011a). People with EDs experience: (a) frequent and intense negative emotions (e.g., anxiety, anger, sadness or guilt); (b) aversive reactions to these emotions, and (c) urges to avoid these emotional experiences, generally using maladaptive strategies (Bullis et al., 2019). There is also considerable evidence that these shared characteristics by individuals presenting EDs may be explained by underlying temperamental characteristics, particularly high neuroticism, thus resulting in emotion dysregulation (Barlow et al., 2014).

The efficacy of the UP for EDs has been supported by numerous randomized control trials, as evidenced in recent systematic reviews (Cassiello-Robbins et al., 2020; Sakiris & Berle, 2019). However, the majority of studies to date have been conducted in highly controlled research settings (e.g., laboratories, private centers or residential treatment centers; Boswell et al., 2012; Thompson-Brenner et al., 2019) and in an individual format. This is important to mention because individuals with EDs often seek help at public health institutions and these organizations are still in need of cost-effective and affordable interventions for common mental health problems (Saxena et al., 2007).

The UP might provide a feasible solution to the aforementioned problems associated with the limited resources for mental health treatments in public settings. For example, the UP has been adapted to be administered in a group format, with is important for cost-effectiveness purposes (Laposa et al., 2017). Additionally, the fact that the UP can be applied to individuals within the full range of EDs facilitates the creation of such groups.

To our knowledge, nine studies have now provided evidence on the effectiveness of group UP for EDs (Cassiello-Robbins et al., 2020). Of these, four were non-controlled studies (Grill et al., 2017; Laposa et al., 2017; Osma et al., 2015; Varkovitzky et al., 2018), two were randomized controlled trials (de Ornelas et al., 2017; Zemestani et al., 2017), one was a non-randomized controlled trial (de Ornelas et al., 2015), and two were open trials (Bullis et al., 2015; Reinholt et al., 2017). Finally, only four of these nine studies were conducted in public mental health settings (de Ornelas et al., 2015; Osma et al., 2015; Reinholt et al., 2017; Varkovitzky et al., 2018).

The results of the previous investigations on the effectiveness of a group UP intervention are promising and moderate-to-large effects on anxiety ($0.45 \leq d \leq 2.25$) and depression symptoms ($0.44 \leq d \leq 3.08$) have been reported. However, a number of limitations have been also identified. For instance, sample sizes have been generally small ($11 \leq n \leq 52$), which limits the generalizability of these findings and raises questions regarding the reliability of study findings, including effect sizes, due to limited power. In fact, as indicated above, effect sizes have been very diverse across studies, ranging from moderate to very large. Additionally, follow-up periods have been either very short (up to 4 months) or non-existent (in 5 out of 9 studies). Only one investigation evaluated the effectiveness of the UP in a longer term (12 month), but again the sample size was small ($n = 6$; Osma et al., 2015).

A final shortcoming refers to the selection of outcomes, which has been characterized by an excessive focus on anxiety and depression symptomatology. For example, only four in nine studies assessed quality of life or personality and affect when exploring the effectiveness of the UP in group format. This is important because the UP aims to lead to enduring changes in personality (i.e., neuroticism) and functioning (i.e., quality of life) (Barlow et al., 2014). So far, the UP applied in a group format appears to

result in moderate-to-large changes in quality of life (Bullis et al., 2015; de Ornelas et al., 2017; Osma et al., 2015; Reinholt et al., 2017), positive and negative affect (Laposa et al., 2017; Osma et al., 2015; Reinholt et al., 2017), and neuroticism (Osma et al., 2015).

As noted earlier, the existing findings in relation to group UP interventions are promising. However, higher quality investigations using large samples, long follow-up periods, and ecological settings (e.g., public mental health settings) are still needed to support the effectiveness and durability of the UP delivered in a group format. The aim of the present study is to investigate the effectiveness of the UP applied in group format in public mental health settings in Spain. Building upon prior research and the described limitations in the literature, this multicenter randomized controlled trial will include a wide range of outcomes (anxiety and depression symptoms, quality of life, positive and negative affect, neuroticism, and extraversion) and will investigate the durability of changes up to 6 months after the end of the UP intervention (9 months after treatment onset) in a large sample of individuals with a wide range of EDs.

Consistent with previous research, we hypothesize that the UP will produce significant moderate-to-large improvements on all outcomes. Also similar to past research (Cassiello-Robbins et al., 2020; Sakiris & Berle, 2019), we expect that a moderation analysis will reveal that both interventions exert comparable effects on outcomes (moderation by treatment condition is not significant) on study outcomes when comparing the UP with the active comparator, that is, non-protocolized individual cognitive-behavioral therapy (CBT) – the usual treatment in public health settings in Spain. Because patients in both conditions are likely to receive a dissimilar number of treatment sessions due to the limited resources available for frequent (i.e., weekly)

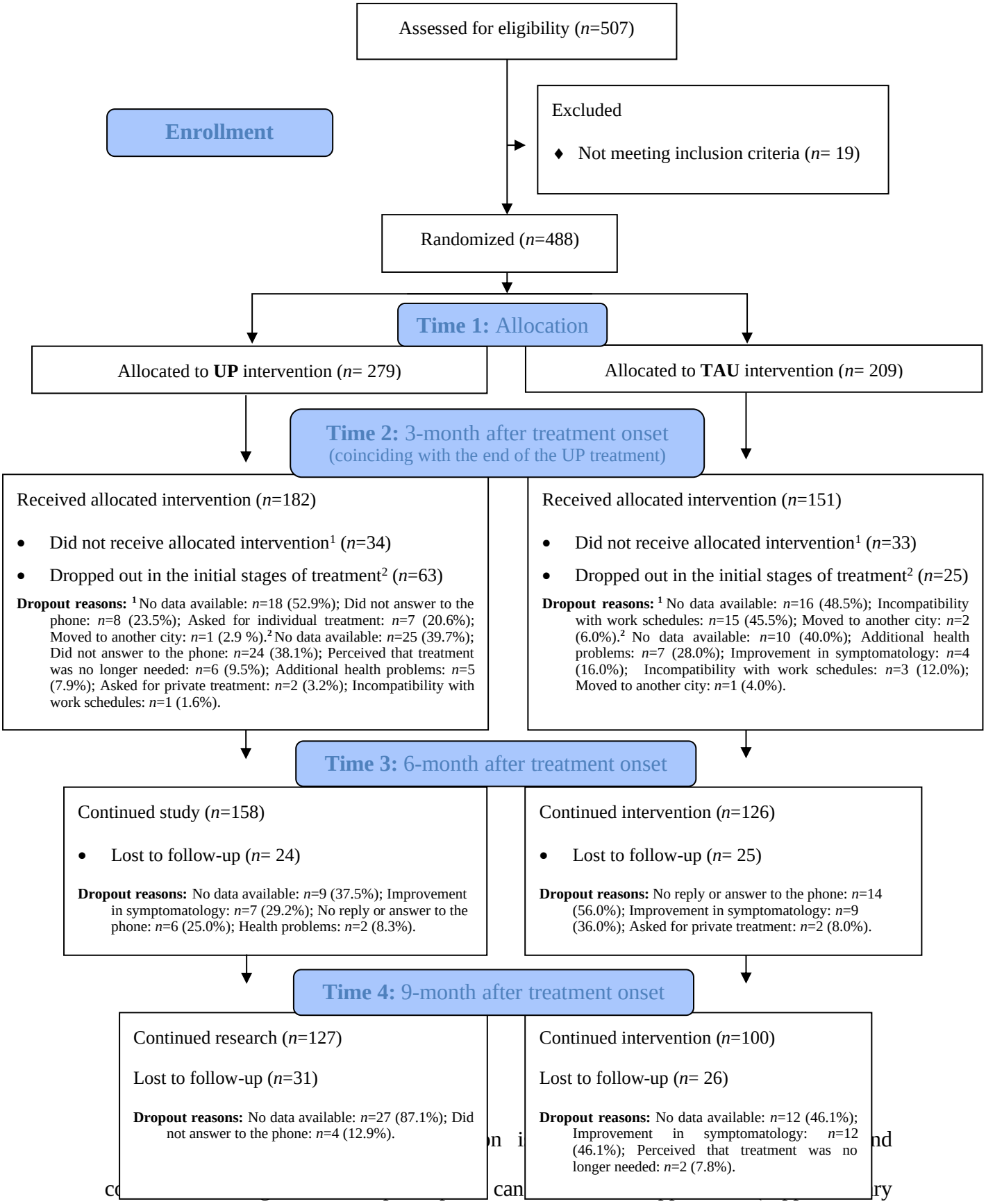
individual therapy, we will also analyze the moderating role of the number of sessions in an exploratory manner.

Method

Participants

The public mental health centers collaborating in the study are described in (Osma et al., 2018). The sample was recruited from 10 public mental health centers and was composed of 488 participants. Participants were randomly allocated to the UP (*n* = 279) or the TAU (*n* = 209). The distribution of participants according to the treatment condition and the center was: Center 1: TAU (*n* = 18), UP (*n*= 26, grouped in 2 groups); Center 2: TAU (*n* = 47), UP (*n* = 82, grouped in 6 groups); Center 3: TAU (*n* = 16), UP (*n* = 22, grouped in 2 groups); Center 4: TAU (*n* = 39), UP (*n* = 37, grouped in 4 groups); Center 5: TAU (*n* = 16), UP (*n* = 12, grouped in 2 groups); Center 6: TAU (*n* = 17), UP (*n* = 18, grouped in 2 groups); Center 7: TAU (*n* = 13), UP (*n* = 16, grouped in 2 groups); Center 8: TAU (*n* = 20), UP (*n* = 21, grouped in 2 groups); Center 9: TAU (*n* = 10), UP (*n* = 26, grouped in 2 groups); Center 10: TAU (*n* = 13), UP (*n* = 19, grouped in 2 groups). Figure 1 shows the CONSORT flow diagram. The sample size was sufficient according to previous calculations published in the protocol of the current study (Osma et al., 2018).

Figure 1. CONSORT diagram illustrating participant flow in the study



material). In the UP group, the mean age of the participants was 42.76 years (*SD* = 11.61) and 79.2% of them were women (*n* = 221). In the TAU condition, their main age was 42.45 years (*SD*=13.01) and 78.0% of participants were women (*n* = 163). In total, 220 participants in the UP (78.9%) and 154 participants in the TAU condition (73.7%) were taking psychotropic medication at the time of enrollment and randomization.

-Insert Table 1 about here-

Measures

Diagnosis. The Spanish version of the Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version (ADIS-IV-L; Di Nardo et al., 1994) translated by Botella and Ballester (1997) was used for diagnostic purposes. Test-retest reliability varies depending on the study from .68 to 1.00 (Brown et al., 2001). The ADIS-IV-L, as opposed to the ADIS-5, was used because the new clinical interview for DSM-5 was not available in Spanish at the time the initial recruitment took place.

Personality. The NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1999) is a self-report inventory which offers a rapid measure of the Big Five personality structure, namely neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. This questionnaire consists of 60 items. In the NEO-FFI, individuals are asked to respond using a 5-point Likert scale ranging from 0 (totally disagree) to 4 (totally agree). We only present data on neuroticism and extraversion because these are the underlying temperamental characteristics proposed in the transdiagnostic model of EDs (Brown & Barlow, 2009). The internal consistency in the present sample was α = .77 for neuroticism and α = .82 for extraversion.

Affect. The Positive and Negative Affect Schedule (PANAS; Sandín et al., 1999; Watson et al., 1988) consists of 20 items that evaluate positive and negative affect, with

10 items for each dimension. Responses use a 5-point Likert scale ranging from 1 (very slightly or not at all) to 5 (extremely). The internal consistency in the present sample was good for both scales, that is, α =.90 for positive affect (PA) and α = .90 for negative affect (NA).

Depression. Depression was evaluated with the Beck Depression Inventory-II and the Overall Depression Severity and Impairment Scale. The Beck Depression Inventory-II (BDI-II; Beck et al.,1996), validated into Spanish by Sanz et al., (2003), is a 21-item self-report questionnaire that measures the severity of depressive symptoms. The scale has a 0-63 range and responses are rated from 0 "absence of depressive symptoms" to 3 "severe levels of depression". The internal consistency of the BDI-II in the present sample was excellent (α = .92). The Overall Depression Severity and Impairment Scale (ODSIS; Bentley et al., 2014), validated in Spanish by Osma et al. (2019), is a 5-item measure that evaluates the frequency and intensity of depressive symptoms and their interference with the person's work or school life and social life. The total score ranges from 0 to 20 and responses use a 5-point Likert scale ranging from 0 (None) to 4 (All the time). The internal consistency of the ODSIS in the present sample was α = .92.

Anxiety. Anxiety was measured with the Beck Anxiety Inventory and the Overall Anxiety Severity and Impairment Scale. The Beck Anxiety Inventory (BAI; Beck & Steer, 1993), validated into Spanish by Sanz et al. (2012), is a 21-item measure of anxiety symptoms. The total score in the BAI ranges from 0 to 63 and responses use a 4-point Likert scale ranging from 0 (Not at all) to 3 (Severely). The internal consistency of the BAI in the present sample was α = .92. The Overall Anxiety Severity and Impairment Scale (OASIS; Norman et al., 2006), validated in Spanish by Osma et al. (2019), is a 5-item measure that evaluates the frequency and intensity of anxiety

symptoms and their interference with the person's work or school life and social life. All items are rated on a 5-point Likert scale ranging from 0 (None) to 4 (All the time). The total scale score ranges from 0 to 20. In the present sample, the Cronbach's alpha was α = .92.

Quality of life. The Quality of Life Index (QLI; Mezzich et al., 2000) consists of 10 items that evaluate different dimensions of quality of life. The total score ranges from 10 to 100 and each item is evaluated on a 10-point numerical rating scale ranging from 1 (Bad) to 10 (Excellent). The Cronbach's alpha in the present sample was α = .86.

Procedure

The ethical and research committees of all collaborating centers approved this ongoing multicenter randomized controlled trial NCT03064477 (March 10, 2017). Participants were stratified according to the severity of anxiety and depression symptomatology (evaluated with the BDI-II and the BAI). Next, using a computer-generated sequence (Randomizer), participants were randomized to the UP or the TAU condition. Each center had their own randomization list, which was provided by an external researcher not belonging with the center.

Inclusion criteria to participate in the study were: (a) having a principal (most interfering and severe) ED disorder diagnosis (i.e., anxiety disorder, mood disorder, adjustment disorder, among others); (b) being aged 18 or over; (c) being fluent in the language in which therapy is performed (Spanish or Catalan in the present study); (d) being able to attend to the evaluation and treatment sessions and signing the informed consent form; (e) patients taking pharmacological treatment were asked to maintain the same dosages and medications for at least 3 months prior to enrolling into the study and during the whole treatment. Exclusion criteria were: (a) presence of a severe condition

that would require to be prioritized for treatment, so that an interaction between both interventions could not be ruled out. These include a severe mental disorder (bipolar disorder, schizophrenia, or an organic mental disorder), suicide risk at the time of assessment, or substance use in the last three months; (b) the patient has previously received 8 or more sessions of psychological treatment with clear and identifiable Cognitive Behavioral Therapy (CBT) principles within the past 5 years.

Regarding therapist expertise, all therapists were licensed psychologists with between 8 and 20 years of experience (n = 8, 44.4%) or clinical psychology residents with 2 to 4 years of experience (n = 10, 55.6%) in delivering CBT in an individual format (TAU). In addition, all therapists and co-therapists in the UP groups received a UP training workshop before starting the intervention. This consisted of 2 or 3 group workshop sessions where therapists were instructed on the delivery of the different treatment modules of the UP. The duration of the course was between 10 and 20 hours, depending on the availability of the center's therapists. In addition to the workshop, all therapists received an individual training during 12 therapy sessions. The individual training consisted of an online supervision before each session or the participation as a co-therapist with an expert delivering the UP intervention, who also assesses treatment fidelity. In both cases, the training was conducted by the leading author, Jorge Osma, who has been certified as a UP Researcher/Trainer by the Unified Protocol Institute.

The 8 treatment modules of the UP were administered during 12 two-hour group treatment sessions, at a rate of one session per week. The treatment modules included: (1) Setting goals and maintaining motivation; (2) Understanding the adaptability of emotions; (3) Mindful emotion awareness; (4) Cognitive flexibility; (5) Countering emotional behaviors; (6) Understanding and confronting physical sensations; (7) Exposure to emotions; (8) Recognizing accomplishments and looking to the future (for

more information see Osma et al. 2018). The UP treatment was conducted over approximately three months. If clinicians considered that there was a need to provide additional sessions to the participants after treatment completion due their clinical status, they were allowed to do so in an individual format (each session would have an approximate duration of 35 minutes). This information was collected and will be described later in the corresponding section. Groups were composed of eight to ten participants and two clinicians (therapist and co-therapist). All participants in the UP condition received the UP workbook (Barlow et al., 2011b).

The TAU at the collaborating public mental health centers is individual, non-protocolized CBT plus pharmacological treatment (i.e., antidepressants and/or anxiolytics). In the TAU condition, participants received the number of sessions established by their therapists according to their clinical judgment and the availability of therapists based on waiting lists in the collaborating settings. The duration of sessions was, on average, 35 minutes each (*SD* = 5.05, ranging from 30 to 40 minutes depending on the center). The most commonly used techniques in this treatment condition included: psychoeducation, cognitive restructuring, exposure techniques, and activity scheduling (used by 90% of the therapists), followed by mindfulness techniques (80%), relaxation techniques (70%), problem-solving training (70%), and communication skills training (50%).

Participants from both conditions were patients who attended mental health units for psychological treatment and were not compensated economically for their participation. Irrespective of the assigned condition, participants completed four evaluations at the same times, that is, at pre-treatment, at 3 months after treatment onset (coinciding with the end of the UP treatment), and 6 and 9 months after treatment onset (follow-ups). Assessments were conducted by trained evaluators who were blind to the

number of sessions performed and the clinical status of the participants in the TAU condition (e.g., clinical discharge or still in treatment). This means that, at the end of the month 9 after treatment onset, patients in the TAU condition received 9 months of episodical (i.e., often monthly, as will be described later) individual CBT, while patients in the UP condition received an intensive 3-month UP treatment in group format. Participants in both conditions were allowed to use pharmacological treatment in addition to psychological intervention. However, they were asked to maintain the same dosages from three months prior to the onset of the intervention and until the end of the study.

Data analysis

Analyses were carried out using SPSS version 22.0 (IBM Corp, 2013). First, the sociodemographic characteristics of the sample were analyzed (n=488). Next, comparisons of patients who received the TAU and the UP condition were investigated using a multivariate analysis of variance (MANOVA) for continuous variables and a chi-square test for categorical variables. The same procedure will be used to compare the characteristics between the participants in the UP condition who required extra treatment sessions and those who did not.

Following this, a linear mixed model analysis with compound symmetry as a covariance structure (AL-Marshadi, 2014) was performed. Main effects of time were used to assess if there were improvements in outcomes throughout the study. In addition to the main effect of time, we also included the main effects of treatment condition and number of sessions. This was done to calculate interaction effects (treatment condition*time, as well as treatment condition*number of sessions* time) that would reveal whether the treatment condition and the number of sessions received interacted with time in the prediction of changes in study outcomes (i.e.,

whether changes in outcomes were different across treatment conditions and dependent on the number of treatment sessions received).

Next, we conducted a post hoc analysis when a significant interaction was observed in the mixed model analyses. To do so, we categorized the number of sessions using a half standard deviation below and above our sample mean (i.e., ≤ 8 sessions, 9 to 12 sessions, and > 12 sessions). Similar to past research (Kerig et al., 2011), by doing this we expected to distinguish subgroups of patients who presented a different evolution in study variables according to the number of sessions they received.

As a final step, as recommended by the literature in multicenter studies (Leichsenring et al., 2013), center effects were included as a random effect in the linear mixed model. Specifically, we included as Center*Time interaction in order to assess the magnitude of the random variance(s) at the between and within-center levels.

Results

Treatment sessions received over time in UP and TAU conditions

The average number of treatment sessions 9 months after treatment onset was 12.49 (*SD* = 2.18, range = 8-18) in the UP group and 6.07 (*SD* = 2.51, range = 3-14) in the TAU condition. After the end of the UP treatment, 18 (7.14%) participants in the UP condition were provided with additional individual UP sessions (mean = 2.22 sessions, *SD* = 1.51, range = 1-5).

Sociodemographic characteristics of the sample and differences between treatment conditions before treatment onset (baseline assessment)

The MANOVA indicated no significant baseline differences in any of the study variables, including age, the BDI, the BAI, and remaining study outcomes when comparing the TAU and the UP ($p > .05$). No baseline differences were found in gender ($\chi 2$ (1) = .106, p = .744) and clinical diagnoses ($\chi 2$ (2) = 1.174, p = .556) between conditions either. Similarly, no differences were found between the participants who required extra UP treatment sessions and those who did not in any of the study outcomes ($p > .05$) and sociodemographic variables, including age (F (1) = .018, p = .895), gender ($\chi 2$ (1) = 2.498, p = .114), and clinical diagnosis ($\chi 2$ (14) = 10.779, p = .703).

Changes in personality, affect, depression, anxiety, and quality of life after the interventions

As seen in Table 2, there was a significant effect of time after the UP application. We found moderate-to-large effect sizes of the UP on neuroticism (F = 48.41, $p < .001$, dof = 488.361, pretreatment to 9 months after treatment onset Cohen's d = 0.72), negative affect (F = 27.66, $p < .001$, dof = 532.93, pretreatment to 9 months after treatment onset Cohen's d = 0.76), the BDI-II (F = 109.29, $p < .001$, dof = 482.15, pretreatment to 9 months after treatment onset Cohen's d = 0.87), the ODSIS (F = 70.74, $p < .001$, dof = 490.04, pretreatment to 9 months after treatment onset Cohen's d = 0.78), the BAI (F = 60.30, $p < .001$, dof = 487.30, pretreatment to 9 months after treatment onset Cohen's d = 0.69), the OASIS (F = 78.49, $p < .001$, dof = 499.65, pretreatment to 9 months after treatment onset Cohen's d =0.90), and the QLI (F = 63.53, $p < .001$, dof = 491.07, pretreatment to 9 months after treatment onset Cohen's d = -0.83). The results also revealed significant moderate effects of the UP on extraversion (F = 19.41, $p < .001$, dof = 470.81, pretreatment to 9 months after

treatment onset Cohen's d = -0.41) and positive affect (F = 25.59, p < .001, dof = 504.20, pretreatment to 9 months after treatment onset Cohen's d = -0.57).

Regarding the TAU condition, a significant effect of time was also observed. Effect sizes were moderate on neuroticism (F = 14.17, p < .001, dof = 392.42, pretreatment to 9 months after treatment onset Cohen's d = 0.45), negative affect (F = 20.22, p < .001, dof = 427.81, pretreatment to 9 months after treatment onset Cohen's d = 0.61), positive affect (F = 5.70, p = .001, dof = 393.69, pretreatment to 9 months after treatment onset Cohen's d = -0.27), the BDI-II (F = 48.73, p < .001, dof = 386.22, pretreatment to 9 months after treatment onset Cohen's d = 0.70), the ODSIS (F = 12.09, p < .001, dof = 384.74, pretreatment to 9 months after treatment onset Cohen's d = 0.41), the BAI (F = 25.31, p < .001, dof = 392.16, pretreatment to 9 months after treatment onset Cohen's d = 0.49), the OASIS (F = 12.31, p <.001, dof = 392.09, pretreatment to 9 months after treatment onset Cohen's d = 0.49) and the QLI (F = 18.46, p < .001, dof = 394.56, pretreatment to 9 months after treatment onset Cohen's d = -0.47). No significant changes were observed on extraversion after the TAU intervention (F = .43, p = .733, dof = 391.01, pretreatment to 9 months after treatment onset Cohen's d = -0.12).

As indicated in Table 2, relevant improvements mostly emerged when comparing the pre-treatment and 3 months after treatment onset (coinciding with end of the UP treatment), especially in the case of the UP condition (all effect sizes over 0.30). The improvements accomplished after 3 months of treatment were generally maintained after treatment termination (during the follow-ups, which occurred 3 and 6 months after the UP treatment finished). This was evidenced by the small effect sizes presented in Table 2 when comparing the 3 months after treatment onset (end of UP treatment) with the 6-month after treatment onset (3 months after UP termination). In this case, the

Cohen's *d*'s were generally close to 0 or smaller than 0.30, except for negative affect in the TAU condition, which obtained a Cohen's *d* of 0.35 (continued to improve). Similar findings were obtained when comparing the results 6 and 9 months after treatment onset, where Cohen's *d* was generally close to 0 or smaller than 0.30.

-Insert Table 2 about here-

Main effects and interaction effects of treatment condition and number of sessions

The results of the main effects and interaction effects of the treatment condition and the number of sessions can be found in Table 3. A main effect of treatment condition was found on the BDI-II (F = 9.01, p = .003, dof = 972.16, Cohen's d = 0.27), the ODSIS (F = 8.10, p = .005, dof = 994.24, Cohen's d = 0.26), the BAI (F = 9.52, p = .002, dof = 986.50, Cohen's d = 0.28), the OASIS (F = 9.92, p = .002, dof = 1016.50, Cohen's d = 0.29), and the QLI (F = 3.93, p = .048, dof = 1018.56, Cohen's d = 0.18). The post hoc analyses of the main effect of treatment condition can be seen in detail in Appendix 2 (supplementary material). The main effect of the number of sessions was not significant on any outcome (p > .05).

-Insert Table 3 about here-

In terms of the interaction effects, a Time*condition effect was evidenced for the BDI-II (F = 4.67, p = .003, dof = 695.11, Cohen's d = 0.20), the ODSIS (F = 4.05, p = .007, dof = 719.88, Cohen's d = 0.18), the BAI (F = 3.96, p = .008, dof = 713.99, Cohen's d = 0.18), the OASIS (F = 4.56, p = .004, dof = 740.59, Cohen's d = 0.20) and the QLI (F = 3.01, p = .030, dof = 736.02, Cohen's d = 0.16). These findings together with an inspection of Table 2 evidenced that the UP resulted in larger improvements in depression (BDI-II and ODSIS), anxiety (BAI and OASIS), and quality of life when

compared with the TAU. In contrast to this, both interventions produced comparable changes in neuroticism, negative affect, extraversion, and positive affect.

Finally, the linear mixed model analyses showed an interaction effect of Time*Condition*Sessions for the BDI-II (F = 2.82, p =.016, dof = 672.88, pretreatment to 9 months after treatment onset Cohen's d = 0.15) and the ODSIS (F = 3.03, p =.010, dof = 693.72, pretreatment to 9 months after treatment onset Cohen's d = 0.16). Post hoc analyses were carried out considering the treatment condition and subgroups according to the number of sessions received (less or equal to 8, from 9 to 12, and more than 12 sessions). The results are reported in Table 4.

In the UP condition, a statistically significant main effect of time was found in the BDI-II in participants who received 9 to 12 sessions (F = 27.77, p < .001) and more than 12 sessions (F = 19.59, p < .001). A statistically significant main effect of time was also found in the ODSIS in subgroups who received 9 to 12 sessions (F = 21.26, p < .001) and more than 12 sessions (F = 15.68, p < .001). The strongest effects of time were observed in the subgroup of participants who received 9 to 12 sessions (Cohen's d over 0.96 in both cases).

In the TAU condition, only the subgroup of participants who received 8 sessions or less showed statistically significant improvements in the BDI-II (F = 38.14, p < .001, Cohen's d = 0.72) and the ODSIS (F = 10.45, p < .001, Cohen's d = 0.52).

-Insert Table 4 about here-

Random effects of Center on study outcomes

The random effects erer statistically non-significant for neuroticism (*Wald Z* = 1.00, p = .317, 95% confidence interval: 0.07 to 3.44), negative affect (*Wald Z* = 1.88, p = .059, 95% confidence interval: 0.77 to 6.16), positive affect (*Wald Z* = 1.13, p = .258,

95% confidence interval: 0.11 to 3.50), the ODSIS (*Wald Z* = 1.89, *p* = .059, 95% confidence interval: 0.07 to 3.44), and the BAI (*Wald Z* = 0.63, *p* = .527, 95% confidence interval: 0.03 to 13.96). For extraversion, the BDI, the OASIS and the QLI convergence issues were found, even when interactions were increased to 3000.

Discussion

To the best of our knowledge, this is the first multicenter randomized controlled trial study conducted in public mental health settings in Spain to show the effectiveness of the UP for the transdiagnostic treatment of EDs in group format. In relation to the effectiveness of the treatments, we hypothesized that both the UP and the TAU would result in significant moderate-to-large improvements in study outcomes. The results obtained in this study have confirmed this hypothesis, with the only exception of extraversion in the case of TAU.

Regarding the results of the UP, we expected that the UP would produce significant improvements with moderate-to-large effects sizes on study outcomes. Consistent with the former idea and in line with past research (Grill et al., 2017; Ornelas et al., 2017; Osma et al. 2015), we found a large positive impact of the UP on all outcomes. Importantly, these changes were maintained in the follow-up periods even though the large majority of the participants in the UP condition (92.86%) had not received additional treatment during the follow-ups (6 and 9 months after treatment onset). Of particular interest for the present study was the finding that improvements in depression (BDI-II and ODSIS), anxiety (BAI and OASIS), and quality of life were larger in the UP condition than in the TAU group. These results add to the existing research supporting the potential clinical utility of the UP in group format for the treatment of EDs in public mental health settings (Reinholt et al., 2017).

In addition to depression and anxiety, another relevant study finding was that the UP produced large reductions in neuroticism and negative affect, which is again consistent with previous literature (Ornelas et al., 2017; Osma et al., 2015). These findings support the idea that the UP is a useful intervention to address emotional dysregulation (high neuroticism), a mechanism believed to be shared by all patients with EDs (Barlow et al., 2014). Also consistent with past research (Grill et al., 2017; Laposa et al., 2017; Osma et al., 2015; Reinholt et al., 2017), the present study results revealed moderate effects of the UP on positive affect. Different to a previous investigation with a much smaller sample (*n*=6; Osma et al., 2015), however, we found a significant moderate effect of the UP on extraversion. While the discrepancies between findings might be attributable to small sample size in the previous investigation, there is previous evidence to support the social benefits of the group format (Burlingame et al., 2013), which might explain the positive results obtained in extraversion after the UP group intervention. Finally, a large effect of the UP was observed on quality of life. This result is again consistent with recent findings from a systematic review and meta-analysis of UP studies, including those applied in group format (Sakiris & Berle, 2019), and supports the utility of the UP delievered in group for the improvement of wide range of outcomes.

It is important to note that the usual psychological treatment in public mental health settings in Spain generally occurs with a 4-to-6 week frequency. Therefore, therapy can last several months or even years and distress due to delayed consecutive appointments – and therefore delayed training of psychological skills - cannot be ruled out. Contrary to this, changes with the UP occurred rapidly after an intensive 3-month intervention and were maintained even though patients were no longer being treated. Therefore, the UP in a group format might provide a more feasible, ethical and effective

alternative as groups can be more easily formed thanks to the inclusion of a wide range of ED diagnoses and allow treating several individuals at the same time with little additional resources (i.e., a co-therapist and a larger room). Ultimately, the adoption of a group format and the selection of the UP as the treatment option mean that patients will receive more treatment sessions in a reduced period of time compared to other existing treatment options (individual therapy or disorder-specific CBT group interventions). Additionally, by implementing UP groups more patients can be treated in less time. Again, this may help alleviate some of the issues associated with limited resources that make the implementation of empirically-supported psychological treatments in public mental health systems difficult.

In addition to the effects of the UP compared with the TAU, the present investigation also explored whether the number of sessions received was a determinant factor moderating the effect that treatments had on outcomes. In this regard, the results support the idea that this variable is not necessarily a decisive one for treatment effectiveness. An interesting finding, however, was observed in relation to the interaction between the number of sessions and the treatment condition in the prediction of depression (BDI-II and ODSIS). In the UP condition, receiving 9 to 12 sessions was associated with greater improvements, while approximately 8 sessions were the optimal in the TAU condition. At an equal number of sessions (9 to 12 sessions), the UP condition obtained the best results. While these results should be taken with caution due to their novelty, we encourage researchers to include the number of sessions into their models when comparing the effectiveness of the UP with other active controls because this might shed light into the recommended number of sessions to achieve the desired results in individuals with EDs.

It is important to note that only a small percentage of participants in the UP condition (7.1%) required additional treatment after the 12-session intervention. The effect of this additional number of sessions was controlled by including this variable (main effect and interaction effect) in the prediction of outcomes. Most importantly, what this small prevalence of participants requiring additional treatment means is that the majority of participants with ED diagnosis found sufficient help with the pre-established intervention. This result is relevant for public health care managers and authorities because it suggests that only a very small proportion of participants will probably require additional resources when the UP in group format is implemented. In addition, the results of this study are likely to be generalizable to mental health systems that are in a situation of collapse or lack of resources. Attending to World Health Organization (WHO) estimates, depression and anxiety disorders have increased between 14.9% and 18.4% from 2005 to 2015 worldwide (WHO, 2017), so if this growing trend is maintained over time, all mental health systems, at least to some extent, will need to cope with this large demand for psychological care by offering effective and efficient treatments. The results of this study can help in this direction.

The results of this study should be considered in light of a number of limitations. First, the assessments were too long. Specifically, it took approximately 3 hours to complete the first evaluation and approximately 45 minutes to fill the rest of the evaluations, since the ADIS was not administered systematically by all clinicians at all assessment times (it was only required during the pre-evaluation to establish clinical diagnoses [inclusion criteria]). This may have reduced some participant's willingness to complete study evaluations. In this sense, the inclusion of a shorter diagnostic interview, such as the MINI (Sheehan et al., 2015) as opposed to the ADIS (Di Nardo et al., 1994) should be considered in future research conducted in public mental health settings.

Similarly, several measures of depression and anxiety symptoms were used here because of the relatively recent nature of the ODSIS and the OASIS. However, there is now enough evidence to support their use as measures of depression and anxiety symptoms (and interference) and both are recommended in the assessment of change with the UP because they were developed according to this psychological perspective (Osma et al., 2019). These changes would facilitate the work of clinical psychologists in the public system because they are much shorter (5 items each) than other traditional measures of depression and anxiety, such as the BDI-II or the BAI (21 items each). Additionally, because this study took place across several public mental health settings, the results might not be generalizable to other contexts. The study took place in many centers (10) and each center contributed with a very different number of participants (from 28 to 129 participants). Additionally, some differences in the training of therapists existed. This could have resulted in patients from one specific center to improve more than those from another center. Fortunately, the center analyses showed no significant center*time effects, which means that comparable improvements were achieved regardless of the center. Another aspect to highlight is the difficulty to describe the TAU intervention in a precise way due to the heterogeneity and variety of techniques used. As some authors suggest, it is necessary to clarify and take this into account in order to generalize the results (Wampold et al., 2011). However, the fact that this study was performed in public health settings as opposed to tightly controlled clinical research contexts is one of the novelties and strengths of this investigation. Lastly, while this study included a follow-up evaluation 9 months after treatment onset (6 months after UP termination), even longer follow-up periods (i.e., 12-months) would be desirable to ensure that results are maintained after treatment completion.

To sum up, the results from this study suggest that an intensive (12-sessions over 3 months) UP group intervention is an effective solution for the treatment of EDs. These findings are especially important due the alarmingly high prevalence of EDs in our society and their associated direct and indirect costs (Ruiz-Rodríguez et al., 2017). It is therefore necessary to develop and apply new forms of psychological evidence-based treatments, as well as to improve the cost-benefit relationship of psychological interventions, especially in public settings where resources are limited. The UP could be one option in this direction.

Data availability statement

The data that support the findings of this study are available from the corresponding author, [Jorge Osma], upon reasonable request.

References

AL-Marshadi, A. H. (2014). Selecting the Covariance Structure in Mixed Model using Statistical Methods Calibration. *Journal of Mathematics and Statistics*, *10*(3), 309–315. https://doi.org/10.3844/jmssp.2014.309.315

Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & Ehrenreich-May, J. (2011a). *The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Therapist Guide.* Oxford University Press.

Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & Ehrenreich-May, J. (2011b). *The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Client Workbook.* Oxford University Press.

Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2014). The Nature, Diagnosis, and Treatment of Neuroticism. *Clinical Psychological Science*, *2*(3), 344–365. https://doi.org/10.1177/2167702613505532

Beck, A., & Steer, R. (1993). *Beck Anxiety Inventory manual*. Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI-II, Beck depression inventory: manual second edition*. Psychological Corporation.

Bentley, K. H., Gallagher, M. W., Carl, J. R., & Barlow, D. H. (2014). Development and validation of the Overall Depression Severity and Impairment Scale. *Psychological Assessment, 26*(3), 815–830. https://doi.org/10.1037/a0036216

Boswell, J. F., Sauer-Zavala, S. E., Gallagher, M. W., Delgado, N. K., & Barlow, D. H. (2012). Readiness to change as a moderator of outcome in transdiagnostic treatment. *Psychotherapy Research, 22*(5), 570–578.

Botella, C., & Ballester, R. (1997). *Trastorno de pánico, evaluación y tratamiento [Panic disorder, evaluation and treatment]* (4ed., text). Martínez Roca.

Brown, T. A., & Barlow, D. H. (2009). A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. *Psychological Assessment, 21*(3), 256–271. https://doi.org/10.1037/a0016608

Brown, T. A., Di Nardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of DSM-IV anxiety and mood disorders: implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 110*(1), 49.

Bullis, J. R., Boettcher, H., Sauer-Zavala, S., Farchione, T. J., & Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clinical Psychology: Science and Practice, 26*(2), e12278. https://doi.org/10.1111/cpsp.12278

Bullis, J. R., Sauer-Zavala, S., Bentley, K. H., Thompson-Hollands, J., Carl, J. R., & Barlow, D. H. (2015). The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Preliminary Exploration of Effectiveness for Group Delivery. In *Behavior Modification* (Vol. 39, Issue 2). Oxford University Press. https://doi.org/10.1177/0145445514553094

Burlingame, G. M., Strauss, B., & Joyce, A. (2013). Change mechanisms and effectiveness of small group treatments. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 640–689). Wiley.

Cassiello-Robbins, C., Southward, M. W., Tirpak, J. W., & Sauer-Zavala, S. (2020). A systematic review of Unified Protocol applications with adult populations: Facilitating widespread dissemination via adaptability. *Clinical Psychology Review*, 101852. https://doi.org/10.1016/j.cpr.2020.101852

Costa, P. T., & McCrae, R. R. (1999). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. TEA Ediciones.

de Ornelas, C. A. C., Egidio, A., & Cardoso, A. (2015). The utilization of unified protocols in behavioral cognitive therapy in transdiagnostic group subjects : A clinical trial. *Journal of Affective Disorders*, *172*, 179–183. https://doi.org/10.1016/j.jad.2014.09.023

de Ornelas, A. C. C., Sanford, J., Boettcher, H., Nardi, A. E., & Barlow, D. (2017). Improvement in quality of life and sexual functioning in a comorbid sample after the unified protocol transdiagnostic group treatment. *Journal of Psychiatric Research*, *93*, 30–36. https://doi.org/10.1016/j.jpsychires.2017.05.013

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV: Lifetime Version (ADIS-IV-L)*. Oxford University Press.

Grill, S. S., Castañeiras, C., & Fasciglione, M. P. (2017). Aplicación grupal del Protocolo Unificado para el tratamiento transdiagnóstico de los trastornos emocionales en población Argentina. *Revista de Psicopatologia y Psicologia Clinica*, *22*(3), 171–181. https://doi.org/10.5944/rppc.vol.22.num.3.2017.18122

IBM Corp. (2013). *IBM SPSS Statistics for Windows, Version 22.0.* (No. 22). IBM Corp.

Kerig, P. K., Schulz, M. S., & Hauser, S. T. (2011). *Adolescence and beyond: Family processes and development*. Oxford University Press.

Laposa, J. M., Mancuso, E., Abraham, G., & Loli-Dano, L. (2017). Unified Protocol Transdiagnostic Treatment in Group Format: A Preliminary Investigation With Anxious Individuals. *Behavior Modification*, *41*(2), 253–268. https://doi.org/10.1177/0145445516667664

Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., Huesing, J., Joraschky, P., Nolting, B., Poehlmann, K., Ritter, V., Stangier, U., Strauss, B., Stuhldreher, N., Tefikow, S., Teismann, T., Willutzki, U., Wiltink, J., & Leibing, E. (2013). Psychodynamic Therapy and Cognitive-Behavioral Therapy in Social Anxiety Disorder: A Multicenter Randomized Controlled Trial. *American Journal of Psychiatry*, *170*(7), 759–767. https://doi.org/10.1176/appi.ajp.2013.12081125

Mezzich, J. E., Ruipérez, M. A., Pérez, C., Yoon, G., Liu, J., & Mahmud, S. (2000). The Spanish Version of the Quiality of Life Index. *The Journal of Nervous and Mental Disease*, *188*(5), 301–305.

Norman, S. B., Hami Cissell, S., Means-Christensen, A. J., & Stein, M. B. (2006). Development and validation of an Overall Anxiety Severity And Impairment Scale (OASIS). *Depression and Anxiety*, *23*(4), 245–249. https://doi.org/10.1002/da.20182

Osma, J., Castellano, C., Crespo, E., & García-Palacios, A. (2015). The unified protocol for transdiagnostic treatment of emotional disorders in group format in a spanish public mental health setting. *Behavioral Psychology/ Psicologia Conductual*, *23*(3), 447–466. https://doi.org/10.1177/0145445514553094

Osma, J., Quilez-Orden, A., Suso-Ribera, C., Peris-Baquero, O., Norman, S. B., Bentley, K. H., & Sauer-Zavala, S. (2019). Psychometric properties and validation of the Spanish versions of the overall anxiety and depression severity and impairment scales. *Journal of Affective Disorders*, *252*, 9–18. https://doi.org/10.1016/j.jad.2019.03.063

Osma, J., Suso-Ribera, C., García-Palacios, A., Crespo-Delgado, E., Robert-Flor, C., Sánchez-Guerrero, A., Ferreres-Galan, V., Pérez-Ayerra, L., Malea-Fernández, A., & Torres-Alfosea, M. Á. (2018). Efficacy of the unified protocol for the treatment of emotional disorders in the Spanish public mental health system using a group format: study protocol for a multicenter, randomized, non-inferiority controlled trial. *Health and Quality of Life Outcomes*, *16*(1), 46. https://doi.org/10.1186/s12955-018-0866-2

Reinholt, N., Aharoni, R., Winding, C., Rosenberg, N., Rosenbaum, B., & Arnfred, S. (2017). Transdiagnostic group CBT for anxiety disorders: the unified protocol in mental health services. *Cognitive Behaviour Therapy*, *46*(1), 29–43. https://doi.org/10.1080/16506073.2016.1227360

Ruiz-Rodríguez, P., Cano-Vindel, A., Muñoz Navarro, R., Medrano, L., Moriana, J. A., Buiza Aguado, C., Jiménez Cabré, G., & González-blanch, C. (2017). Impacto económico y carga de los trastornos mentales comunes en España: Una revisión sistemática y crítica. *Ansiedad y Estres*, *23*(2–3), 118–123. https://doi.org/10.1016/j.anyes.2017.10.003

Sakiris, N., & Berle, D. (2019). A systematic review and meta-analysis of the Unified Protocol as a transdiagnostic emotion regulation based intervention. *Clinical Psychology Review*, *72*, 101751. https://doi.org/10.1016/j.cpr.2019.101751

Sandín, B., Chorot, P., Lostao, L., Joiner, T. E., Santed, M. A., & Valiente, R. M. (1999). Escalas PANAS de afecto positivo y negativo: validación factorial y convergencia transcultural. *Psicothema*, *11*(1), 37–51.

Sanz, J., García-Vera, M. P., & Fortún, M. (2012). The Beck Anxiety Inventory (BAI): Psychometric properties of the Spanish version in patients with psychological disorders. *Behavioral Psychology/ Psicologia Conductual*, *20*(3), 563–583.

Sanz, J., Navarro, M. E., & Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II). Propiedades psicométricas en estudiantes universitarios. *Análisis y Modificación de Conducta*, *29*(124), 239–288.

Saxena, S., Thornicroft, G., Knapp, M., & Whiteford, H. (2007). Resources for mental health: scarcity, inequity, and inefficiency. *The Lancet*, *370*(9590), 878–889. https://doi.org/10.1016/S0140-6736(07)61239-2

Sheehan, D., Janavs, J., Baker, R., Sheehan, K. H., Knapp, E., & Sheehan, M. (2015). *Mini international neuropsychiatric interview–version 7.0. 0 DSM-5*. Medical Outcomes Systems.

Thompson-Brenner, H., Boswell, J. F., Espel-Huynh, H., Brooks, G., & Lowe, M. R. (2019). Implementation of transdiagnostic treatment for emotional disorders in residential eating disorder programs: A preliminary pre-post evaluation. *Psychotherapy Research*, *29*(8), 1045–1061. https://doi.org/10.1080/10503307.2018.1446563

Varkovitzky, R. L., Sherrill, A. M., & Reger, G. M. (2018). Effectiveness of the Unified Protocol for Transdiagnostic Treatment of Emotional Disorders Among Veterans With Posttraumatic Stress Disorder: A Pilot Study. *Behavior Modification*, *42*(2), 210–230. https://doi.org/10.1177/0145445517724539

Wampold, B. E., Budge, S. L., Laska, K. M., Del Re, A. C., Baardseth, T. P., Flűckiger, C., Minami, T., Kivlighan II, D. M., & Gunn, W. (2011). Evidence-based treatments for depression and anxiety versus treatment-as-usual: a meta-analysis of direct comparisons. *Clinical Psychology Review*, *31*(8), 1304–1312.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063.

World Health Organization. (2017). *Depression and other common mental disorders: Global health estimates*. Author.

Zemestani, M., Imani, M., & Ottaviani, C. (2017). A preliminary investigation on the effectiveness of unified and transdiagnostic cognitive behavior therapy for patients with comorbid depression and anxiety. *International Journal of Cognitive Therapy*, *10*(2), 175–185. https://doi.org/10.1521/ijct.2017.10.2.175