MANUSCRIPT PUBLISHED IN THE PSYCHOLOGICAL RECORD

Acceptance and commitment therapy focused on repetitive negative thinking for child depression: A randomized multiple-baseline evaluation

Daniela M. Salazar[1] (danielasalazar1989@gmail.com)

Francisco J. Ruiz[1] (franciscoj.ruizj@konradlorenz.edu.co)

Eduar S. Ramírez[1] (eduars.ramirez@gmail.com)

Verónica Cardona-Betancourt[2] (veritofbl@hotmail.com)

[1]Fundación Universitaria Konrad Lorenz

[2]Universitat Oberta de Catalunya

Declarations of interest: none

Correspondence address: Francisco J. Ruiz, franciscoj.ruizj@konradlorenz.edu.co, Fundación Universitaria Konrad Lorenz, Carrera 9 bis, Nº 62-43, Bogotá (Cundinamarca, Colombia), Phone: (+57 1) 347 23 11 ext. 185

Abstract

The current study analyzes the efficacy of acceptance and commitment therapy (ACT) focused on repetitive negative thinking (RNT) for child depression. A randomized, nonconcurrent, multiple-baseline design was conducted with 9 children, aged between 8 and 13 years, who showed a main diagnosis of child depression. Measures of psychological inflexibility, RNT, and generalized pliance were administered on a weekly basis throughout the study, whereas measures of emotional symptoms and parents' report of problematic behavior were applied at pretreatment, posttreatment, and the 4-week follow-up. All participants showed evidence of a treatment effect for psychological inflexibility and RNT. The standardized mean difference effect sizes for single-case experimental designs were very large for these measures. No participant showed the diagnosis of child depression or comorbid disorders at the 4-week follow-up. Pretreatment to follow-up changes in emotional symptoms and problematic behavior reported by parents were statistically significant, with large effect sizes. RNT-focused ACT interventions for child depression deserve further empirical tests.

Key words: Child depression; Acceptance and commitment therapy; Relational frame theory; Repetitive negative thinking.

Acceptance and commitment therapy focused on repetitive negative thinking for child depression: A randomized multiple-baseline evaluation

Recent estimations show that approximately 1 to 3% of children suffer from depression (e.g., Costello, Erkanli, & Angold, 2006; Egger & Angold, 2006; Ford, Goodman, & Meltzer, 2003), with few gender differences. These rates of depression contrast significantly with those found in adolescents and adults, where females show higher rates of depression, and the prevalence is about 5 to 10% in most studies (e.g., Avenevoli, Swendsen, He, Burstein, & Merikangas, 2015; Kessler & Bromet, 2013; Kessler et al., 2003; Merikangas et al., 2010; Thapar, Collishaw, Pine, & Thapar, 2012). In spite of its lower prevalence, child depression is an important concern because of its high comorbidity with other emotional and behavioral disorders (Maughan, Collishaw, & Stringaris, 2013) and its persistence. For instance, one-third of the children identified with clinically significant levels of depressive symptoms remained symptomatic in a 2-year longitudinal study (DuBois, Felner, Bartels, & Silverman, 1995). Additionally, child depression has been associated with important negative consequences such as low academic performance (DuBois et al., 1995), disrupted parent-child attachment (Brumariu & Kerns, 2010), poor physical health (Fekkes, Pijpers, Fredriks, Vogels, & Verloove-Vanhorick, 2006), unsatisfying social relationship (Perren & Alsaker, 2009), family dysfunction (Sander & McCarty, 2005), higher risk of alcohol problems (Maughan et al., 2013), and mortality by suicide (Rao, Weissman, Martin, & Hammond, 1993).

Psychological interventions for child depression have been considerably less investigated than for adolescent and adult depression, with no child treatment achieving a well-established empirical status (Weersing, Jeffreys, Do, Schwartz, & Bolano, 2017). Regarding cognitive behavioral therapy (CBT) as a broad intervention, Weersing et al.

found that, among the seven high-quality studies identified in their review, only one study showed positive findings favoring CBT versus waitlist or psychologically inert controls (Kahn, Kehle, Jenson, & Clark, 1990). Accordingly, CBT only met the criteria for possibly efficacious treatment for child depression. Less high-quality studies were found for behavior therapy (BT), which also met the criteria for possibly efficacious treatment.

Considering broader eligibility study criteria, Zhou et al. (2015) found that interpersonal therapy (IPT) and CBT had lower effects in treating child depression compared to adolescent depression. Nevertheless, they concluded that IPT and CBT should be considered as the best available approaches for child and adolescent depression. More recently, Yang et al. (2017) identified nine studies analyzing the efficacy of CBT against control conditions. At posttreatment, CBT was more effective than control conditions, but the weighted effect size was small to medium (d = 0.41, 95% CI [0.18, 0.64]. Additionally, subgroup analyses showed that CBT was more effective than nontreatment conditions but equally effective than waitlist or psychological placebo.

In summary, the treatments for child depression tested so far have obtained mixed evidence, with small to medium effect sizes. Accordingly, further research is needed to establish the efficacy of CBT, BT, and IPT for child depression. Also, new psychological interventions need to be developed and tested for child depression. Interestingly, few studies have been conducted testing the efficacy of contextual cognitive-behavioral therapies for child depression.

In recent years, there has been a growing interest in identifying transdiagnostic processes involved in emotional disorders and developing psychological interventions targeting them, such as acceptance and commitment therapy (ACT; Hayes, Strosahl, & Wilson, 1999), metacognitive therapy (MCT; Wells, 2009), and rumination-focused

cognitive-behavioral therapy for depression (RF-CBT; Watkins, 2016). For instance, ACT was initially developed for the treatment of experiential avoidance (Hayes & Wilson, 1994) and has been redefined in broader terms for the treatment of psychological inflexibility (Hayes & Strosahl, 2004). Regarding MCT and RF-CBT, they were developed focusing on dismantling dysfunctional patterns of worry and rumination, which have been included under the term repetitive negative thinking (RNT; Ehring & Watkins, 2008; Watkins, 2008). However, to our knowledge, none of these approaches have been analyzed in child depression: ACT has been analyzed in the treatment of depression and anxiety disorders in adolescents (Hayes, Boyd, & Sewell, 2011; Petts, Duenas, & Gaynor, 2017; Swain, Hancock, Dixon, & Bowman, 2015), whereas MCT in the treatment of generalized anxiety disorders in children (Esbjørn, Normann, Christiansen, & Reinholdt-Dunne, 2018).

In the last few years, brief ACT protocols have been developed and tested in adults, which explicitly include the links among experiential avoidance, RNT, and psychological inflexibility (Dereix-Calonge, Ruiz, Sierra, Peña-Vargas, & Ramírez, 2019; Ruiz, et al., 2018; Ruiz, García-Beltrán, Monroy-Cifuentes, & Suárez-Falcón, in press; Ruiz, Luciano, Flórez, & Suárez-Falcón, submitted; Ruiz, Riaño-Hernández, Suárez-Falcón, & Luciano, 2016). This approach has been termed RNT-focused ACT. Briefly, psychological inflexibility entails the dominance of private experiences over chosen values and contingencies in guiding action (Bond et al., 2011). One of the main processes involved in psychological inflexibility is experiential avoidance, which is a pattern of verbal regulation based on deliberate efforts to either avoid or escape from discomfiting private experiences (Hayes, Wilson, Gifford, Follette, & Strosahl, 1996; Luciano & Hayes, 2001). Ruiz et al. (2016) suggested that RNT in the form of worry and rumination is an especially maladaptive experiential avoidance strategy because: (a) RNT tends to be the first reaction

to fear, the perception of not attaining personally relevant goals, and feelings of incoherence; (b) RNT tends to prolong negative affect (e.g., Newman & Llera, 2011), which usually leads to (c) engagement in additional experiential avoidance strategies in an attempt to reduce prolonged discomfort (e.g., Caselli et al., 2013; Nolen-Hoeksema, Stice, Wade, & Bohon, 2007; Wells, 2009); and (d) the repetition of this cycle generates an inflexible and maladaptive repertoire.

The practical implication of this account is that ACT protocols primarily focused on dismantling unconstructive RNT should produce quick changes and be particularly effective for the treatment of emotional disorders. The study by Ruiz et al. (2016) showed that a one-session, RNT-focused ACT protocol was sufficient to significantly reduce RNT, with very large effect sizes, among adult participants suffering from mild to moderate emotional disorders. Subsequent studies have shown that brief RNT-focused ACT protocol (2- to 3-session protocols) obtained very large effect sizes in treating moderate and severe emotional disorders, mainly depression and generalized anxiety disorders (Ruiz, et al., 2018; Ruiz, García-Beltrán, et al., in press; Ruiz, Luciano, et al., submitted). Additionally, Dereix-Calonge et al. (2019) showed that a web-based RNT-focused protocol was effective in reducing emotional symptoms and improving valued living in clinical psychology trainees compared to a waitlist control.

To our knowledge, RNT-focused ACT protocols have not been tested in children in spite of the fact that RNT is a frequent phenomenon in children (Henker, Whalen, & O'Neil, 1995; Păsărelu et al., 2016; Silverman, LaGreca, & Wasserstein, 1995). Specifically, rumination has been closely associated with concurrent levels of depressive symptoms in children (Abela, Vanderbilt, & Rochon, 2004) and predicts their increase over time (Abela, Aydin, & Auerbach, 2007; Abela, Brozina, & Haigh, 2002). Accordingly, the

aim of this study was to evaluate the efficacy of a 3-session RNT-focused ACT protocol in child depression. A nonconcurrent, randomized, multiple-baseline design was conducted where the effect of the protocol was directly replicated in 9 participants with the main diagnosis of child depression. The SCRIBE statement (Tate et al., 2016) was followed to guide the reporting of this single-case experimental design.

Method

Participants

Nine children aged between 8 and 13 years participated in the study. Participants were recruited through advertisements in social media beginning with the question: "Do you think your child is irritable or sad?" The parents of 15 children showed interest in the study and attended an assessment interview with their children. The inclusion criteria were: (a) child between 8-13 years old, (b) presenting the main diagnosis of child depression according to the Mini International Neuropsychiatric Interview for Kids and Adolescents (MINI KID; Sheehan, Shytle, Milo, Janavs, & Lecrubier, 2009) diagnostic interview and clinician's judgment, and (c) showing a verbal intelligent quotient (IQ) higher than 70 according to the Kaufman Brief Intelligence Test (K-BIT, Kaufman & Kaufman, 1990). The latter inclusion criterion was selected to guarantee that the child had a minimum verbal repertoire to conduct the intervention. The exclusion criteria were: (a) current psychological/psychiatric treatment, (b) having a psychological diagnosis prior to the study, and (c) presenting a high risk of suicide according to the MINI KID. The second exclusion criterion was adopted to avoid recruiting children with significant experience with assessment and therapeutic contexts, which might act as an extraneous variable. None participant was excluded for this reason.

The application of the inclusion and exclusion criteria led to the rejection of 6

potential participants: 1 individual was younger than 8 years and 5 did not meet the depression criteria as the main diagnosis. The final sample consisted of 9 participants (4 girls; mean age = 10.22, *SD* = 2.11). Table 1 shows the demographic data of the participants and diagnostic categories met. Six participants showed comorbid disorders (oppositional defiant disorder in 4 participants, attention-deficit hyperactivity disorder in 3, generalized anxiety disorder in 1, and separation anxiety in 1). Verbal IQ scores on the K-BIT ranged from 76 to 126 (*M* = 100.78, *SD* = 18.16)

INSERT TABLE 1 ABOUT HERE

Design and Variables

A three-arm, nonconcurrent, randomized multiple-baseline design across participants was implemented. Each cohort consisted of 3 participants. Participants in Cohort 1 received the intervention after collecting 4 weeks of baseline, participants in Cohort 2 after collecting 5 weeks of baseline, and participants in Cohort 3 after collecting 6 weeks of baseline. The randomization was conducted using the web-based tool Research Randomized (Urbaniak & Plous, 2013). The implemented randomization procedure was conducted because it significantly improves the internal validity of multiple baseline designs (Kratochwill & Levin, 2010). The independent variable of the study was the staggered introduction of a 3-session RNT-focused ACT protocol. Dependent variables were measures of psychological inflexibility, RNT, generalized pliance, emotional symptoms, diagnostic categories met according to the MINI KID, and internalizing/externalizing symptoms according to the parent with closer contact with the child. Measures of psychological inflexibility, RNT, and generalized pliance were applied on a weekly basis, whereas the remaining measures were administered at pretreatment, posttreatment, and 4-week follow-up to avoid participants' burden.

Instruments

Avoidance and Fusion Questionnaire – Youth (AFQ-Y; Greco, Lambert, & Baer, 2008; Spanish version by Salazar et al., 2019). The AFQ-Y consists of 17 items, which are rated on a 5-point Likert-type scale (4 = *very true*; 0 = *not at all true*) and measures psychological inflexibility (e.g., "The bad things I think about myself must be true," "I push away thoughts and feelings that I don't like"). The AFQ-Y was originally developed and validated in USA (Greco et al., 2008). The original study found an alpha of .90 and a one-factor structure. The AFQ-Y has shown a one-factor structure and excellent psychometric properties (alpha of .89) in Colombia (Salazar et al., 2019). The mean score of the AFQ-Y in a large Colombian nonclinical sample was 25.70 (*SD* = 15.19).

Perseverative Thinking Questionnaire (PTQ-C; Bijttebier, Raes, Vasey, Bastin, & Ehring, 2015; Spanish version by Ruiz, Salazar, et al., in press). The PTQ-C consists of 15 items with a 5-point Likert-type scale (4 = *almost always*, 0 = *never*) that measure RNT in children and adolescents (e.g., "The same thoughts keep going through my mind again and again"). The PTQ-C showed excellent psychometric properties (alpha of .93) and a one-factor structure in Colombia (Ruiz, Salazar, et al., in press). The mean score of the PTQ-C in a large Colombian nonclinical sample was 23.16 (*SD* = 15.21).

Generalized Pliance Questionnaire – Children (GPQ-C; Salazar, Ruiz, Flórez, & Suárez-Falcón, 2018). The GPQ-C consists of 8 items that are responded on a 5-point Likert-type scale (5 = *always true*, 1 = *never true*). The questionnaire is the result of reducing the original GPQ for adults (Ruiz, Suárez-Falcón, Barbero-Rubio, & Flórez, 2019) by removing items with typical adult content and changing the wording of some items from the original version to facilitate children's understanding. The GPQ-C showed good internal consistency (alpha of .81) in Colombian children and a one-factor structure

(Salazar et al., 2018). The mean score of the GPQ-C in a large Colombian nonclinical sample was 20.30 (*SD* = 7.83).

Depression, Anxiety, and Stress Scale – Children (DASS-C; Szabó, submitted) The DASS-C is an adaptation of the DASS-21 (Lovibond & Lovibond, 1995) for children. It is a 24-item, 4-point Likert-type scale (3 = *applies most of the time*, 0 = *does not apply*) consisting of sentences describing negative emotional states (e.g., "I felt tense and uptight"). It contains three subscales (Depression, Anxiety, and Stress) and has shown good internal consistency and convergent and discriminant validity. The method described in Muñiz, Elosua, and Hambleton (2013) was used to translate the DASS-C into Spanish. Alpha values in a previous study with a large Colombian nonclinical sample were acceptable (.78, .79, and .69; Salazar et al., 2018), with a mean score for the overall scale of 19.40 (*SD* = 12.92), for Depression 5.18 (*SD* = 5.02), for Anxiety 5.82 (*SD* = 5.37), and for Stress 8.91 (*SD* = 4.89).

Mental International Neuropsychiatric Interview for Kids and Adolescents (MINI KID; Sheehan et al., 2009; Spanish adaptation by Colón-Soto, Díaz, Soto, & Santana, 2005). The MINI KID is a brief diagnostic interview that explores the main psychiatric disorders of Axis I of the DSM-IV-TR and the CIE-10. The administration of the MINI KID takes approximately 15 minutes and consists of different modules identified by letters belonging to a specific diagnostic category. The questions in each module have a YES or NO answer. At the beginning of each module, there are filter questions that allow advancing more quickly in the interview by ruling out the presence of specific disorders.

Child Behavior Checklist for Ages 6-18 years old (CBCL/6-18; Achenbach & Rescorla, 2001) The CBCL is a questionnaire used to assess behavioral issues in children ages 6-18 years old as reported by the parents. It consists of 113 items that are responded

on a 3-point Likert-type scale (2 = *very true or often true*, 0 = *not true*). The instrument has shown excellent internal consistency and validity. The CBCL assesses a wide range of behavior domains including anxiety/depression, withdrawal/depression, somatic concerns, social problems, thought problems, attention problems, rule-breaking behavior, and aggressive behavior. The CBCL provides an overall score, and internalizing, externalizing and mixed problem scores.

Kaufman Brief Intelligence Test - 2 (KBIT-2; Kaufman & Kaufman, 1990). The KBIT-2 is a brief (approximately 20 minutes) intelligence test for individuals from 4 to 90 years old. It was designed for traditional brief assessment purposes, such as screening, conducting periodic cognitive reevaluations, and assessing cognitive functioning when it is a secondary consideration. It assesses both verbal and nonverbal intelligence. Only the verbal scale was administered, which has two types of items that evaluate crystallized ability: verbal knowledge and riddles.

RNT-focused ACT Protocol

The protocol consisted of three, individual, 40-minute sessions. It was based on the relational frame theory's (RFT; Hayes, Barnes-Holmes, & Roche, 2001) definition of psychological flexibility and the formation of the self (Luciano, 2017; Luciano, Valdivia-Salas, & Ruiz, 2012; Ruiz & Perete, 2015; Törneke, Luciano, Barnes-Holmes, & Bond, 2016) and on previous similar protocols used in Ruiz et al. (2016, 2018). The aim of the protocol was to develop the ability to discriminate ongoing triggers for worry/rumination, take distance from them (i.e., defusion), and behave according to what is most important at that moment for the individual in the long term (i.e., values).

Table 2 presents the content of the three protocol sessions (a complete description of the protocol can be found at https://osf.io/xavhw/). The aims of Session 1 were: (a) to

establish the differentiation between psychological inflexibility (PI) and psychological flexibility (PF) reactions through multiple examples, (b) to practice the differentiation between PI and PF, (c) to examine options for PI and PF in the child's daily life, and (d) to establish the child's commitment to realize whether she was reacting in an inflexible or flexible way toward her ongoing private experiences until the next session. The objectives of Session 2 were: (a) to review the experience since the last session and advances in discrimination of PI and PF, (b) to identify the counterproductive effects of RNT and practice defusing from its triggers, and (c) to establish the commitment to continue practicing the differentiation between PI and PF, and to try not to engage in counterproductive RNT. Lastly, the aims of Session 3 were: (a) to review examples of inflexible and flexible reactions since the last session, (b) to develop defusion skills through multiple exemplar training, and (c) to identify valued actions and barriers and to establish the committed actions for the next weeks.

INSERT TABLE 2 ABOUT HERE

Procedure

The study was conducted in the Clinical Psychology laboratory of a Colombian university. The procedures of the study were approved by the Internal Ethics Committee. The parents who showed interest in the research were invited to an assessment and informative session with their children, led by the first and/or third authors. In this session, the parents responded to the CBCL, and the K-BIT, MINI KID, AFQ-Y, PTQ-C, GPQ-C, and DASS-C were administered to the children.

Parents' of children who did not meet the inclusion criteria were given options for inexpensive psychological treatment. If the children were eligible, the study functioning was presented to the parents and the child, and both signed the informed consents. All

eligible individuals agreed to participate in the study. Afterward, participants and experimenters agreed on how the children would respond to the AFQ-Y, PTQ-C, and GPQ-C on a weekly basis. During the following weeks (4 to 6 weeks depending on the cohort randomly assigned to the participant), participants provided the baseline data.

After collecting the baseline data, the protocol was implemented in an individual format exclusively with the children. The ACT protocol was implemented by the first author in all cases. She was a doctoral student who had received about 60 h of formal training in ACT during the last two years (approximately 30 hours in the general ACT model and 30 h of training in RNT-focused ACT protocols). The second author, who is an experienced ACT researcher and has acted as a therapist in several clinical studies, trained and supervised the therapist. Once the intervention had finished, the participants provided data for posttreatment and follow-up on a weekly basis. Blinding procedures were not implemented because the study only involved one intervention.

Data Analysis

In this section, we present the statistical analyses conducted to: (a) explore trends in baseline for the AFQ-Y, PTQ-C, and GPQ-C, (b) the procedure followed to select the statistical analyses after conducting a visual analysis, (c) the Bayesian approach followed to analyze the evidence for a treatment effect, (d) the design-comparable standardized mean difference computed, and (e) the Bayesian repeated-measures ANOVA to analyze the effect of the intervention on the DASS-C and CBCL.

Analysis of trend in baseline. To assess the presence of significant trends in the baseline, the Theil-Sen slope (Sen, 1968; Vannest, Parker, Davis, Soares, & Smith, 2012) was computed before introducing the intervention. The Theil-Sen slope is a nonparametric linear regression slope that does not assume any particular data distribution. It has stronger

power/precision than the Koenig and Tukey nonparametric slopes. The Theil-Sen slope approximates the efficiency of linear regression when data meet all parametric assumptions and it significantly exceeds efficiency when data are very nonnormal and skewed (Vannest et al., 2012). Accordingly, although it is not very frequent in psychology studies, the Theil-Sen slope is the method of choice in medicine and physical sciences for making decisions with time-series data. The Theil-Sen slope was computed using the online calculator provided by Vannest, Parker, and Gonen (2011).

Graphical analysis and selection of statistical analyses. Following a bottom-up analysis of single-case experimental designs (SCED; Parker & Vannest, 2012), the results were first graphed and, subsequently, statistical analyses for SCED were selected and computed. In general, the data showed baselines with no significant trends. At the follow-up, participants' scores usually reached stability at the last three follow-up observations (2-week to 4-week follow-ups). These observations are the most relevant ones in terms of the clinical significance of the findings. Accordingly, we decided to focus the statistical analysis of each participant on all baseline data and the last three follow-up points (see a similar rationale in Au et al., 2017; Parker & Vannest, 2012; Ruiz et al., 2018). This decision has the advantage of avoiding modeling linear improvement trends in the treatment phase, which could lead to lines exceeding the range scale of the questionnaires used.

Bayesian analysis of significant change for SCED. According to the previous decision, we selected the Jeffreys-Zellner-Siow + Auto-Regressive Bayesian hypothesis testing for single-subject designs (JZS+AR model; de Vries & Morey, 2013, 2015). The JZS+AR Bayesian model is an adaptation of the JZS *t*-test (Rouder, Speckman, Sun, Morey, & Iverson, 2009) that accounts for the serial dependence typical of single-subject

designs with an autoregressive (AR(1)) model. It provides a Bayes factor (B_{ar}), which quantifies the relative evidence in the data for the hypothesis of intervention effect (i.e., the true means of both phases differ: $B_{ar} > 1$) and for the hypothesis of no intervention effect (i.e., the true mean in the baseline equals the true mean in the intervention phase: $B_{ar} < 1$). The Bayes factor can be also seen as the extent to which a rational person should adjust his beliefs, expressed as relative odds, in favor of the hypothesis of intervention effect according to the data (de Vries & Morey, 2013). Bayes factors were interpreted according to the guidelines provided by Jeffreys (1961) and Wagenmakers, Wetzels, Borsboom, and van der Maas (2011): 1 = No evidence of treatment effect; 1-3 = Anecdotal evidence of treatment effect; 3-10 = Substantial or moderate evidence of treatment effect; 10-30 = Strong evidence of treatment effect; 30-100 = Very strong evidence of treatment effect; and >100 = Extreme evidence of treatment effect (note that $B_{ar} < 1$ are interpreted in the same way, but favoring the hypothesis of no treatment effect).

One of the distinctive features of Bayesian statistics is that they include prior expectations of the parameters (e.g., the intervention effect). These prior expectations are expressed by prior distributions that receive high density at plausible parameter values and low density at implausible parameter values (Lee, 2004). Prior distributions can be determined based on previous research, expert knowledge, scale boundaries, and statistical considerations (de Vries & Morey, 2013).

To propose prior distributions, the JZS+AR model uses an estimation of two relevant parameters: (a) an effect size of the intervention effect, termed δ, consisting of standardizing the difference in true means between phases; and (b) a parameter for the lag 1 (p) autocorrelation, termed b. De Vries and Morey (2013) suggested three prior distributions for δ in which it is located at 0 and follows a Cauchy distribution that differ in

width according to a factor termed *r* (suggested *r* values of 0.5, 1.0, and 2.0). This factor is equal to half the inter-quartile range of the distribution (i.e., there is a 50% prior probability that the effect size will be found within -0.5 and 0.5, -1.0 and 1.0, and -2.0 and 2.0, respectively). The authors advocated using $r = 1$ by default because, in SCED, effect sizes tend to be larger than in group studies (e.g., Beeson & Robey, 2006; Parker & Vannest, 2009). Additionally, the authors suggested three prior distributions for the lag 1 autocorrelation ($b = 1$, $b = 5$, $b = 15$) and advocated for the use of $b = 5$. This prior distribution reflects the expectation of positive but low autocorrelations, while also considering values of .4 or .5 plausible. This is consistent with the literature in SCED showing that autocorrelation in this type of studies is reasonably low (e.g., Parker et al., 2005).

Following the guidelines of de Vries and Morey (2013) and the results obtained in similar studies (Ruiz, et al., 2016, 2018), we selected a value of $r = 1$ for the prior distribution of δ. However, we also conducted a Bayesian sensitivity analysis that investigated the robustness of the results with *r* values of 0.5 and 2.0, which posit higher density in the Cauchy distribution at, respectively, medium and very large effect sizes. Conducting sensitivity analyses is frequently suggested by Bayesian statisticians to investigate whether the results obtained are excessively dependent on the selected prior distribution (Gelman et al., 2014). Regarding the prior distribution of the autocorrelation, we followed the suggestion provided by de Vries and Morey (2013) of choosing $b = 5$. All analyses with the JZS+AR model were conducted in the BayesSingleSub R package (de Vries & Morey, 2015). Due to prior evidence showing the effect of RNT-focused ACT protocols (Ruiz et al., 2016, 2018), we conducted one-sided Bayes factor, testing the hypothesis that $\delta = 0$ against the alternative that $\delta > 0$.

Design-comparable standardized mean difference. To obtain an overall estimate of the effect size of the intervention, the design-comparable effect size for multiple-baseline designs developed by Pustejovsky, Hedges, and Shadish (2014) was computed. This standardized mean difference effect size for SCED shares the same metric as Cohen's *d*, typically used in group designs, which facilitates the direct comparison and integration through meta-analysis of the results obtained in both types of designs. This *d*-statistic has a formal mathematical development, requires at least three cases for computation, and corrects for small sample bias using Hedges' *g*. It is an extension of the standardized mean difference advocated by Hedges, Pustejovsky, and Shadish (2012, 2013) that uses restricted maximum likelihood estimation. It offers the possibility of obtaining the *d*-statistic by controlling for baseline trend and taking into account change in slope. The R package scdhlm was used to compute this *d*-statistic (Pustejovsky, 2016), following the guidelines provided by Valentine, Tanner-Smith, and Pustejovsky (2016). According to the global visual inspection of the dataset, we modeled baselines without trends including both fixed and random effects for level. The treatment phase was modeled with linear trends with both fixed and random effects at level and slope.

Bayesian repeated-measures ANOVA. To analyze the results on the DASS-C and the CBCL, a Bayesian repeated-measures ANOVA was conducted with JASP 0.9.01 (JASP Team, 2018). JASP provides a graphical interface of the R package BayesFactor, which permits the computation of Bayes factors in standard designs (e.g., *t*-tests, ANOVA, regression). The Bayesian ANOVA framework advocated by Rouder, Morey, Verhagen, Swagman, and Wagenmakers (2017) suggests Cauchy prior distributions in which the effect size of the factor, termed δ, is located at 0, and the researcher can modify the parameter *r* between the recommended values of 0.2 to 1.0. This parameter represents the

width of the distribution (higher values of *r* places more density at higher effect sizes). The authors advocated using $r = 0.5$ by default. However, we also conducted a Bayesian sensitivity analysis that investigated the robustness of the results with *r* values of 0.2 and 0.8, which posit higher density in the Cauchy distribution at, respectively, small and large effect sizes. Cohen's *d* was computed with JASP for pretreatment to posttreatment differences and for pretreatment to the 4-week follow-up.

Results

Within-participant results

The raw data of this study can be obtained at https://osf.io/7r3gn/. Figure 1 shows the scores' evolution on psychological inflexibility (AFQ-Y), RNT (PTQ-C), and generalized pliance (GPQ-C). The Theil-Sen slope revealed that P9 showed a statistically significant improving trend in the AFQ-Y, whereas P3 and P5 showed improving and deteriorating trends for the GPQ-C, respectively. Accordingly, we decided not to compute the JZS+AR analysis with these measures in these participants. The results of the Theil-Sen slope can be seen at https://osf.io/a7z6q/. Visual inspection shows that the ACT protocol was very effective in decreasing scores on psychological inflexibility and RNT in all participants, whereas the change in generalized pliance was modest.

INSERT FIGURE 1 ABOUT HERE

Table 3 shows the effect sizes and B_{ar} on the *JZS+AR* Bayesian model. All participants showed at least strong evidence (i.e., $B_{ar} > 10$) of intervention effect according to Bayes factors in the AFQ-Y and PTQ-C. Specifically, all participants showed extreme evidence in the PTQ-C, and 7 out of 8 in the AFQ-Y. Regarding the GPQ-C, only 4 out of 7 participants showed evidence of intervention effect (P1, P6, P8, and P9). Overall, the Bayesian sensitivity analysis conducted with alternative prior distributions showed that the

results were relatively robust (see the results of the sensitivity analysis at https://osf.io/7bp3f/). In other words, the Bayes factors did not vary in a way that made the interpretation of the results significantly different.

INSERT TABLE 3 ABOUT HERE

Table 1 shows that no participant showed the diagnosis of child depression according to the MINI KID at posttreatment or at the 4-week follow-up. At posttreatment, of the 6 participants who showed comorbid disorders, only P6 showed the diagnoses of attention-deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). No comorbid disorders were identified at the 4-week follow-up.

Between-participant results

Results on weekly measures. Figure 2 shows the mean results across participants in the AFQ-Y, PTQ-C, and GPQ-C. During the baseline, the mean scores on all measures were one standard deviation higher than the mean scores in nonclinical Colombian children (see Table 4). After introducing the intervention, the scores on all measures began to decrease gradually. At posttreatment, the scores on the AFQ-Y and PTQ-C were about one standard deviation below the mean scores in nonclinical participants. At the 4-week follow-up, the scores on the AFQ-Y and PTQ-C stabilized at low levels, and the scores on the GPQ-C decreased to approach the mean scores in nonclinical participants. Table 4 also shows that the *d*-statistics for SCED were very large for the AFQ-Y (d = 3.74, 95% CI [2.43, 5.43]) and PTQ-C (d = 3.14, 95% CI [1.88, 4.85]) and large for the GPQ-C (d = 1.14, 95% CI [0.01, 2.32]).

INSERT TABLE 4 ABOUT HERE

INSERT FIGURE 2 ABOUT HERE

Results on self-reported emotional symptoms. With regard to emotional symptoms, Table 5 shows that participants obtained high scores on the DASS-Total and each of its three subscales (approximately, scores one standard deviation higher than the mean scores in nonclinical participants). The Bayesian repeated-measures ANOVA revealed very strong evidence for the hypothesis of intervention effect for the DASS-Total and the subscales (DASS-Total: *BF* = 53465.77; DASS-Depression: *BF* = 172.95; DASS-Anxiety: *BF* = 67.42; DASS-Stress: *BF* = 4163000). The sensitivity analyses conducted with alternative prior distributions showed that the results were robust (DASS-Total: *BF* = 22542.2 for *r* = .20, *BF* = 77990.9 for *r* = .80; DASS-Depression: *BF* = 82.34 for *r* = .20, *BF* = 212.11 for *r* = .80; DASS-Anxiety: *BF* = 34.09 for *r* = .20, *BF* = 82.53 for *r* = .80; DASS-Stress: *BF* = 1945000 for *r* = .20, *BF* = 6328000 for *r* = .80). The effect sizes were very large both at the posttreatment and at the 4-week follow-up (DASS-Total: *d* = 2.57 and 2.12; DASS-Depression: *d* = 1.24 and 1.22; DASS-Anxiety: *d* = 1.57 and 1.18; DASS-Stress: *d* = 2.61 and 3.11). After treatment, participants showed scores below the mean of nonclinical participants in the total scores and each of the subscales.

Results on behavioral issues reported by the parents. Lastly, Table 5 also shows the scores of the parent with closer contact with the child on the CBCL. Mean scores on the CBCL were in the clinical range. The scores on all subscales decreased at the posttreatment and at the 4-week follow-up. The Bayesian repeated-measures ANOVA revealed strong evidence for the hypothesis of intervention effect for the CBCL-Total and the subscales (CBCL-Total: *BF* = 21.50; CBCL-Internalizing: *BF* = 29.62; CBCL-Externalizing: *BF* = 10.04; CBCL-Mixed: *BF* = 22.22). The sensitivity analyses showed that the results were also robust with regard to the CBCL (CBCL-Total: *BF* = 12.18 for *r* = .20, *BF* = 23.60 for *r* = .80; CBCL-Internalizing: *BF* = 16.40 for *r* = .20, *BF* = 33.86 for *r* = .80; CBCL-

Externalizing: *BF* = 6.43 for *r* = .20, *BF* = 10.22 for *r* = .80; CBCL-Mixed: *BF* = 12.43 for *r* = .20, *BF* = 24.49 for *r* = .80). The effect sizes were large both at the posttreatment and at the 4-week follow-up (CBCL-Total: *d* = 0.91 and 1.21; CBCL-Internalizing: *d* = 1.07 and 1.35; CBCL-Externalizing: *d* = 0.71 and 1.06; CBCL-Mixed: *d* = 0.93 and 1.22).

Discussion

Recent research has shown that very brief RNT-focused ACT protocols can have very large effect sizes in treating emotional disorders in adults (Ruiz et al., 2016, 2018, in press). This study adapted the previous RNT-focused ACT protocols to the work with child depression. A 3-session protocol was designed and its efficacy was analyzed with nine children suffering from child depression as the main diagnosis (six participants showed comorbid disorders). A three-arm, nonconcurrent, randomized multiple-baseline design across participants was conducted. Self-reports of psychological inflexibility (i.e., AFQ-Y), RNT (i.e., PTQ-C), and generalized pliance (i.e., GPQ-C) were administered on a weekly basis, whereas measures of emotional symptoms (DASS-C) and parent-reported problematic behavior (CBCL) were administered at pretreatment, posttreatment, and at the 4-week follow-up. Overall, participants showed scores one standard deviation higher than the mean scores in nonclinical Colombian children and baselines did not show significant improving or deteriorating tendencies.

All participants showed evidence of treatment effect in psychological inflexibility and RNT, whereas 4 out of 7 participants did so in generalized pliance. The standardized mean difference effect sizes for SCED were very large (AFQ-Y: *d* = 3.74; PTQ-C: *d* = 3.14; GPQ-C: *d* = 1.14). Importantly, these effect sizes are in the same metric as the between-group Cohen's *d*. At posttreatment, no participants showed the diagnosis of child depression according to the MINI KID (P6 continued to show the diagnoses of ADHD and

ODD). At the 4-week follow-up, none of the participants suffered from child depression or any other psychological disorder. Effect sizes were also very large for emotional disorders as reported by the children (DASS-C Total: $d = 2.12$ at the 4-week follow-up) and problematic behavior as reported by the parents (CBCL Total: $d = 1.21$ at the 4-week follow-up).

Although the results of this study are very promising and encourage the development of brief RNT-focused ACT protocols for children, some limitations are worth noting. Firstly, as opposed to concurrent multiple baseline designs, the non-concurrent multiple baseline design used in this study cannot control for history or maturation effects that might occur simultaneously with the application of the intervention (Harvey, May, & Kennedy, 2004). However, we think the weaknesses of nonconcurrent multiple baseline designs are not especially significant in this case because: (a) only the results of participants with no improvement trends in baseline are reported; (b) although there are only 4 to 6 measurement points of baseline, they represent weekly measures, which indicated that the baseline showed no improvement trend across at least one month; (c) history confounding effects seem to be less relevant when the intervention effect is replicated in 9 participants with relatively similar results across them; and (d) the interventions were implemented at different data points (after collecting 4, 5, or 6 baseline points), which reduces the possibility that the time point in which the intervention was implemented would have had a relevant effect. Additionally, the randomization of the participants to one of the three cohorts significantly increases the internal validity of the experimental design (Kratochwill & Levin, 2010). Secondly, a general limitation of usual multiple baseline designs is their lack of active control conditions that control for the nonspecific effects of therapy. Thirdly, the current study relied mostly on self-report measures, and the MINI KID was not applied

by a blind evaluator. Further studies might evaluate the intervention effect including independent clinician-administered assessments and daily measures of the children's functioning. Fourthly, we did not administer the DASS-C on a weekly basis to avoid participants' burden. In this study, we were more interested in analyzing the changes in process measures in more detail (i.e., psychological inflexibility, RNT, and generalized pliance). Future studies should include more frequent assessments of emotional symptoms. Lastly, only one therapist implemented the interventions, which reduces the external validity of the study. Subsequent studies might employ several therapists trained in the intervention.

The effect sizes obtained in this study are unusually large. For instance, the meta-analysis conducted by Yang et al. (2017) found that CBT yields weighted effect sizes of d = 0.41 (95% CI [0.18, 0.64]) for child depression. This contrasts with the effect sizes obtained in the current study in terms of emotional symptoms (d = 2.12 at the 4-week follow-up in the DASS-Total). However, the experimental design of this study cannot explain why the ACT protocol reached these unusually large effect sizes. Following Ruiz et al. (2016), this could be due to two main reasons: (a) the protocol simultaneously addressed the three strategies to promote psychological flexibility (Törneke et al., 2016) during the sessions, and (b) the protocol was focused on disrupting the first and most pervasive reaction to discomfiting thoughts and emotions (i.e., worry/rumination), which extends discomfort and supports further EA strategies.

In conclusion, this study constitutes an initial and very promising step in the analysis of brief RNT-focused ACT protocols for the treatment of child depression. Further studies might conduct randomized controlled trials to compare the effect of the ACT

protocol with waitlist control conditions or brief versions of empirically established treatments.

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Availability of Data and Materials

The raw data of this study can be downloaded at https://osf.io/7r3gn/. The results of the statistical analyses that are not reported in the manuscript for the sake of brevity can be found at https://osf.io/a7z6q/ and https://osf.io/7bp3f/. The RNT-focused ACT protocol employed in this study can be downloaded at https://osf.io/xavhw/

References

Abela, J. R. Z., Aydin, C. M., & Auerbach, R. P. (2007). Responses to depression in children: Reconceptualizing the relation among response styles. *Journal of Abnormal Child Psychology, 35*, 913-927.

Abela, J. R. Z., Brozina, K., & Haigh, E. (2002). An examination of the response styles theory of depression in third and seventh grade children: A short term longitudinal study. *Journal of Abnormal Child Psychology, 30*, 515-527.

Abela, J. R. Z., Vanderbilt, E., & Rochon, A. (2004). A test of the integration of the response styles and social support theories of depression in third and seventh grade children. *Journal of Social and Clinical Psychology, 23*, 653–674.

Achenbach, T. M., & Rescorla, L. A. (2001). *The manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Au, T. M., Sauer-Zavala, S., King, M. W., Petrocchi, N., Barlow, D. H., & Litz, B. T. (2017). Compassion-based therapy for trauma-related shame and post-traumatic stress: Initial evaluation using a multiple baseline design. *Behavior Therapy, 48*, 207-221.

Avenevoli, S., Swendsen, J., He, J. P., Burstein, M., & Merikangas, K. R. (2015). Major depression in the National Comorbidity Survey–Adolescent Supplement: Prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*, 37-44.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review, 16*, 161-169.

Bijttebier, P., Raes, F., Vasey, M. W., Bastin, M., & Ehring, T. W. (2015). Assessment of repetitive negative thinking in children: The Perseverative Thinking Questionnaire–Child Version (PTQ-C). *Journal of Psychopathology and Behavioral Assessment, 37*, 164-170.

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., . . . Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire – II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*, 676-688.

Brumariu, L. E., & Kerns, K. A. (2010). Parent–child attachment and internalizing symptoms in childhood and adolescence: A review of empirical findings and future directions. *Development and Psychopathology, 22*, 177-203.

Caselli, G., Gemelli, A., Querci, S., Lugli, A. M., Canfora, F., Annovi, C., ...Watkins, E. R. (2013). The effect of rumination on craving across the continuum of drinking behaviour. *Addictive Behaviors, 38*, 2879-2883.

Colón-Soto, M., Díaz, V., Soto, O., & Santana, C. (2005). *MINI KID Mini International Neuropsychiatric Interview*. Tampa, FL: University of South Florida.

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry, 47*, 1263-1271.

Dereix-Calonge, I., Ruiz, F. J., Sierra, M. A., Peña-Vargas, A., & Ramírez, E. S. (2019). Acceptance and commitment training focused on repetitive negative thinking for clinical psychology trainees: A randomized controlled trial. *Journal of Contextual Behavioral Science, 12*, 81-88.

de Vries, R. M., & Morey, R. D. (2013). Bayesian hypothesis testing for single-subject designs. *Psychological Methods, 18*, 165-185.

de Vries, R. M., & Morey, R. D. (2015). A tutorial on computing Bayes factors for single-subject designs. *Behavior Therapy, 46*, 809-823.

DuBois, D. L., Felner, R. D., Bartels, C. L., & Silverman, M. M. (1995). Stability of self-reported depressive symptoms in a community sample of children and adolescents. *Journal of Clinical Child Psychology, 24*, 386-396.

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, 47*, 313-337.

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy, 1*, 192-205.

Esbjørn, B. H., Normann, N., Christiansen, B. M., & Reinholdt-Dunne, M. L. (2018). The efficacy of group metacognitive therapy for children (MCT-c) with generalized anxiety disorder: An open trial. *Journal of Anxiety Disorders, 53*, 16-21.

Fekkes, M., Pijpers, F. I., Fredriks, A. M., Vogels, T., & Verloove-Vanhorick, S. P. (2006). Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. *Pediatrics, 117*, 1568-1574.

Ford, T., Goodman, R., & Meltzer, H. (2003). The British Child and Adolescent Mental Health Survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 1203-1211.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). Boca Raton, FL: CRC Press.

Greco, L. A., Lambert, W., & Baer, R. A. (2008). Psychological inflexibility in childhood and adolescence: Development and evaluation of the Avoidance and Fusion Questionnaire for Youth. *Psychological Assessment, 20*, 93-102.

Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems. *Journal of Behavioral Education, 13*, 267-276.

Hayes, L., Boyd, C. P., & Sewell, J. (2011). Acceptance and commitment therapy for the treatment of adolescent depression: A pilot study in a psychiatric outpatient setting. *Mindfulness, 2*, 86-94.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory. A post-Skinnerian account of human language and cognition*. New York: Kluwer Academic Press.

Hayes, S. C., & Strosahl, K. D. (Eds.). (2004). *A practical guide to acceptance and commitment therapy*. New York: Springer Science & Business Media.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy. An experiential approach to behavior change*. New York: Guilford Press.

Hayes, S. C., & Wilson, K. G. (1994). Acceptance and commitment therapy: Altering the verbal support for experiential avoidance. *The Behavior Analyst, 17*, 289-303.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. D. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology, 64*, 1152.

Hedges, L. G., Pustejovsky, J. E., & Shadish, W. R. (2012). A standardized mean difference effect size for single-case designs. *Research Synthesis Methods, 3*, 224-239.

Hedges, L. G., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods, 4*, 324-341.

Henker, B., Whalen, C. K., & O'Neil, R. (1995). Worldly and workaday worries: Contemporary concerns of children and young adolescents. *Journal of Abnormal Child Psychology, 23*, 685-702.

JASP Team (2018). JASP (version 0.9.0.1) [computer software]. Retrieved from https://jasp-stats.org/.

Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.

Kahn, J. S., Kehle, T. J., Jenson, W. R., & Clark, E. (1990). Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. *School Psychology Review, 19*, 196-211.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test (KBIT-2)* (2nd ed.). Circle Pines, MN: AGS, American Guidance Service.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., & Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of American Medical Association, 289*, 3095-3105.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health, 34*, 119-138.

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*, 124-144.

Lee, P. M. (2004). *Bayesian statistics: An introduction* (3rd ed.). New York: Wiley.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*, 335-343.

Luciano, C. (2017). The self and responding to the own's behavior. Implications of coherence and hierarchical framing. *International Journal of Psychology and Psychological Therapy, 17*, 267-275.

Luciano, C., & Hayes, S. C. (2001). Trastorno de evitación experiencial [Experiential avoidance disorder]. *International Journal of Clinical and Health Psychology, 1*, 109-157.

Luciano, C., Valdivia-Salas, S., & Ruiz, F. J. (2012). The self as the context for rule-governed behavior. In L. McHugh & I. Stewart (Eds.), *The self and perspective taking: Research and applications* (pp. 143-160). Oakland, CA: Context Press.

Maughan, B., Collishaw, S., & Stringaris, A. (2013). Depression in childhood and adolescence. *Journal of the Canadian Academy of Child and Adolescent Psychiatry, 22*, 35-40.

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., ...Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement

(NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry, 49,* 980-989.

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). International Test Commission Guidelines for test translation and adaptation. *Psicothema, 25,* 151-157.

Newman, M. G., & Llera, S. J. (2011). A novel theory of experiential avoidance in generalized anxiety disorder: A review and synthesis of research supporting a contrast avoidance model of worry. *Clinical Psychology Review, 31,* 371-382.

Nolen-Hoeksema, S., Stice, E., Wade, E., & Bohon, C. (2007). Reciprocal relations between rumination and bulimic, substance abuse, and depressive symptoms in female adolescents. *Journal of Abnormal Psychology, 116,* 198-207.

Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., Garcia de Alba, R., Baugh, F. G., & Sullivan, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34,* 116-132.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40,* 357–367.

Parker, R. I., & Vannest, K. J. (2012). Bottom-up analysis of single-case research designs. *Journal of Behavioral Education, 21,* 254-265.

Păsărelu, C. R., Dobrean, A., Balazsi, R., Predescu, E., Şipos, R., & Lupu, V. (2016). The Penn State Worry Questionnaire for Children: Age, gender and clinical invariance. *Child Psychiatry & Human Development, 48,* 359-369.

Perren, S., & Alsaker, F. D. (2009). Depressive symptoms from kindergarten to early school age: Longitudinal associations with social skills deficits and peer victimization. *Child and Adolescent Psychiatry and Mental Health, 3,* 28-38.

Petts, R. A., Duenas, J. A., & Gaynor, S. T. (2017). Acceptance and commitment therapy for adolescent depression: Application with a diverse and predominantly socioeconomically disadvantaged sample. *Journal of Contextual Behavioral Science, 6,* 134-144.

Pustejovsky, J. E. (2016). scdhlm: A web-based calculator for between-case standardized mean differences (Version 0.3.1) [Web application]. Retrieved from: https://jepusto.shinyapps.io/scdhlm

Pustejovsky, J. E., Hedges, L. V., & Shadish, W. R. (2014). Design-comparable effect sizes in multiple baseline designs: A general modeling framework. *Journal of Educational and Behavioral Statistics, 39*, 368-393.

Rao, U., Weissman, M. M., Martin, J. A., & Hammond, R. W. (1993). Childhood depression and risk of suicide: A preliminary report of a longitudinal study. *Journal of the American Academy of Child & Adolescent Psychiatry, 32*, 21-27.

Rouder, J. N., Morey, R. D., Verhagen, A. J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods, 22*, 304–321.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225-237.

Ruiz, F. J., Flórez, C. L., García-Martín, M. B., Monroy-Cifuentes, A., Barreto-Montero, K., García-Beltrán, D. M., ...Gil-Luciano, B. (2018). A multiple-baseline evaluation of a brief acceptance and commitment therapy protocol focused on repetitive negative thinking for moderate emotional disorders. *Journal of Contextual Behavioral Science, 9*, 1-14.

Ruiz, F. J., García-Beltrán, D. M., Monroy-Cifuentes, A., & Suárez-Falcón, J. C. (in press). Single-case experimental design evaluation of RNT-focused acceptance and commitment therapy in GAD with couple-related worry. *International Journal of Psychology and Psychological Therapy*.

Ruiz, F. J., Luciano, C., Flórez, C. L., & Suárez-Falcón, J. C. (submitted). Effect of a brief RNT-focused acceptance and commitment therapy for comorbid GAD and depression: A multiple-baseline evaluation.

Ruiz, F. J., & Perete, L. (2015). Application of a relational frame theory account of psychological flexibility in young children. *Psicothema, 27*, 114-119.

Ruiz, F. J., Riaño-Hernández, D., Suárez-Falcón, J. C., & Luciano, C. (2016). Effect of a one-session ACT protocol in disrupting repetitive negative thinking: A randomized multiple-baseline design. *International Journal of Psychology and Psychological Therapy, 16*, 213-233.

Ruiz, F. J., Salazar, D. M., Suárez-Falcón, J. C., Peña-Vargas, A., Ehring, T., Barreto-Zambrano, M. L., & Gómez-Barreto, M. P. (in press). Psychometric properties and measurement invariance across gender and age-group of the Perseverative Thinking Questionnaire–Children (PTQ-C) in Colombia. *Assessment*.

Ruiz, F. J., Suárez-Falcón, J. C., Barbero-Rubio, A., & Flórez, C. L. (2019). Development and initial validation of the Generalized Pliance Questionnaire. *Journal of Contextual Behavioral Science, 12*, 189-198.

Salazar, D. M., Ruiz, F. J., Flórez, C. L., & Suárez-Falcón, J. C. (2018). Psychometric properties of the Generalized Pliance Questionnaire – Children. *International Journal of Psychology and Psychological Therapy, 18*, 271-284.

Salazar, D. M., Ruiz, F. J., Suárez-Falcón, J. C., Barreto-Zambrano, M. L., Gómez-Barreto, M. P., & Flórez, C. L. (2019). Psychometric properties of the Avoidance and Fusion Questionnaire–Youth in Colombia. *Journal of Contextual Behavioral Science, 12*, 305-313.

Sander, J. B., & McCarty, C. A. (2005). Youth depression in the family context: Familial risk factors and models of treatment. *Clinical Child and Family Psychology Review, 8*, 203-219.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*, 1379-1389.

Sheehan, D., Shytle, D., Milo, K., Janavs, J., & Lecrubier, Y. (2009). *MINI International Neuropsychiatric Interview for Children and Adolescents, English Version 6.0*. Tampa, FL: University of South Florida.

Silverman, W. K., LaGreca, A. M., & Wasserstein, S. (1995). What do children worry about? Worries and their relation to anxiety. *Child Development, 66*, 671-686.

Swain, J., Hancock, K., Dixon, A., & Bowman, J. (2015). Acceptance and commitment therapy for children: A systematic review of intervention studies. *Journal of Contextual Behavioral Science, 4*, 73-85.

Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., ...Sampson, M. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology*, *4*, 10-31.

Thapar, A., Collishaw, S., Pine, D. S., & Thapar, A. K. (2012). Depression in adolescence. *The Lancet, 379*, 1056-1067.

Törneke, N., Luciano, C., Barnes-Holmes, Y., & Bond, F. W. (2016). Relational frame theory and three core strategies in understanding and treating human suffering. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley handbook of contextual behavioral science* (pp. 254-272). Oxford, UK: Wiley-Blackwell.

Urbaniak, G. C., & Plous, S. (2013). Research Randomizer, Version 4.0 (Web-based application. Available at www.randomizer.org

Valentine, J. C., Tanner-Smith, E. E., & Pustejovsky, J. E. (2016). *Between-case standardized mean difference effect sizes for single-case designs: A primer and tutorial using the scdhlm web application*. Oslo, Norway: The Campbell Collaboration.

Vannest, K. J., Parker, R. I., Davis, J. L., Soares, D. A., & Smith, S. L. (2012). The Theil-Sen slope for high stakes decision from progress monitoring. *Behavioral Disorders, 37*, 271-280.

Vannest, K. J., Parker, R. I., & Gonen, O. (2011). Single-case research: Web-based calculator for SCR analysis, Version 1.0 (Web-based application). College Station: Texas A&M University. Available from www.singlecaseresearch.org

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of Psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 3*, 426-432.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*, 163-206.

Watkins, E. R. (2016). *Rumination-focused cognitive-behavioral therapy for depression*. New York: Guilford Press.

Weersing, V. R., Jeffreys, M., Do, M. C. T., Schwartz, K. T., & Bolano, C. (2017). Evidence base update of psychosocial treatments for child and adolescent depression. *Journal of Clinical Child & Adolescent Psychology, 46*, 11-43.

Wells, A. (2009). *Metacognitive therapy for anxiety and depression*. New York: Guilford Press.

Yang, L., Zhou, X., Zhou, C., Zhang, Y., Pu, J., Liu, L., ...Xie, P. (2017). Efficacy and acceptability of cognitive behavioral therapy for depression in children: A systematic review and meta-analysis. *Academic Pediatrics, 17*, 9-16.

Zhou, X., Hetrick, S. E., Cuijpers, P., Qin, B., Barth, J., Whittington, C. J., ...Zhang, Y. (2015). Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. *World Psychiatry, 14*, 207-222.

Table 1

Demographical Data, K-BIT Scores, and Diagnoses at Baseline and the 4-Week Follow-Up

	Sex	Age	Grade	K-BIT Verbal IQ	Diagnoses baseline	Diagnoses posttreatment	Diagnoses 4-week follow-up
P1	F	10	5th	109	Depression, GAD	None	None
P2	M	8	3rd	110	Depression, separation anxiety, ADHD (combined)	None	None
P3	M	9	3rd	101	Depression, ODD	None	None
P4	F	8	3rd	126	Depression, ADHD (combined), ODD	None	None
P5	M	13	5th	76	Depression	None	None
P6	M	11	6th	91	Depression, ADHD (attention), ODD	ADHD (attention), ODD	None
P7	M	8	3rd	125	Depression, ODD	None	None
P8	F	13	7th	88	Depression	None	None
P9	F	12	5th	81	Depression	None	None

Note. ADHD = Attention-Deficit Hyperactivity Disorder; GAD = Generalized Anxiety Disorder, K-BIT = Kaufman – Brief Intelligence Test, ODD = Oppositional Defiant Disorder.

Table 2

Summary of the ACT Protocol

Phase	Aims	Therapeutic interactions
Session 1 (40 min)	1. Differentiating between psychological inflexibility (PI) and psychological flexibility (PF) reactions.	▪ Stating that we have thoughts and feelings all day and we have to choose between: (a) being the "wise king" by doing things that make us to feel proud of ourselves because we want to be that way, we are being responsible and growing as a person; (b) being the "slave" of our thoughts and feelings because we move away from what is important for us.
	2. Practicing the differentiation between experiential avoidance (EA) and repetitive negative thinking (RNT) reactions (PI) and valued actions (VA).	▪ Introducing two examples of EA and RNT and two examples of VA, and situating them on a rug depicting the two options as opposed directions. ▪ Asking the child to classify six new examples of EA and RNT (3) and VA (3).
	3. Examining options for PI and PF in daily life.	▪ Examining how the child can be the "wise king" or the "slave" in 8 daily life situations. Emphasis on RNT as the beginning of choosing to be the "slave." ▪ Summarizing what things seem important for the child in view of the previous responses.
	4. Commitments for the next session	▪ Asking the child to realize to what direction her actions go and to try to be the "wise king."
Session 2 (40 min)	1. Review experience since the last session.	▪ Exploration of actions as the "wise king" or the "slave" since the last session.
	2. Identifying the counterproductive effect of RNT and defusing from triggers.	▪ Introducing RNT as being the "slave" of thoughts. ▪ Identify two typical situations in which the child engages in RNT. ▪ Go around exercise: while doing something symbolically valuable, the therapist shows a trigger for RNT on a card and the participant stops her actions and begins the RNT process going around a chair in circles. Every time the participant makes a loop, she says the next thought of the chain and chooses to make another loop (the same process is repeated 8 times). Then, the participant is invited to engage again in the valued action and choose just to observe the triggers for RNT and go back to the valued action. This exercise is done with the two situations identified in the previous point. ▪ Balls as triggers exercise: The child is asked to walk toward a valued direction, but the therapist throws triggers for RNT in the form of small balls. The child can choose between avoiding the contact of the balls and not advancing or letting the balls contact her and advancing toward the valued direction. ▪ Eye-contact exercise (Hayes et al., 1999): The participant and therapist look into each other's eyes for 2 min while noticing every thought and emotion and choosing to continue.
	3. Commitment for the next session.	▪ Asking what the child was learnt in the session. ▪ Asking the child to realize in what direction her actions go and to try not to engage in counterproductive RNT.
Session 3 (40 min)	1. Review experience since the last session.	▪ Exploration of actions as the "wise king" or the "slave" since the last session.

2. Developing defusion skills.	▪ Multiple-exemplar training in defusion based on Luciano et al. (2011) and Ruiz and Perete (2015): hierarchically framing the ongoing thoughts and feelings with the deictic "I" and providing regulatory functions to that discrimination. ▪ Playing where's Wally and free association exercise (based on Wells, 2009): The child is asked to look for Wally and the therapist says 8 words separated approximately by 10 s. The participant has to notice what thought comes to her mind and chooses between entangling with it and following the search for Wally. ▪ Daydreaming and worrying exercise (based on Wells, 2009): The participant is invited to daydream for 2 minutes. Each 20 s, the therapist asks the participant to notice what she was thinking and how she could choose between following or stopping the process. The same process was repeated with a worry. ▪ Writing with the nondominant hand: The participant writes for 2 min with her nondominant hand while noticing the discomfort, not entangling with it, and choosing to continue.
3. Identification of valued actions	▪ The therapist asks the participant to list some valued actions that the she could do instead of being entangled with her thoughts ("things that make her proud at the end of the day"), to identify barriers, and establish commitment to practice the exercises in order to choose to be the "wise king."

Table 3

Results in the JZS+AR Analysis for each Participant and Measure with a Prior Distribution with $r = 1$.

		P1	P2	P3	P4	P5	P6	P7	P8	P9	%
AFQ-Y – Psychological Inflexibility	δ	6.81	7.04	6.13	12.60	6.81	23.61	2.92	20.21	--	
	B_{ar}	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	**24.4**	**>100**	--	100%
PTQ-C – Repetitive Negative Thinking	δ	8.82	9.06	5.32	6.80	5.21	5.49	6.47	16.72	8.22	
	B_{ar}	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	**>100**	100%
GPQ-C – Generalized Pliance	δ	1.90	0.29	--	1.02	--	1.46	-0.77	2.42	3.36	
	B_{ar}	**9.09**	0.71	--	2.19	--	**3.76**	0.29	**15.4**	**23.8**	57.1%

Note. B_{ar} = Bayes Factors of the JZS+AR model. $B_{ar} > 1$ supports the hypothesis of intervention effect. $B_{ar} > 3$ are in bold to highlight where at least substantial evidence of treatment effect was found.

Table 4

Means and Standard Deviations in each Self-Report Measure at Baseline, Posttreatment, and 4-Week Follow-Up

	Baseline	Post	4-week F-U	*d*-statistic for SCED	
	M (SD)	*M (SD)*	*M (SD)*	*d (SE)*	*95% CI*
AFQ-Y – Psychological Inflexibility	43.87 (8.16)	12.22 (9.12)	7.33 (6.14)	3.74 (0.80)	[2.43, 5.43]
PTQ-C – Repetitive Negative Thinking	40.16 (9.71)	7.67 (7.47)	7.00 (5.32)	3.14 (0.80)	[1.88, 4.85]
GPQ-C – Generalized Pliance	29.29 (2.78)	26.11 (4.89)	22.44 (7.67)	1.14 (0.60)	[0.01, 2.32]

Note. AFQ-Y = Avoidance and Fusion Questionnaire - Youth; GPQ-C = Generalized Pliance Questionnaire – Children; PTQ-C = Perseverative Thinking Questionnaire - Children.

Table 5

Means, Standard Deviations, Bayes Factors and 95% Confidence Intervals in the Scores of the DASS-C and CBCL

	Pre	Post	4-week F-U	Pre vs. Post	Pre vs. 4-week F-U
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*d* 95% *CI*	*d* 95% *CI*
DASS-C – Total	37.78 (14.39)	10.44 (12.69)	6.56 (7.23)	2.57 [1.16, 3.96]	2.12 [0.89, 3.32]
DASS-C – Depression	12.11 (8.91)	2.89 (3.95)	1.56 (2.88)	1.24 [0.33, 2.09]	1.22 [0.32, 2.08]
DASS-C – Anxiety	8.44 (4.50)	3.00 (4.21)	2.67 (2.45)	1.57 [0.55, 2.55]	1.18 [0.30, 2.03]
DASS-C – Stress	17.22 (3.73)	4.56 (4.93)	2.33 (2.50)	2.61 [1.18, 4.01]	3.11 [1.47, 4.73]
CBCL – Total	68.67 (46.39)	36.00 (29.63)	20.78 (19.18)	0.91 [0.11, 1.70]	1.21 [0.32, 2.07]
CBCL – Internalizing	17.00 (11.12)	8.78 (8.27)	5.56 (5.57)	1.07 [0.22, 1.89]	1.35 [0.41, 2.26]
CBCL – Externalizing	20.22 (15.20)	11.00 (9.23)	5.44 (5.05)	0.71 [-0.05, 1.43]	1.06 [0.21, 1.87]
CBCL – Mixed	31.44 (21.09)	16.22 (13.10)	9.78 (9.15)	0.93 [0.12, 1.70]	1.22 [0.32, 2.08]

Note. CBCL = Child Behavior Checklist; DASS-C = Depression Anxiety and Stress Scale – Children; F-U = Follow-up

Figure 1. Scores on psychological inflexibility (AFQ-Y), repetitive negative thinking (PTQ-C), and generalized pliance (GPQ-C) for each participant.

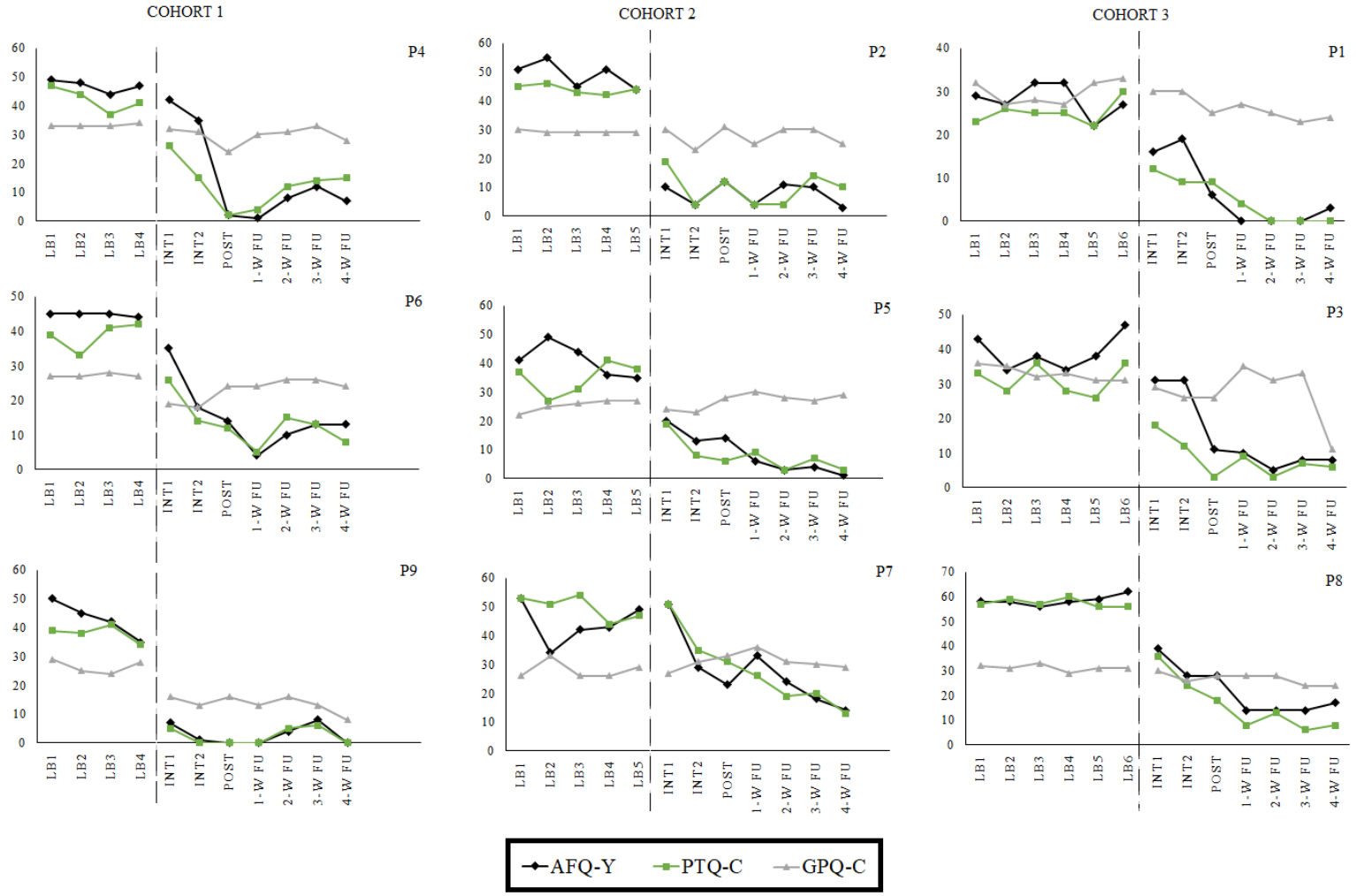

Figure 2. Mean scores' evolution on psychological inflexibility, repetitive negative thinking, and generalized pliance. Bars represent 95% credibility intervals