Behavioral Experiments for Intolerance of Uncertainty:

A Randomized Clinical Trial for Adults with Generalized Anxiety Disorder

Michel J. Dugas[1], Kathryn A. Sexton[2], Elizabeth A. Hebert[3], Stéphane Bouchard[1], Jean-Philippe Gouin[4], and Roz Shafran[5]

[1] Département de Psychoéducation et de Psychologie, Université du Québec en Outaouais;

[2] Ottawa, Ontario, Canada

[3] Department of Clinical Health Psychology, University of Manitoba

[4] Department of Psychology, Concordia University

[5] Great Ormond Street Institute of Child Health, University College London

Second revision: March 17th 2022

Declarations of interest: none

Author Note

This study was registered with ClinicalTrials.gov (Identifier NCT02552108). We have no conflicts of interest to disclose. This study was funded by Grant MOP-142311 awarded by the Canadian Institutes of Health Research to Michel J. Dugas, Stéphane Bouchard, Jean-Philippe Gouin, and Roz Shafran. We are grateful to Drs. Andrée Côté and Robert Blais for their help with the recruitment of participants.

All research at Great Ormond Street Hospital NHS Foundation Trust and UCL Great Ormond Street Institute of Child Health is made possible by the NIHR Great Ormond Street Hospital Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Correspondence to Michel J. Dugas, Département de Psychoéducation et de Psychologie, Université du Québec en Outaouais, 283 Alexandre-Taché Blvd, P. O. Box 1250, Station Hull, Gatineau, Québec, Canada, J8X 3X7. Email: michel.dugas@uqo.ca

Abstract

Sophisticated multicomponent treatments for adults with generalized anxiety disorder (GAD) have been developed over the past three decades. Although these comprehensive treatments have produced encouraging results, they appear to be less efficacious than treatments for other anxiety disorders. The goal of this randomized controlled trial is to test a newly developed, highly focused treatment for adults with GAD: *Behavioral Experiments for Intolerance of Uncertainty*. Sixty (60) participants (51 women, 9 men), with a mean age of 34.60 years (range: 19 to 67 years), were randomized to either treatment (n = 30) or wait-list control (n = 30). Treatment consisted of 12 weekly one-hour sessions in which participants learned to use behavioral experiments to test their catastrophic beliefs about uncertainty. Assessments were conducted at pre-, mid- and postcondition, and at 6- and 12-month follow-up. The primary outcome was the severity of GAD, and secondary outcomes were worry, depression, somatic anxiety, and intolerance of uncertainty. Using growth curve modeling, we found that (1) the treatment group was superior to the wait-list group in terms of change from pre- to posttest on all outcomes; (2) the combined sample (once wait-listed participants received treatment) evidenced large and significant decreases on all outcomes; and (3) treatment gains were either maintained or increased over the 12-month follow-up period of the study. The new treatment is a promising treatment option for adults with GAD considering that it may be as efficacious as more comprehensive evidence-based psychological treatments for GAD.

Keywords: Generalized anxiety disorder; cognitive-behavioral treatment; intolerance of uncertainty; behavioral experiments.

Behavioral Experiments for Intolerance of Uncertainty: A Randomized Clinical Trial for Adults with Generalized Anxiety Disorder

Generalized anxiety disorder (GAD) is a chronic and debilitating condition that is characterized by excessive and uncontrollable worry and anxiety about various life domains (American Psychiatric Association [APA], 2013). There exist a number of empirically supported, cognitive-behavioral treatments (CBT) for adults with GAD. Examples include metacognitive therapy (Wells & King, 2006), acceptance-based behavior therapy (Roemer & Orsillo, 2007), and CBT with a focus on intolerance of uncertainty (Ladouceur et al., 2000). These treatments have many things in common: (a) they are based on models of GAD that involve multiple affective, cognitive, and behavioral components; (b) they include several intervention strategies that target the model components; and (c) although they have received considerable empirical support, their efficacy appears to lag behind that of CBT protocols for other anxiety disorders (Hunot et al., 2010). Thus, despite being relatively sophisticated, CBT protocols for GAD have not produced optimal outcomes.

CBT with a focus on intolerance of uncertainty (CBT-IU) provides a good example of an established multicomponent treatment for GAD. The treatment includes four main modules: the reevaluation of the usefulness of worry, behavioral exposure to uncertainty, problem-solving training, and imaginal exposure (Robichaud et al., 2019). CBT-IU has been tested in five randomized clinical trials, with results showing that it is more efficacious than wait-list control (e.g., Ladouceur et al., 2000), supportive therapy (Gosselin et al., 2006), and to a lesser degree, applied relaxation (Dugas et al., 2010). However, a modified version of CBT-IU was found to be less effective than metacognitive therapy in one study directly comparing both treatments (van der Heiden et al., 2012). Although the results of the trials of CBT-IU are encouraging overall,

they also show that many clients do not fully benefit from the treatment; indeed, 45% to 50% do not achieve high endstate functioning at posttreatment. These numbers are similar to those obtained with other CBT protocols for GAD (Cuijpers et al., 2014).

One way to potentially increase the efficacy and clinical usefulness of treatment is to sharpen its focus. This can be accomplished by selecting the most important target construct and focusing exclusively on the chosen construct throughout treatment. Although each component of the model underlying CBT-IU has empirical support, the data show that intolerance of uncertainty (IU) is by far the most critical component. Over 25 years of basic and applied research has shown a consistent and robust relationship between IU and GAD. Despite the transdiagnostic nature of IU (for a review, see Carleton, 2012), the relationship between IU and GAD is not accounted for by shared variance with other anxiety disorders, mood disorders, negative affect, perfectionism or need for control (Buhr & Dugas, 2006; Norton et al., 2005). Data also suggest that IU is a *causal risk factor* for high levels of worry and GAD. Experimental studies have shown that change in IU leads to corresponding changes in GAD symptoms (e.g., Grenier & Ladouceur, 2004), and a five-year longitudinal study found that IU predicts subsequent levels of worry during adolescence (Dugas et al., 2012). Thus, data from experimental and longitudinal studies suggest that intolerance of uncertainty plays a key role in the etiology of GAD.

Safety Behaviors and GAD

Although some authors have argued that the judicious use of safety behaviors can be helpful in the early stages of therapy (e.g., Rachman et al., 2008), the weight of the evidence shows that safety behaviors contribute to the development and maintenance of anxiety disorders (e.g., Deacon & Maack, 2008; Olatunji et al., 2011). The literature on safety behaviors in GAD has lagged behind that of other anxiety disorders, possibly because the diagnostic criteria of

GAD do not include behavioral symptoms (see APA, 2013). However, recent theorizing is converging towards the notion that individuals with GAD frequently use safety behaviors (e.g., reassurance seeking) to increase their feelings of certainty (Beesdo-Baum et al., 2012). The consequence of using such safety behaviors is that rather than learning to cope with uncertainty, individuals with GAD attempt to decrease or avoid uncertainty in their daily lives. Ultimately, the use of certainty-seeking safety behaviors interferes with new learning about one's ability to cope with uncertainty, which ultimately maintains intolerance of uncertainty and GAD. It should be noted that worry itself can function as a safety behavior because thinking about many potential negative outcomes can be a way to decrease feelings of uncertainty about the future. However, because worry is largely a spontaneous and intrusive mental activity (as opposed to a deliberate safety behavior such as overpreparation), it is it is less amenable to direct change during treatment. Stated differently, even if one considers worry to be a form of safety behavior, its spontaneous nature precludes it from being the direct target of a treatment aiming to increase tolerance of uncertainty.

A New Treatment for GAD

Given the theoretical and empirical association between intolerance of uncertainty, safety behaviors, and GAD, we suggest that treatments for GAD should aim to decrease certainty-seeking safety behaviors. Though most established treatments do in fact address these behaviors, they may not do so in sufficient depth or in ways that explicitly promote new learning about uncertainty. However, behavioral experiments, a cognitive-behavioral technique with high evidential value, may offer important treatment advantages (see McMillan & Lee, 2010 for a review). Behavioral experiments are a personalized intervention strategy that require an individual to formulate specific predictions before entering an exposure-type situation. In the

treatment of GAD, this cognitive-behavioral technique can be used to systematically address safety behaviors and explicitly promote new learning about the contextual (What will happen?) and emotional (How will I feel?) sequelae of uncertainty. Behavioral experiments encourage expectancy violations, which may enhance inhibitory learning and retrieval (Gallistel & Gibbon, 2000) and may thereby promote new learning and superior treatment outcomes (for a review of inhibitory learning, see Craske et al., 2014). In summary, we propose that behavioral experiments are ideal to address intolerance of uncertainty in GAD because (1) they have strong evidential value; (2) they rely on experiential learning to promote change; (3) they are consistent with recent theorizing of fear reduction; and (4) when used as a vehicle for exposure, behavioral experiments promote the violation of expectations since they require clients to make specific predictions before entering exposure-type situations.

Based on the aforementioned considerations, Hebert and Dugas (2019) developed and tested a focused, single-component treatment for GAD: *Behavioral Experiments for Intolerance of Uncertainty*. The authors tested the new treatment, which relies exclusively on behavioral experiments to address safety behaviors and beliefs about uncertainty, in a case replication series with seven participants. Pre- to posttreatment effect sizes were large for all outcomes: severity of GAD symptoms, $d = 2.06$; worry, $d = 1.13$; depression, $d = 2.08$; and somatic anxiety, $d = 1.64$. Thus, it appears that focusing exclusively on intolerance of uncertainty using behavioral experiments may be a promising avenue of inquiry for the treatment of GAD.

The Current Study

The current study tests the new focused treatment in a larger sample within a randomized clinical trial. The experimental design consists of a 2 (conditions) X 5 (assessments) mixed factorial design, with repeated measures on the second factor. Sixty (60) participants were

randomly allocated to the experimental condition (Behavioral Experiments for Intolerance of Uncertainty) or the control condition (12-week Waiting List). Assessments were carried out at pretreatment, midtreatment, posttreatment, and at 6- and 12-month follow-ups.

The main goal of the study is to compare the effects of the new treatment to a waiting list on the symptoms of GAD, associated psychopathology (worry, depression, somatic anxiety) and cognitive vulnerability (intolerance of uncertainty). Hypothesis 1 is that relative to participants in the wait-list condition, those in the treatment condition will experience greater decreases in the severity of GAD, worry, depression, somatic anxiety, and intolerance of uncertainty. Hypothesis 2 states that behavioral experiments will lead to clinically significant change in the severity of GAD, in associated psychopathology, and in cognitive vulnerability. Finally, Hypothesis 3 is that treatments gains on all dependent variables will be maintained or augmented over the 12-month follow-up.

Method

Participants

The handling of study participants was in accordance with established ethical guidelines and the study was approved by the Human Research Ethics Committee of the *Université du Québec en Outaouais* (UQO). In the following paragraphs, we report how we determined our sample size, all manipulations, and all measures in the study. This study's design was preregistered; see https://register.clinicaltrials.gov/prs/app/action/SelectProtocol?sid=S0005S21&selectaction=Edit&uid=U0002UMS&ts=2&cx=9wug08.

The sample consisted of 60 French-speaking adults (51 women) with a primary diagnosis of GAD. Mean age was 34.60 years (*SD* = 12.47), with a range of 19 to 67 years. Eight (8)

participants had completed high school, 19 had finished college, and 33 had a university degree. Fifty-four (54) participants self-identified as White/European Canadian, 3 as Black/African Canadian, 1 as Hispanic/Latinx, 1 as Asian Canadian, and 1 as Multi-ethnic. In addition, 30 participants were single, separated or divorced, 13 were married, and 17 were living in common-law relationships. Finally, 34 participants were employed, 4 were students, 15 were both employed and in school, and 7 were unemployed.

At initial assessment, the mean duration of GAD was 12.22 years (*SD* = 13.05; range 1-54) and the mean severity score for GAD was 5.48 (*SD* = 0.86; range 4-7) on the 9-point (0 to 8) *Clinician's Severity Rating* of the *Anxiety Disorders Interview Schedule for DSM-IV* (ADIS-IV; Di Nardo et al., 1994). Comorbid conditions were diagnosed in 40 participants, with 25 having one comorbid condition, 9 having two comorbid conditions, and 6 having three comorbid conditions. Secondary conditions were social anxiety disorder (*n* = 14), panic disorder (*n* = 9), agoraphobia (*n* = 9), specific phobia (*n* = 9), major depressive disorder (*n* = 7), posttraumatic stress disorder (*n* = 5), obsessive-compulsive disorder (*n* = 4), adjustment disorder (*n* = 2), dysthymic disorder (*n* = 1), and hypochondriasis (*n* = 1). Finally, 24 participants (40%) were taking anxiolytic or antidepressant medication and 17 (28.3%) had previously received CBT for an anxiety or mood disorder.

Procedure

Descriptions of the study were posted in local medical clinics and universities from November 2015 to May 2017. A total of 236 interested participants contacted the study coordinator in response to the posters. Following initial telephone screening, 92 participants were assessed by structured diagnostic interview. The final sample consisted of 60 adults with a

primary diagnosis of GAD; 30 were randomized to the immediate treatment condition and 30 were randomized to the waiting list (see Figure 1 for a description of the flow of participants).

Individuals responding to the posters were screened initially over the telephone by a doctoral student using the *Telephone Screening Interview* (verbal consent was obtained prior to beginning the interview). The interview, which was used in previous studies (e.g., Dugas et al., 2010; Ladouceur et al., 2000), takes 20 to 30 minutes to administer. The goal of the telephone interview is to screen out individuals who clearly do not meet the study's inclusion criteria. Following the phone screening, potential participants were invited to the Psychological Services Clinic of the host university (UQO) to sign the study consent form and receive a formal assessment by a second doctoral-level psychology student using the ADIS-IV. Following the assessment, a team meeting was be held to discuss the diagnostic impression from the ADIS-IV and to review inclusion/exclusion criteria. Excluded individuals were contacted by the interviewer and given appropriate referral. For individuals accepted into the study, the research coordinator applied a random allocation sequence using the Research Randomizer website (https://www.randomizer.org) and contacted participants to inform them of their condition allocation.

All participants accepted into the study (N = 60) were invited to the clinic for a final intake assessment session, in which they completed a series of precondition study questionnaires, including those used in the current study (*Penn State Worry Questionnaire, Beck Depression Inventory-II, Beck Anxiety Inventory, Intolerance of Uncertainty Scale*). The total time for the intake assessment session was approximately 90 minutes. Wait-listed participants were also assessed (ADIS-IV and self-report questionnaires) after 6 and 12 weeks on the waiting list. Participants randomized to the treatment condition were offered 12 weekly 60-minute treatment

sessions. Following the 3rd treatment session, they completed measures of common therapy factors (*Working Alliance Inventory-Short Form, Credibility and Expectancy Scale for GAD*). The remaining assessments were conducted at mid- and posttreatment, and at 6- and 12-month follow-up. These included administration of the ADIS-IV and completion of self-report questionnaires. Assessments and therapy were never conducted by the same person.

Inclusion criteria

Inclusion criteria were the following: (a) at least 18 years of age; (b) primary diagnosis of GAD; (c) no change in medication type or dose in 4 to 12 weeks before study entry (4 weeks for benzodiazepines, 12 weeks for antidepressants and hypnotics); (d) willingness to keep medication status stable while participating in the study; (e) no evidence of suicidal intent (based on clinical judgement); (f) no evidence of current substance abuse, schizophrenia or bipolar disorder; (g) no current participation in other trials; and (h) no evidence of anxiety symptoms due to a general medical condition based on clinical judgement (e.g., hypoglycemia, anemia).

Measures

Measure of Primary Outcome

The ***Anxiety Disorders Interview Schedule for DSM-IV*** (ADIS-IV; Di Nardo et al., 1994) assesses anxiety disorders, and screens for mood disorders, somatoform disorders, psychoactive substance use disorders, psychotic disorders, and medical problems. The interview provides information on the presence of disorders with severity ratings on a 9-point *Clinician's Severity Rating* scale ranging from 0 (*absent or none*) to 8 (*very severe or very severely disturbing/disabling*). Hereafter, the *Clinician's Severity Rating* from the ADIS-IV will simply be referred to as the CSR. In previous studies, we obtained reliability scores ranging from $\kappa = .66$ to $\kappa = .70$ for the presence and severity of GAD with the ADIS-IV.

Measures of Secondary Outcomes

The ***Penn State Worry Questionnaire*** (PSWQ; Meyer et al., 1990) includes 16 items that assess excessive worry. Items are rated on a 5-point scale, ranging from 1 (*not at all typical of me*) to 5 (*very typical of me*). The PSWQ has high internal consistency, α = .86 to .95, and good test-retest reliability over four weeks, *r* = .74 to .93 (Molina & Borkovec, 1994). In the current sample, the internal consistency of the PSWQ at intake was α = .83.

The ***Beck Depression Inventory-II*** (BDI-II; Beck et al., 1996) includes 21 groups of 4 items reflecting different levels of depressive symptoms (e.g., sadness, loss of interest, indecision). Respondents indicate which item within each group best describes them over the past 2 weeks, with scores ranging from 0 to 3. The BDI-II has very good internal consistency, α = .92, and excellent test-retest reliability over a 1-week period, *r* = .93 (Beck et al., 1996). The internal consistency of the BDI-II at intake was α = .84 in the present sample.

The ***Beck Anxiety Inventory*** (BAI; Beck et al., 1988) is a 21-item measure of anxiety (primarily somatic anxiety) experienced during the past week. Items are rated on a 4-point scale, ranging from 0 (*not at all*) to 3 (*severely*). The BAI has high internal consistency, α = .92, and good test-rest reliability, *r* = .81, at one week in clinical samples (Beck et al., 1988). In the current study, the internal consistency of the BAI at intake was α = .87.

The ***Intolerance of Uncertainty Scale*** (IUS; Freeston et al., 1994) is a self-report measure consisting of 27 items assessing negative beliefs about uncertainty. Items are rated on a 5-point scale from 1 (*Not at all characteristic of me*) to 5 (*Entirely characteristic of me*), with higher scores reflecting greater intolerance of uncertainty. The IUS shows excellent internal consistency, α = .91, and evidence of convergent and divergent validity (Freeston et al., 1994). In the present study, the internal consistency of the IUS at intake was α = .91.

Measures of Common Therapy Factors

The ***Working Alliance Inventory-Short Form*** (WAI-SF; Horvath & Greenberg, 1989) is a 12-item measure of the client's participatory relationship with the therapist. The WAI-SF measures agreement on therapeutic goals, agreement on the tasks of therapy, and the strength of the interpersonal bond between the client and therapist. Items are rated on a 7-point point scale ranging from 1 (*never*) to 7 (*always*), with higher scores reflecting a stronger alliance. The internal consistency of the WAI-SF was $\alpha = .79$ at intake in the present sample.

The ***Credibility and Expectancy Scale for GAD*** (CES-GAD; Ladouceur et al., 2000) is based on the CES developed by Borkovec and Nau (1972). Items are rated on a 5-point scale ranging from 1 (*extremely weak*) to 5 (*extremely strong*). The internal consistency of the CES was $\alpha = .86$ in the original validation study (Borkovec & Nau). The internal consistency of the CES-GAD was $\alpha = .83$ in the present sample.

Experimental Conditions

Behavioral Experiments for IU

The treatment consisted of 12 weekly 60-minute sessions. In Sessions 1 and 2, participants were socialized to CBT and provided with information about the symptoms of GAD. Participants began basic symptom monitoring, and learned that worry, anxiety and safety behaviors are normal reactions to uncertainty. In sessions 3 to 10, participants tested their beliefs about uncertainty through a series of behavioral experiments. *Behavioral experiments involved: (1) selecting a specific thought to be tested (e.g., "uncertainty at work makes it impossible to be productive"); (2) designing a detailed experiment to test out the thought (e.g., taking on new and previously avoided responsibilities at work); (3) making specific predictions about what will happen (contextual prediction) and how they will feel (emotional prediction) during the*

experiment; (4) monitoring the contextual and emotional outcome of the experiment; and (5) comparing the predictions to the outcomes of the experiment and reflecting on what was learned. Sessions 11 and 12 consisted of relapse prevention, which involved reflecting on the new beliefs about uncertainty and behavioral patterns acquired during treatment.

Waiting List

Participants allocated to the control condition were placed on a 12-week waiting list. During the waiting period, they were contacted by telephone every 2 weeks to briefly assess their condition and screen for suicidal ideation (which none reported). Participants on the waiting list were assessed in person 6 and 12 weeks after condition allocation (mid- and post-waiting list, respectively). Following the post-waiting list assessment, they were offered the study treatment with full assessments at mid- and posttreatment, and at 6- and 12-month follow-up.

Therapists

Four graduate students with previous experience in CBT were responsible for the treatment of all participants. Two of the therapists completed their doctoral training and the other two completed their predoctoral internship during the study. The primary author (MJD), a clinical psychologist with 30 years of experience in CBT, held 90-minute weekly group supervision meetings with all therapists throughout the study.

Data Analytic Approach

Baseline differences between the conditions on demographic and clinical variables were examined using the Statistical Package for the Social Sciences (SPSS version 25; IBM Corp.). Similarly, a series of t-tests were conducted in SPSS to compare therapists' treatment outcomes. We took a liberal approach in order to increase the chance of detecting any baseline differences or therapist effects if present and did not adjust for family-wise Type I error in these analyses.

We employed growth curve modeling to assess both short- and long-term outcomes. The multilevel modeling program *Hierarchical Linear Modeling* (HLM 7; Raudenbush et al., 2011) was used to estimate growth curves for each participant, and to model the effect of condition (Treatment or Waiting List) on the rate of change. We also used HLM to estimate growth curves over follow-up. Separate analyses were performed for each of the measured outcomes. To better model outcome while taking account of the rate of attrition in this study, all participants who began treatment were retained in the analyses and all available data were used to estimate the expected intercepts and rates of change (slopes) for each individual over time. Effect sizes for the rates of change (slope) over time and the relative difference in the final outcome scores between the two conditions are presented as pseudo R^2, the percentage of variability in the outcome explained. We used model-adjusted least square means with *SEs* (unadjusted for multiple outcomes) to present the effect estimates of (1) the slope over time for the wait-list condition and (2) the difference between the slope of the treatment condition and the slope of the wait-list condition. Of note, the effect estimate of the difference between the slopes in the wait-list and treatment conditions is conceptually similar to an Group X Time interaction in a mixed-model ANOVA. To achieve this, the conditions were coded with the wait-list condition as the reference group (Condition = 0) relative to the treatment condition (Condition = 1), and the intercept (Time = 0) was set at posttest for the analyses of short-term outcomes and at the 12-month follow-up point for the analyses of long-term outcomes.

Clinically significant change was assessed in two ways. First, a reliable change index (Jacobson & Truax, 1991) was calculated for each participant on each outcome measure from pretreatment to posttreatment, and from pretreatment to 6-month and 12-month follow-ups. Second, endstate functioning was assessed for each participant on each outcome measure.

Following guidelines set forth by Jacobson and Truax, and using existing population and clinical norms for each measure, a cutoff point was calculated to evaluate whether each participant's posttreatment or follow-up score fell within the normal or clinical population range. Endstate functioning was defined as follows: participants who met criteria on 0 or 1 outcome measure were characterized as low, on 2 or 3 measures were deemed moderate, and on 4 or 5 measures were designated as high endstate functioning.

Power Analysis

Optimal Design Software program, which is a power program designed specifically for HLM, was used to calculate power. The study was powered *a priori* in order to detect a medium effect ($f = 0.25$) based on the effect sizes reported in prior studies with anxious symptomatology. With three assessment times, specifying a power of 0.90 and an alpha level of 0.05, and an estimated drop-out rate of 16.99% (mean dropout rate of CBT for GAD; Gersh et al., 2017), we sought to recruit at least 57 participants to achieve an N of 47 entering the study's follow-up phase. The data and all measures used in the current study can be obtained from the primary author (MJD).

Results

Preliminary Analyses

Treatment adherence was assessed by two doctoral students who listened to audiotapes of all sessions for 8 participants or 14% of the sample (each student rated one randomly selected participant per therapist). The students used an intervention checklist to rate therapist interventions and noted whether every item was properly addressed. Treatment integrity was 95% for Therapist 1, 100% for Therapist 2, 99% for Therapist 3, and 99% for Therapist 4, suggesting that the study therapists closely adhered to the treatment manual.

We compared therapists in terms of the treatment outcomes attained by the participants under their care. Specifically, we compared pre-, mid- and posttreatment scores on the primary outcome (CSR) for participants treated by each therapist. For each of the four therapists, participants' mean CSR scores significantly decreased from pre- to midtreatment and from mid- to posttreatment. Further, there were no significant differences between participants' mean CSR scores at each time point across the four therapists. Given the lack of significant differences between therapists at each time point and the similar patterns of improvement over time, we did not distinguish between therapists in the multilevel analyses of treatment outcome.

Final preliminary analyses compared the groups (Treatment vs. Waiting List) on pretest variables to see if they were comparable at the point of intake. We found no differences between the groups on demographic variables (age, sex, ethnic origin, education level, employment status, and marital status), clinical variables (duration of GAD, number of comorbid conditions, medication use, and previous experience with CBT), or pretest scores on the measures of dependent variables (CSR, PSWQ, BDI-II, BAI, and IUS). The multilevel analyses of outcome measures over time were therefore conducted without controlling for any demographic, clinical or pretest variables. See Supplemental Online Material for detailed demographic and clinical characteristics of participants in each condition.

Short-Term Outcomes: Treatment vs. Waiting List

Table 1 presents descriptive statistics for the measures of primary (CSR) and secondary outcomes (PSWQ, BDI-II, BAI, IUS) at pretest, midtest and posttest in the treatment and wait-list conditions. We evaluated the hypothesis that individuals receiving treatment would show greater improvement on all measures relative to those in the wait-list condition during the 12-week course of treatment by comparing the rates of change (slopes) in each measure pretest to

posttest. For the CSR (primary outcome), though the slope decreased over time in both the treatment and wait-list conditions, there was a significant effect of condition, with the treatment group showing significantly greater decreases in the severity of GAD over time.[1]

For the PSWQ (worry), the BDI-II (depressive symptoms) and the BAI (somatic anxiety), the pattern of results was similar to that of the CSR; although the slope decreased over time in both conditions, the rate of change was significantly greater in the treatment condition than in the control condition for each measure. As for the IUS (intolerance of uncertainty), we found no change in the slope from pretest to posttest in the wait-list condition, but a significant decrease in the slope in the treatment condition. In addition, the decrease in IUS scores was significantly greater in the treatment condition than in the wait-list condition. Effect estimates of the slope over time for all measures in the wait-list condition and in the treatment condition relative to wait-list, and the pretest to posttest effect size for this relative difference, are shown in Table 2.

Short-Term Outcomes: Combined Sample

After a 12-week delay, wait-listed participants were offered the study treatment, which resulted in a combined sample of 57 participants who started treatment (3 participants did not complete the wait-list period). A total of 48 completed treatment. Forty-five (45) individuals at least partially completed the assessment at 6-month follow-up, and 36 at least partially completed the 12-month follow-up. Descriptive statistics for all measures at each measurement time for the combined treatment sample are presented in Table 3.

Over the course of the 12-week treatment, the rate of change (slope) in the primary outcome measure in the combined treatment sample differed significantly from a slope of zero, evidencing a significant drop in GAD severity from pretreatment to posttreatment (the CSR

[1] For all main analyses, the addition of a quadratic term to assess for non-linear change did not significantly contribute to the models that were tested. Therefore, non-linear results are not reported.

slope coefficient = -1.39, *SE* = 0.11, $p < .001$, pseudo $R^2 = .80$). This linear change over time accounted for 80% of the within-participant variability in CSR scores from pretreatment to posttreatment. For the four secondary outcome measures, the rate of change (or slope) differed significantly from a slope of zero over the course of the 12-week treatment: PSWQ slope, coefficient = -11.08, *SE* = 0.92, $p < .001$, pseudo $R^2 = .69$; BDI-II slope, coefficient = -6.55, *SE* = 0.70, $p < .001$, pseudo $R^2 = .50$; BAI slope, coefficient = -8.09, *SE* = 0.80, $p < .001$, pseudo $R^2 = .62$; and IUS slope, coefficient = -15.16, *SE* = 1.49, $p < .001$, pseudo $R^2 = .50$. Thus, from pretest to posttest, there were large and significant decreases on all measures in the combined sample.

Long-Term Outcomes: Combined Sample

To test the hypothesis that treatment would lead to continued progress over follow-up, we compared the slope for GAD severity (CSR) with a slope of zero (a slope of zero denotes no change over time). There were further decreases in CSR scores over the 12 months following the end of treatment (CSR slope coefficient = -0.30, *SE* = 0.11, $p < .01$, pseudo $R^2 = .38$). In other words, further decreases in symptoms were observed over the follow-up period. This linear pattern of change over time accounted for 38% of the within-participant variability in CSR scores over the follow-up period.

We also compared the slope for each secondary outcome measure over the 12-month follow-up period with a slope of zero. For treatment completers in the combined sample ($n = 48$), the linear slopes for the four secondary outcome measures did not differ significantly from a slope of zero: PSWQ slope, coefficient = -0.26, *SE* = 0.89, $p > .05$, pseudo $R^2 = .36$; BDI-II slope, coefficient = -0.08, *SE* = 0.74, $p > .05$, pseudo $R^2 = .25$; BAI slope, coefficient = 0.85, *SE* = 0.73, $p > .05$, pseudo $R^2 = .15$; and IUS slope, coefficient = -0.96, *SE* = 1.58, $p > .05$, pseudo $R^2 = .11$. These findings suggest that treatment gains were maintained on each secondary outcome.

Clinically Significant Change

Frequencies and percentages of participants meeting criteria for reliable change and endstate functioning across study measures at posttreatment and follow-up are presented in Table 4. On the measure of worry (PSWQ), 47 of 57 (82.5%) participants fell within range of the normal population at posttreatment, as did 39 of 48 (81.3%) treatment completers (or 90.7% of 43 respondents) at 6-month follow-up, and 29 of 48 (60.4%) treatment completers (or 80.9% of 35 respondents) at 12-month follow-up. On the IUS, a measure of the primary target of the intervention, 40 of 57 (70.2%) participants met criteria for reliable change at posttreatment, as did 33 of 48 (68.8%) treatment completers (or 76.7% of 43 respondents) at 6-month follow-up, and 27 of 48 (56.3%) treatment completers (or 77.1% of 35 respondents) at 12-month follow-up.

Common Therapy Factors and Medication

Following the third treatment session, participants rated the quality of the therapeutic alliance (WAI-SF) as well as treatment credibility and expectations of change (CES-GAD). Mean scores were 74.13 (*SD* = 8.13) on the WAI-SF and 25.98 (*SD* = 2.86) on the CES-GAD for the 56 participants having completed three sessions. Scores on the WAI-SF were comparable to those typically reported in previous treatment studies (for a review, see Sturgiss et al., 2019). As for the CES-GAD, scores obtained in the current study were almost identical to those reported in earlier studies of CBT-IU (Dugas et al., 2010; Ladouceur et al., 2000). Thus, the quality of the therapeutic alliance as well as the credibility of treatment and expectations of change appear to be as strong as with a multicomponent treatment for GAD.

As mentioned previously, 24 of 60 participants (40%) were taking anxiolytic or antidepressant medication at the beginning of the study. For the combined sample at posttreatment, 20 of 49 participants (40.8%) continued to use anxiolytic or antidepressant

medication. Of the 24 participants who were taking medication at intake, 3 did not complete posttreatment assessments and one discontinued their medication. Thus, the treatment had a negligible impact on medication use in this study.

Discussion

The current study provides support for the use of behavioral experiments to increase tolerance of uncertainty in adults with GAD. The results show that, compared to a waiting list, the focused treatment led to greater decreases in the severity of GAD, associated psychopathology (worry, depressive symptoms, somatic anxiety), and cognitive vulnerability (intolerance of uncertainty). The findings also reveal large and significant decreases on all outcomes for the total sample, once participants in the control condition received treatment after a 12-week waiting period. Finally, we found that treatment gains on all outcomes were either maintained or increased (for the severity of GAD) over the 12-month follow-up period of the study. Thus, it appears that a single-component treatment, *Behavioral Experiments for Intolerance of Uncertainty*, represents a promising treatment option for individuals with GAD.

Although cross-study comparisons should be made with extreme caution, they can nonetheless be informative in terms of ruling out large differences between findings. Keeping this in mind, it appears that Behavioral Experiments for Intolerance of Uncertainty (a single-component treatment) and CBT with a Focus on IU (a multicomponent treatment) may lead to similar outcomes. Overall, the results of the current study appear to be (at least) comparable to those of previous clinical trials of the multicomponent treatment in a wait-list design. One notable difference, however, may be that the single-component treatment produces greater decreases in intolerance of uncertainty. In previous trials of the multicomponent treatment (Dugas et al., 2010; Gosselin et al., 2006; Ladouceur et al., 2000; van der Heiden, 2012), within-

group effect sizes on the IUS (for the intent-to-treat sample) ranged from $d = 0.58$ to $d = 0.72$. By comparison, the within-group effect sizes on the IUS (also intent-to-treat) were $d = 1.72$ in the Hebert and Dugas (2019) study and $d = 1.49$ in the current study. It may be that by focusing exclusively on intolerance of uncertainty using behavioral experiments, the single-component treatment produces larger improvements in cognitive vulnerability for GAD. This possibility, however, awaits testing in a clinical trial directly comparing both treatments.

It should be noted that, in addition to the encouraging findings of the current study, the new treatment presents a number of important advantages compared to more comprehensive treatments. Most importantly, an added benefit of using a single treatment strategy focusing exclusively on intolerance of uncertainty is that the resulting treatment is highly parsimonious. A number of authors have called for a move away from traditional assessments of the quality of therapy and towards newer assessment models that take into consideration factors such as treatment parsimony (e.g., Cougle, 2012; Mazzucchelli et al., 2009). As argued by these authors, given equivalent outcomes between two different treatments, the more parsimonious approach should be preferred. Why? First, more parsimonious treatment options may require less clinical training and may be easier to disseminate. For full-time therapists who have limited time to devote to learning new treatment protocols, the amount of time required to learn new treatment procedures is an important consideration. Second, more parsimonious treatments may increase treatment integrity; the lower the number of treatment components, the greater the likelihood that the treatment will be implemented as intended. Given that treatment integrity may predict positive outcomes, treatments that are more likely to be administered as intended can offer important advantages. Third, clients may adhere more easily to treatments that are parsimonious. Relatedly, treatment receipt and enactment may be limited if treatments are overly inclusive.

Finally, from a theoretical point of view, the active ingredients of therapy can be more easily identified when treatments have fewer components. Conversely, multicomponent treatment protocols often require dismantling studies to tease apart active and non-active ingredients. Thus, it appears that assessments of treatment quality that do not take into consideration the issue of parsimony may not be ideal; ease of training, ease of dissemination, treatment integrity, client adherence, and identification of active ingredients are important factors to consider in addition to treatment outcome.

Relatedly, previous attempts to increase the efficacy of CBT for GAD using more comprehensive methods have been met with limited success. For example, studies that have combined CBT with pharmacotherapy have produced inconclusive results (e.g., Bond et al., 2002). Other studies, which have attempted to increase the efficacy of CBT for GAD by offering more comprehensive CBT, have also produced inconclusive results (e.g., Newman et al., 2011). The current study, which tested a more focused treatment, not only holds the promise of contributing to our understanding and ability to treat GAD, but also of leading the way to increasing the dissemination and accessibility of psychological treatments for GAD.

Having acknowledged the advantages of parsimonious treatments, one should nonetheless keep in mind that comprehensive treatment options also have important advantages. Considering that individuals with GAD may be a particularly heterogenous group, multicomponent treatments can offer therapists greater flexibility in emphasizing interventions that are either highly valued by their clients or appear to be directly related to their clients' greatest vulnerabilities. For example, because intolerance of uncertainty is associated with ineffective problem solving (Clarke et al., 2017) and cognitive avoidance (Koerner & Dugas,

2006), some clients with GAD may prefer (or particularly benefit from) problem-solving training or imaginal exposure, both of which are not germane to the new treatment.

Strengths and Limitations

One of the main strengths of the current study is that the treatment intervention (behavioral experiments) and treatment target (intolerance of uncertainty) are solidly grounded in the anxiety disorders literature. As mentioned previously, behavioral experiments, which require clients to formulate specific predictions and to compare their predictions to actual outcomes following an exposure-type situation, are directly in line with contemporary learning theories of fear reduction (e.g., Craske et al., 2014). As for intolerance of uncertainty, it bears repeating that over 25 of research have supported its central role in the etiology and treatment of GAD (for a review, see Robichaud et al., 2019). Consequently, both the specific intervention and target of the new treatment constitute an important strength of the current study.

A second strength of the study is that the treatment is based on a model of GAD that is closely tied to a generic cognitive-behavioral model of psychopathology (see e.g., Tolin, 2016). Namely, individuals who are intolerant of uncertainty tend to make catastrophic misinterpretations of uncertainty when faced with situations that are novel, unpredictable or ambiguous. These misinterpretations then lead to emotional (anxiety), cognitive (worry), and behavioral (avoidance, safety behaviors) symptoms. Simply stated, the model of GAD and the generic cognitive-behavioral model of psychopathology both emphasize the activation of latent core beliefs by precipitating events, leading to biased information processing and emotional-cognitive-behavioral symptoms. Given the numerous and heterogenous cognitive-behavioral models of GAD currently under study, the development of treatments that are closely tied to the

general theory of CBT may hold promise for moving the field forward in terms of a unified understanding of anxiety and other forms of psychopathology.

One important limitation of the current study is that it was conducted in a single site with a relatively homogenous (largely white, female) community sample, which restricts the generalizability of the findings to other settings and populations. Although the use of a single site is a tangible limitation of the study, it is lessened by the fact that we recruited participants from community medical clinic waiting lists. Further, compared to multisite studies, single-site studies offer important advantages in terms of internal validity (e.g., consistency of procedures, uniformity of care), particularly in the early stages of treatment validation. In terms of the sample, it can be argued that the disadvantages of having a homogenous sample is offset by the advantages of having a Francophone sample because the latter are noticeably under-represented in the clinical literature.

Another limitation of the study relates to the use of a wait-list design. By using such a control condition, we are not in a position to disentangle the effects of treatment specific factors (e.g., change in certainty-seeking safety behaviors, change in beliefs about uncertainty) from those of common therapy factors (e.g., time spent with therapist, therapeutic alliance, treatment motivation). Although we are aware that the use of a supportive therapy control condition would have allowed us to disentangle specific and common effects, we nonetheless opted for wait-list control to limit the costs of the trial. From a treatment development perspective, the use of a wait-list control condition at this point of treatment validation is appropriate. According to best practice guidelines for treatment development (Hayes et al., 2013), the validation of new treatment procedures should follow a cost-effective sequence starting with case replication (i.e., Hebert & Dugas, 2019), followed by graded control conditions requiring more and more

participants (e.g., waiting list, supportive therapy, competing treatment). By using a wait-list condition, we were in a position to collect controlled data on the new treatment in a cost-effective manner, which is a judicious and ethical choice at this stage of treatment development.

A third limitation of the study relates to the putative mechanism underlying the behavioral experiments. Although every effort was made to ensure that exposure to uncertainty was at the heart of each behavioral experiment, we are not in a position to rule out the possibility that other constructs related to GAD were simultaneously targeted. For example, an experiment that involves waiting 48 hours before attempting to solve a problem (exposure to the uncertainty of not knowing if the problem can be solved) may also foster greater acceptance of emotional distress (a central construct in acceptance-based behavior therapy). In future studies, it would therefore be important to test hypothesized mediators from competing models.

In conclusion, it appears that a focused, single-component treatment, *Behavioral Experiments for Intolerance of Uncertainty*, represents an interesting treatment option for adults with GAD. In addition to the encouraging findings reported above, the focused treatment has the advantage of being parsimonious and of being compatible with general cognitive-behavioral models of psychopathology. We suggest that treatment models that emphasize specific targets (e.g., intolerance of uncertainty) while recognizing common change processes (e.g., prioritizing behavioral change within a learning theory framework) may represent an interesting avenue for the further development and dissemination of CBT.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Beck, A. T., Epstein, N., Brown, G. K., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897. http://doi.org/10.1037/0022-006X.56.6.893

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory Manual* (2nd ed.). San Antonio, TX: Psychological Corporation.

Beesdo-Baum, K., Jenjahn, E., Hofler, M, Lueken, U, Becker, E.S., & Hoyer, J. (2012). Avoidance, safety behavior, and reassurance seeking in generalized anxiety disorder. *Depression and Anxiety, 29,* 948-957. http://doi.org/10.1002/da.21955

Bond, A. J., Winegrove, J., Curran, H. V., & Lader, M. H. (2002). Treatment of generalized anxiety disorder with a short course of psychological therapy, combined with buspirone or placebo. *Journal of Affective Disorders, 72,* 267-271. http://doi.org/10.1016/S0165-0327(01)00469-4

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3,* 257-260. http://doi.org/10.1016/0005-7916(72)90045-6

Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders, 20,* 222-236. http://doi.org/10.1016/j.janxdis.2004.12.004

Carleton, R. N. (2012). The intolerance of uncertainty construct in the context of anxiety disorders: Theoretical and practical perspectives. *Expert Review of Neurotherapeutics, 12*, 937-947. http://doi.org/10.1586/ern.12.82

Clarke, J. B., Ford, M., Heary, S., Rodgers, J., & Freeston, M. H. (2017). The relationship between negative problem orientation and worry: A meta-analytic review. *Psychopathology Review, a4,* 319-340. https://doi.org/10.5127/pr.034313

Cougle, J. R. (2012). What makes a quality therapy? A consideration of parsimony, ease, and efficiency. *Behavior Therapy, 43,* 468-481. http://doi.org/10.1016/j.beth.2010.12.007

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T.,Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behavior Research and Therapy, 58,* 10-23. http://doi.org/10.1016/j.brat.2014.04.006

Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clinical Psychology Review, 34,*130-40. http://doi.org/10.1016/j.cpr.2014.01.002

Deacon, B., & Maack, D. J. (2008). The effects of safety behaviors on the fear of contamination: An experimental investigation. *Behaviour Research and Therapy, 46,* 537-47. http://doi.org/10.1016/j.brat.2008.01.010

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV). Psychological Corporation: San Antonio, TX.

Dugas, M. J., Brillon, P., Savard, P., Turcotte, J., Gaudet, A., Ladouceur, R., Leblanc, R., & Gervais, N. J. (2010). A randomized clinical trial of cognitive-behavioral therapy and applied relaxation for adults with generalized anxiety disorder. *Behavior Therapy, 41,* 46-58. http://doi.org/10.1016/j.beth.2008.12.004

Dugas, M. J., Laugesen, N., & Bukowski, W. M. (2012). Intolerance of uncertainty, fear of anxiety, and adolescent worry. *Journal of Abnormal Child Psychology, 40,* 863-870. http://doi.org/10.1007/s10802-012-9611-1

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17,* 791-802. http://doi.org/10.1016/0191-8869(94)90048-5

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*, 289-344. http://doi.org/10.1037/0033-295X.107.2.289

Gersh, E., Hallford, D. J., Rice, S. M., Kazantzis, N., Gersh, H., Gersh, B., & McCarty, C. A. (2017). Systematic review and meta-analysis of dropout rates in individual psychotherapy for generalized anxiety disorder. *Journal of Anxiety Disorders, 52*, 25-33. http://dx.doi.org/10.1016/j.janxdis.2017.10.001

Gosselin, P., Ladouceur, R., Morin, M., Dugas, M. J., & Baillargeon, L. (2006). Benzodiazepine discontinuation among adults with GAD: A randomized trial of cognitive-behavioral therapy. *Journal of Consulting and Clinical Therapy, 74*, 908-919. http://doi.org/10.1037/0022-006X.74.5.908

Grenier, S., & Ladouceur, R. (2004). Manipulation de l'intolérance à l'incertitude et inquiétudes [Manipulation of intolerance of uncertainty and worries]. *Canadian Journal of Behavioural Science, 36,* 56–65. https://doi.org/10.1037/h0087216

Hayes, S. C., Long, D. M., Levin, M. E., & Follette, W. C. (2013). Treatment development: Can we find a better way? *Clinical Psychology Review, 33*, 870–882. https://doi.org/10.1016/j.cpr.2012.09.009

Hebert, E. A., & Dugas, M. J. (2019). Behavioral experiments for intolerance of uncertainty: Challenging the unknown in the treatment of generalized anxiety disorder. *Cognitive and Behavioral Practice, 26,* 421-436. https://doi.org/10.1016/j.cbpra.2018.07.007

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223-233. http://doi.org/10.1037/0022-0167.36.2.223

Hunot, V., Churchill, R., Teixeira, V., & Silva de Lima, M. (2010). Psychological therapies for generalised anxiety disorder (review). *Cochrane Database of Systematic Reviews 2007, Issue 1*, Art. No. CD001848. http://doi.org/10.1002/14651858.CD001848.pub4

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

Koerner, N., & Dugas, M. J. (2006). A cognitive model of generalized anxiety disorder: The role of intolerance of uncertainty. In G. C. L. Davey & A. Wells (Eds.), *Worry and its psychological disorders: Theory, assessment and treatment* (pp. 201-216). Chichester: John Wiley and Sons.

Ladouceur, R., Dugas, M. J., Freeston, M. H., Léger, E., Gagnon, F., & Thibodeau, N. (2000). Efficacy of a cognitive-behavioral treatment for generalized anxiety disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 68*, 957-964. http://doi.org/10.1037/0022-006X.68.6.957

Mazzucchelli, T., Kane, R., & Rees, C. (2009). Behavioral activation treatments for depression in adults: A meta-analysis and review. *Clinical Psychology: Science and Practice, 16*, 383-411. https://doi.org/10.1111/j.1468-2850.2009.01178.x

McMillan, D., & Lee, R. (2010). A systematic review of behavioral experiments vs. exposure alone in the treatment of anxiety disorders: A case of exposure while wearing the emperor's new clothes? *Clinical Psychology Review, 30,* 467-478. http://doi.org/10.1016/j.cpr.2010.01.003

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behavior Research and Therapy, 28,* 487-495. http://doi.org/10.1016/0005-7967(90)90135-6

Molina, S., & Borkovec, T. D. (1994). The Penn State Worry Questionnaire: Psychometric properties and associated characteristics. In G. C. L. Davey & F. Tallis (Eds.), Worrying: Perspectives on theory, assessment and treatment (pp. 265–283). New York: John Wiley & Sons.

Newman, M. G., Castonguay, L. G., Borkovec, T. D., Fisher, A. J., Boswell, J. F., Szkodny, L. E., & Nordberg, S. S. (2011). A randomized controlled trial of cognitive-behavioral therapy for generalized anxiety disorder with integrated techniques from emotion-focused and interpersonal therapies. *Journal of Consulting and Clinical Psychology, 79,* 171-181. http://doi.org/10.1037/a0022489

Norton, P. J., Sexton, K. A., Walker, J. R., & Norton, G. R. (2005). Hierarchical model of vulnerabilities for anxiety: Replication and extension with a clinical sample. *Cognitive Behavior Therapy, 34,* 50-63. http://doi.org/10.1080/16506070410005401

Olatunji, B. O., Etzel, E. N., Tomarken, A. J., Ciesielski, B. G., & Deacon, B. (2011). The effects of safety behaviors on health anxiety: An experimental investigation. *Behaviour Research and Therapy, 49,* 719-28. http://doi.org/10.1016/j.brat.2011.07.008

Rachman, S., Radomsky, A. S., & Shafran, R. (2008). Safety behaviour: A reconsideration. *Behaviour Research and Therapy, 46,* 163-173. http://doi.org/10.1016/j.brat.2007.11.008

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM: Linear and nonlinear modeling.* Skokie, IL: Scientific Software International.

Robichaud, M., Koerner, N., & Dugas, M. J. (2019). *Cognitive-behavioral treatment for generalized anxiety disorder: From science to practice (2nd ed.).* New York: Routledge.

Roemer, L., & Orsillo, S. M. (2007). An open trial of acceptance-based behavior therapy for generalized anxiety disorder. *Behavior Therapy, 38,* 72–85. http://doi.org/10.1016/j.beth.2006.04.004

Sturgiss, E. A., Rieger, E., Haesler, E., Ridd, M. J., Douglas, K., & Galvin, S. L. (2019). Adaption and validation of the Working Alliance Inventory for general practice: Qualitative review and cross-sectional surveys. *Family Practice, 36,* 516-522. http://doi.org/10.1093/fampra/cmy113

Tolin, D. F. (2016). *Doing CBT: A comprehensive guide to working with behaviors, thoughts, and emotions.* New York: The Guilford Press.

van der Heiden, C., Muris, P., & van der Molen, H. T. (2012). Randomized controlled trial of the effectiveness of metacognitive therapy and intolerance-of-uncertainty therapy for generalized anxiety disorder. *Behavior Research and Therapy, 50,* 100-109. http://doi.org/10.1016/j.brat.2011.12.005

Wells, A., & King, P. (2006). Metacognitive therapy for generalized anxiety disorder: An open trial. *Journal of Behavior Therapy and Experimental Psychiatry, 37,* 206-212. http://doi.org/10.1016/j.jbtep.2005.07.002

Table 1

Means and Standard Deviations on Outcome Measures in Both Conditions at Pretest, Midtest, and Posttest

Measure and condition	Pretest *n* = 60 *M (SD)*	Midpoint *n* = 53 *M (SD)*	Posttest *n* = 50 *M (SD)*
CSR			
TRT	5.48 (0.83)	4.35 (1.21)	2.46 (1.41)
WL	5.47 (0.91)	5.44 (0.75)	5.17 (0.94)
PSWQ			
TRT	67.43 (7.51)	55.77 (9.77)	44.13 (8.83)
WL	64.00 (6.48)	61.48 (7.80)	60.00 (8.33)
BDI-II			
TRT	21.70 (9.74)	13.46 (11.71)	7.96 (7.34)
WL	20.70 (7.97)	16.52 (9.60)	15.81 (8.70)
BAI			
TRT	25.23 (11.76)	14.31 (11.06)	8.30 (5.84)
WL	24.23 (9.68)	21.15 (10.90)	18.89 (9.96)
IUS			
TRT	85.07 (19.48)	68.15 (21.98)	50.48 (15.73)
WL	79.70 (15.45)	78.52 (20.48)	79.30 (20.29)

Note. CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; TRT = treatment; WL = waiting list; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

Table 2

Results of Multilevel Analysis of Outcome Measures, Including Random Slopes

Outcome	Fixed effects	Intercept (posttest)[a] coeff. (*SE*)	Pseudo R^2	Slope coeff. (*SE*)	Pseudo R^2
CSR	WL (ref grp)[b]	5.18*** (0.16)	.–	-0.16*** (0.06)	
	TRT vs. WL[c]	-2.46*** (0.32)	.63	-1.27*** (0.15)	.86
PSWQ	WL (ref grp)	59.95*** (1.63)	.–	-1.96* (0.77)	
	TRT vs. WL	-15.92*** (2.46)	.56	-9.73*** (1.46)	.69
BDI-II	WL (ref grp)	15.38*** (1.69)	.–	-7.79** (2.39)	
	TRT vs. WL	-2.41*** (0.59)	.24	-4.42*** (1.22)	.78
BAI	WL (ref grp)	18.93*** (1.93)	.–	-2.61** (0.86)	
	TRT vs. WL	-11.23*** (2.38)	.39	-5.81*** (1.49)	.39
IUS	WL (ref grp)	79.30*** (4.06)	.–	-0.07 (1.78)	
	TRT vs. WL	-28.83*** (5.27)	.47	-17.33*** (2.78)	.82

Note. coeff. = coefficient; *SE* = standard error; Pseudo R^2 = proportion of between-person (intercept) or within-person (slope) variance explained; ref grp = reference group; CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; WL = waiting list; TRT = treatment; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

[a] Pretest coded as -2, midpoint coded as -1, and posttest coded as 0, such that the intercept reflects the estimated fixed effects at posttest (Time = 0).

[b] WL condition coded as 0; this effect therefore reflects the effect of Time in the WL condition.

[c] TRT condition coded as 1; this effect therefore reflects the effect of Time in the TRT condition relative to the effect of Time in the reference group (the WL condition).

*p < .05 ** p < .01 *** p < .001

Table 3

Means and Standard Deviations on Outcome Measures for Combined Sample at all Measurement Times

Measure	Pretreatment (n = 57) *M (SD)*	Midtreatment (n = 52) *M (SD)*	Posttreatment (n = 48) *M (SD)*	6 months ($n \geq 43$[a]) *M (SD)*	12 months ($n \geq 35$[b]) *M (SD)*
CSR	5.33 (0.88)	4.35 (1.10)	2.63 (1.50)	2.16 (1.55)	2.03 (2.56)
PSWQ	63.91 (8.69)	53.44 (10.47)	43.59 (10.92)	44.63 (9.77)	43.43 (10.74)
BDI-II	18.91 (9.65)	13.44 (9.85)	8.02 (6.86)	8.26 (6.33)	7.80 (8.33)
BAI	22.22 (11.31)	14.23 (9.33)	8.48 (6.28)	10.14 (5.73)	9.94 (9.73)
IUS	82.33 (19.90)	67.92 (20.49)	51.93 (17.97)	51.52 (17.54)	50.46 (17.33)

Note. CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

[a] n = 45 completed the ADIS; n = 43 completed the self-report measures. [b] n = 36 completed the ADIS; n = 35 completed the self-report measures at 12-month follow-up.

Table 4

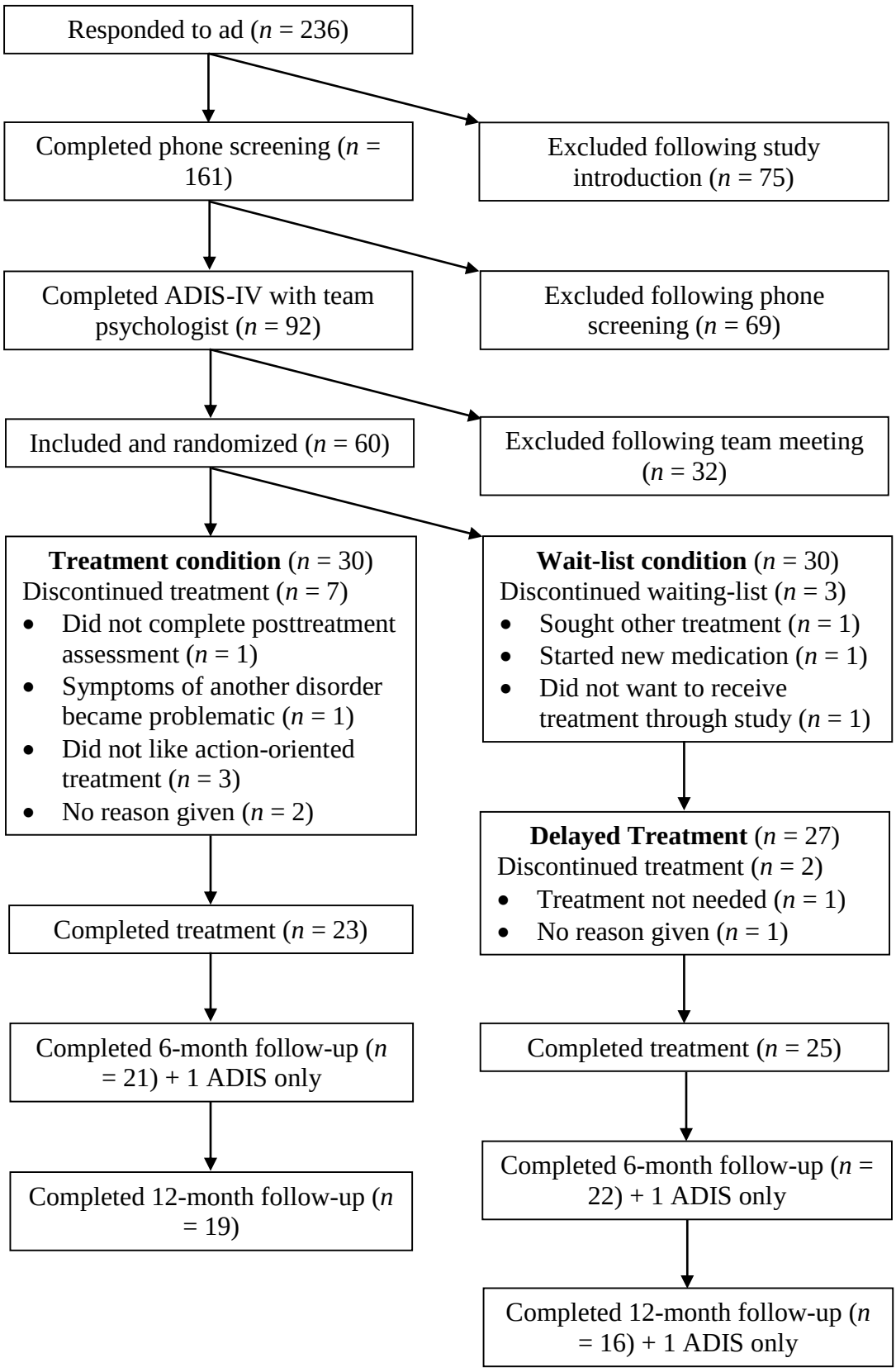
Frequency and Percentages of Participants (n = 57) Meeting Criteria for Reliable Change and Endstate Functioning at Posttreatment, 6-Month Follow-Up, and 12-Month Follow-Up

	Reliable change		Endstate functioning	
No. of measures	Frequency	%	Frequency	%
Posttreatment				
0-1	14	24.6	10	17.5
2-3	13	22.8	13	22.8
4-5	30	52.6	34	59.7
6-month follow-up (*n* = 43)				
0-1	6	14.0	3	7.0
2-3	14	32.6	14	32.6
4-5	23	53.5	26	60.5
12-month follow-up (*n* = 35)				
0-1	4	11.4	3	8.6
2-3	10	28.6	9	25.7
4-5	21	60.0	23	65.7

Note. At posttreatment, 25 participants (43.9%) met criteria on 4-5 measures for both reliable change and high endstate functioning. At 6-month follow-up, 18 participants (41.9%) met criteria on 4-5 measures for both reliable change and high endstate functioning. At 12-month follow-up, 18 participants (51.4%) met criteria on 4-5 measures for both reliable change and high endstate functioning.

Figure 1

Flow of Participants Through the Trial



Behavioral Experiments for Intolerance of Uncertainty:

A Randomized Clinical Trial for Adults with Generalized Anxiety Disorder

Second revision: March 17th 2022

Declarations of interest: none

Abstract

Sophisticated multicomponent treatments for adults with generalized anxiety disorder (GAD) have been developed over the past three decades. Although these comprehensive treatments have produced encouraging results, they appear to be less efficacious than treatments for other anxiety disorders. The goal of this randomized controlled trial is to test a newly developed, highly focused treatment for adults with GAD: *Behavioral Experiments for Intolerance of Uncertainty*. Sixty (60) participants (51 women, 9 men), with a mean age of 34.60 years (range: 19 to 67 years), were randomized to either treatment (n = 30) or wait-list control (n = 30). Treatment consisted of 12 weekly one-hour sessions in which participants learned to use behavioral experiments to test their catastrophic beliefs about uncertainty. Assessments were conducted at pre-, mid- and postcondition, and at 6- and 12-month follow-up. The primary outcome was the severity of GAD, and secondary outcomes were worry, depression, somatic anxiety, and intolerance of uncertainty. Using growth curve modeling, we found that (1) the treatment group was superior to the wait-list group in terms of change from pre- to posttest on all outcomes; (2) the combined sample (once wait-listed participants received treatment) evidenced large and significant decreases on all outcomes; and (3) treatment gains were either maintained or increased over the 12-month follow-up period of the study. The new treatment is a promising treatment option for adults with GAD considering that it may be as efficacious as more comprehensive evidence-based psychological treatments for GAD.

Keywords: Generalized anxiety disorder; cognitive-behavioral treatment; intolerance of uncertainty; behavioral experiments.

Behavioral Experiments for Intolerance of Uncertainty: A Randomized Clinical Trial for Adults with Generalized Anxiety Disorder

Generalized anxiety disorder (GAD) is a chronic and debilitating condition that is characterized by excessive and uncontrollable worry and anxiety about various life domains (American Psychiatric Association [APA], 2013). There exist a number of empirically supported, cognitive-behavioral treatments (CBT) for adults with GAD. Examples include metacognitive therapy (Wells & King, 2006), acceptance-based behavior therapy (Roemer & Orsillo, 2007), and CBT with a focus on intolerance of uncertainty (Ladouceur et al., 2000). These treatments have many things in common: (a) they are based on models of GAD that involve multiple affective, cognitive, and behavioral components; (b) they include several intervention strategies that target the model components; and (c) although they have received considerable empirical support, their efficacy appears to lag behind that of CBT protocols for other anxiety disorders (Hunot et al., 2010). Thus, despite being relatively sophisticated, CBT protocols for GAD have not produced optimal outcomes.

CBT with a focus on intolerance of uncertainty (CBT-IU) provides a good example of an established multicomponent treatment for GAD. The treatment includes four main modules: the reevaluation of the usefulness of worry, behavioral exposure to uncertainty, problem-solving training, and imaginal exposure (Robichaud et al., 2019). CBT-IU has been tested in five randomized clinical trials, with results showing that it is more efficacious than wait-list control (e.g., Ladouceur et al., 2000), supportive therapy (Gosselin et al., 2006), and to a lesser degree, applied relaxation (Dugas et al., 2010). However, a modified version of CBT-IU was found to be less effective than metacognitive therapy in one study directly comparing both treatments (van der Heiden et al., 2012). Although the results of the trials of CBT-IU are encouraging overall,

they also show that many clients do not fully benefit from the treatment; indeed, 45% to 50% do not achieve high endstate functioning at posttreatment. These numbers are similar to those obtained with other CBT protocols for GAD (Cuijpers et al., 2014).

One way to potentially increase the efficacy and clinical usefulness of treatment is to sharpen its focus. This can be accomplished by selecting the most important target construct and focusing exclusively on the chosen construct throughout treatment. Although each component of the model underlying CBT-IU has empirical support, the data show that intolerance of uncertainty (IU) is by far the most critical component. Over 25 years of basic and applied research has shown a consistent and robust relationship between IU and GAD. Despite the transdiagnostic nature of IU (for a review, see Carleton, 2012), the relationship between IU and GAD is not accounted for by shared variance with other anxiety disorders, mood disorders, negative affect, perfectionism or need for control (Buhr & Dugas, 2006; Norton et al., 2005). Data also suggest that IU is a *causal risk factor* for high levels of worry and GAD. Experimental studies have shown that change in IU leads to corresponding changes in GAD symptoms (e.g., Grenier & Ladouceur, 2004), and a five-year longitudinal study found that IU predicts subsequent levels of worry during adolescence (Dugas et al., 2012). Thus, data from experimental and longitudinal studies suggest that intolerance of uncertainty plays a key role in the etiology of GAD.

Safety Behaviors and GAD

Although some authors have argued that the judicious use of safety behaviors can be helpful in the early stages of therapy (e.g., Rachman et al., 2008), the weight of the evidence shows that safety behaviors contribute to the development and maintenance of anxiety disorders (e.g., Deacon & Maack, 2008; Olatunji et al., 2011). The literature on safety behaviors in GAD has lagged behind that of other anxiety disorders, possibly because the diagnostic criteria of

GAD do not include behavioral symptoms (see APA, 2013). However, recent theorizing is converging towards the notion that individuals with GAD frequently use safety behaviors (e.g., reassurance seeking) to increase their feelings of certainty (Beesdo-Baum et al., 2012). The consequence of using such safety behaviors is that rather than learning to cope with uncertainty, individuals with GAD attempt to decrease or avoid uncertainty in their daily lives. Ultimately, the use of certainty-seeking safety behaviors interferes with new learning about one's ability to cope with uncertainty, which ultimately maintains intolerance of uncertainty and GAD. It should be noted that worry itself can function as a safety behavior because thinking about many potential negative outcomes can be a way to decrease feelings of uncertainty about the future. However, because worry is largely a spontaneous and intrusive mental activity (as opposed to a deliberate safety behavior such as overpreparation), it is it is less amenable to direct change during treatment. Stated differently, even if one considers worry to be a form of safety behavior, its spontaneous nature precludes it from being the direct target of a treatment aiming to increase tolerance of uncertainty.

A New Treatment for GAD

Given the theoretical and empirical association between intolerance of uncertainty, safety behaviors, and GAD, we suggest that treatments for GAD should aim to decrease certainty-seeking safety behaviors. Though most established treatments do in fact address these behaviors, they may not do so in sufficient depth or in ways that explicitly promote new learning about uncertainty. However, behavioral experiments, a cognitive-behavioral technique with high evidential value, may offer important treatment advantages (see McMillan & Lee, 2010 for a review). Behavioral experiments are a personalized intervention strategy that require an individual to formulate specific predictions before entering an exposure-type situation. In the

treatment of GAD, this cognitive-behavioral technique can be used to systematically address safety behaviors and explicitly promote new learning about the contextual (What will happen?) and emotional (How will I feel?) sequelae of uncertainty. Behavioral experiments encourage expectancy violations, which may enhance inhibitory learning and retrieval (Gallistel & Gibbon, 2000) and may thereby promote new learning and superior treatment outcomes (for a review of inhibitory learning, see Craske et al., 2014). In summary, we propose that behavioral experiments are ideal to address intolerance of uncertainty in GAD because (1) they have strong evidential value; (2) they rely on experiential learning to promote change; (3) they are consistent with recent theorizing of fear reduction; and (4) when used as a vehicle for exposure, behavioral experiments promote the violation of expectations since they require clients to make specific predictions before entering exposure-type situations.

Based on the aforementioned considerations, Hebert and Dugas (2019) developed and tested a focused, single-component treatment for GAD: *Behavioral Experiments for Intolerance of Uncertainty*. The authors tested the new treatment, which relies exclusively on behavioral experiments to address safety behaviors and beliefs about uncertainty, in a case replication series with seven participants. Pre- to posttreatment effect sizes were large for all outcomes: severity of GAD symptoms, $d = 2.06$; worry, $d = 1.13$; depression, $d = 2.08$; and somatic anxiety, $d = 1.64$. Thus, it appears that focusing exclusively on intolerance of uncertainty using behavioral experiments may be a promising avenue of inquiry for the treatment of GAD.

The Current Study

The current study tests the new focused treatment in a larger sample within a randomized clinical trial. The experimental design consists of a 2 (conditions) X 5 (assessments) mixed factorial design, with repeated measures on the second factor. Sixty (60) participants were

randomly allocated to the experimental condition (Behavioral Experiments for Intolerance of Uncertainty) or the control condition (12-week Waiting List). Assessments were carried out at pretreatment, midtreatment, posttreatment, and at 6- and 12-month follow-ups.

The main goal of the study is to compare the effects of the new treatment to a waiting list on the symptoms of GAD, associated psychopathology (worry, depression, somatic anxiety) and cognitive vulnerability (intolerance of uncertainty). Hypothesis 1 is that relative to participants in the wait-list condition, those in the treatment condition will experience greater decreases in the severity of GAD, worry, depression, somatic anxiety, and intolerance of uncertainty. Hypothesis 2 states that behavioral experiments will lead to clinically significant change in the severity of GAD, in associated psychopathology, and in cognitive vulnerability. Finally, Hypothesis 3 is that treatments gains on all dependent variables will be maintained or augmented over the 12-month follow-up.

Method

Participants

The handling of study participants was in accordance with established ethical guidelines and the study was approved by the Human Research Ethics Committee of [MASKED]. In the following paragraphs, we report how we determined our sample size, all manipulations, and all measures in the study. This study's design was preregistered; see [MASKED].

The sample consisted of 60 French-speaking adults (51 women) with a primary diagnosis of GAD. Mean age was 34.60 years (*SD* = 12.47), with a range of 19 to 67 years. Eight (8) participants had completed high school, 19 had finished college, and 33 had a university degree. Fifty-four (54) participants self-identified as White/European Canadian, 3 as Black/African Canadian, 1 as Hispanic/Latinx, 1 as Asian Canadian, and 1 as Multi-ethnic. In addition, 30

participants were single, separated or divorced, 13 were married, and 17 were living in common-law relationships. Finally, 34 participants were employed, 4 were students, 15 were both employed and in school, and 7 were unemployed.

At initial assessment, the mean duration of GAD was 12.22 years (*SD* = 13.05; range 1-54) and the mean severity score for GAD was 5.48 (*SD* = 0.86; range 4-7) on the 9-point (0 to 8) *Clinician's Severity Rating* of the *Anxiety Disorders Interview Schedule for DSM-IV* (ADIS-IV; Di Nardo et al., 1994). Comorbid conditions were diagnosed in 40 participants, with 25 having one comorbid condition, 9 having two comorbid conditions, and 6 having three comorbid conditions. Secondary conditions were social anxiety disorder (*n* = 14), panic disorder (*n* = 9), agoraphobia (*n* = 9), specific phobia (*n* = 9), major depressive disorder (*n* = 7), posttraumatic stress disorder (*n* = 5), obsessive-compulsive disorder (*n* = 4), adjustment disorder (*n* = 2), dysthymic disorder (*n* = 1), and hypochondriasis (*n* = 1). Finally, 24 participants (40%) were taking anxiolytic or antidepressant medication and 17 (28.3%) had previously received CBT for an anxiety or mood disorder.

Procedure

Descriptions of the study were posted in local medical clinics and universities from November 2015 to May 2017. A total of 236 interested participants contacted the study coordinator in response to the posters. Following initial telephone screening, 92 participants were assessed by structured diagnostic interview. The final sample consisted of 60 adults with a primary diagnosis of GAD; 30 were randomized to the immediate treatment condition and 30 were randomized to the waiting list (see Figure 1 for a description of the flow of participants).

Individuals responding to the posters were screened initially over the telephone by a doctoral student using the *Telephone Screening Interview* (verbal consent was obtained prior to

beginning the interview). The interview, which was used in previous studies (e.g., [MASKED]), takes 20 to 30 minutes to administer. The goal of the telephone interview is to screen out individuals who clearly do not meet the study's inclusion criteria. Following the phone screening, potential participants were invited to the Psychological Services Clinic of the host university ([MASKED]) to sign the study consent form and receive a formal assessment by a second doctoral-level psychology student using the ADIS-IV. Following the assessment, a team meeting was be held to discuss the diagnostic impression from the ADIS-IV and to review inclusion/exclusion criteria. Excluded individuals were contacted by the interviewer and given appropriate referral. For individuals accepted into the study, the research coordinator applied a random allocation sequence using the Research Randomizer website (https://www.randomizer.org) and contacted participants to inform them of their condition allocation.

All participants accepted into the study (N = 60) were invited to the clinic for a final intake assessment session, in which they completed a series of precondition study questionnaires, including those used in the current study (*Penn State Worry Questionnaire, Beck Depression Inventory-II, Beck Anxiety Inventory, Intolerance of Uncertainty Scale*). The total time for the intake assessment session was approximately 90 minutes. Wait-listed participants were also assessed (ADIS-IV and self-report questionnaires) after 6 and 12 weeks on the waiting list. Participants randomized to the treatment condition were offered 12 weekly 60-minute treatment sessions. Following the 3rd treatment session, they completed measures of common therapy factors (*Working Alliance Inventory-Short Form, Credibility and Expectancy Scale for GAD*). The remaining assessments were conducted at mid- and posttreatment, and at 6- and 12-month

follow-up. These included administration of the ADIS-IV and completion of self-report questionnaires. Assessments and therapy were never conducted by the same person.

Inclusion criteria

Inclusion criteria were the following: (a) at least 18 years of age; (b) primary diagnosis of GAD; (c) no change in medication type or dose in 4 to 12 weeks before study entry (4 weeks for benzodiazepines, 12 weeks for antidepressants and hypnotics); (d) willingness to keep medication status stable while participating in the study; (e) no evidence of suicidal intent (based on clinical judgement); (f) no evidence of current substance abuse, schizophrenia or bipolar disorder; (g) no current participation in other trials; and (h) no evidence of anxiety symptoms due to a general medical condition based on clinical judgement (e.g., hypoglycemia, anemia).

Measures

Measure of Primary Outcome

The ***Anxiety Disorders Interview Schedule for DSM-IV*** (ADIS-IV; Di Nardo et al., 1994) assesses anxiety disorders, and screens for mood disorders, somatoform disorders, psychoactive substance use disorders, psychotic disorders, and medical problems. The interview provides information on the presence of disorders with severity ratings on a 9-point *Clinician's Severity Rating* scale ranging from 0 (*absent or none*) to 8 (*very severe or very severely disturbing/disabling*). Hereafter, the *Clinician's Severity Rating* from the ADIS-IV will simply be referred to as the CSR. In previous studies, we obtained reliability scores ranging from $\kappa = .66$ to $\kappa = .70$ for the presence and severity of GAD with the ADIS-IV.

Measures of Secondary Outcomes

The ***Penn State Worry Questionnaire*** (PSWQ; Meyer et al., 1990) includes 16 items that assess excessive worry. Items are rated on a 5-point scale, ranging from 1 (*not at all typical of*

me) to 5 (*very typical of me*). The PSWQ has high internal consistency, α = .86 to .95, and good test-retest reliability over four weeks, *r* = .74 to .93 (Molina & Borkovec, 1994). In the current sample, the internal consistency of the PSWQ at intake was α = .83.

The ***Beck Depression Inventory-II*** (BDI-II; Beck et al., 1996) includes 21 groups of 4 items reflecting different levels of depressive symptoms (e.g., sadness, loss of interest, indecision). Respondents indicate which item within each group best describes them over the past 2 weeks, with scores ranging from 0 to 3. The BDI-II has very good internal consistency, α = .92, and excellent test-retest reliability over a 1-week period, *r* = .93 (Beck et al., 1996). The internal consistency of the BDI-II at intake was α = .84 in the present sample.

The ***Beck Anxiety Inventory*** (BAI; Beck et al., 1988) is a 21-item measure of anxiety (primarily somatic anxiety) experienced during the past week. Items are rated on a 4-point scale, ranging from 0 (*not at all*) to 3 (*severely*). The BAI has high internal consistency, α = .92, and good test-rest reliability, *r* = .81, at one week in clinical samples (Beck et al., 1988). In the current study, the internal consistency of the BAI at intake was α = .87.

The ***Intolerance of Uncertainty Scale*** (IUS; Freeston et al., 1994) is a self-report measure consisting of 27 items assessing negative beliefs about uncertainty. Items are rated on a 5-point scale from 1 (*Not at all characteristic of me*) to 5 (*Entirely characteristic of me*), with higher scores reflecting greater intolerance of uncertainty. The IUS shows excellent internal consistency, α = .91, and evidence of convergent and divergent validity (Freeston et al., 1994). In the present study, the internal consistency of the IUS at intake was α = .91.

Measures of Common Therapy Factors

The ***Working Alliance Inventory-Short Form*** (WAI-SF; Horvath & Greenberg, 1989) is a 12-item measure of the client's participatory relationship with the therapist. The WAI-SF

measures agreement on therapeutic goals, agreement on the tasks of therapy, and the strength of the interpersonal bond between the client and therapist. Items are rated on a 7-point point scale ranging from 1 (*never*) to 7 (*always*), with higher scores reflecting a stronger alliance. The internal consistency of the WAI-SF was α = .79 at intake in the present sample.

The ***Credibility and Expectancy Scale for GAD*** (CES-GAD; Ladouceur et al., 2000) is based on the CES developed by Borkovec and Nau (1972). Items are rated on a 5-point scale ranging from 1 (*extremely weak*) to 5 (*extremely strong*). The internal consistency of the CES was α = .86 in the original validation study (Borkovec & Nau). The internal consistency of the CES-GAD was α = .83 in the present sample.

Experimental Conditions

Behavioral Experiments for IU

The treatment consisted of 12 weekly 60-minute sessions. In Sessions 1 and 2, participants were socialized to CBT and provided with information about the symptoms of GAD. Participants began basic symptom monitoring, and learned that worry, anxiety and safety behaviors are normal reactions to uncertainty. In sessions 3 to 10, participants tested their beliefs about uncertainty through a series of behavioral experiments. *Behavioral experiments involved: (1) selecting a specific thought to be tested (e.g., "uncertainty at work makes it impossible to be productive"); (2) designing a detailed experiment to test out the thought (e.g., taking on new and previously avoided responsibilities at work); (3) making specific predictions about what will happen (contextual prediction) and how they will feel (emotional prediction) during the experiment; (4) monitoring the contextual and emotional outcome of the experiment; and (5) comparing the predictions to the outcomes of the experiment and reflecting on what was learned.*

Sessions 11 and 12 consisted of relapse prevention, which involved reflecting on the new beliefs about uncertainty and behavioral patterns acquired during treatment.

Waiting List

Participants allocated to the control condition were placed on a 12-week waiting list. During the waiting period, they were contacted by telephone every 2 weeks to briefly assess their condition and screen for suicidal ideation (which none reported). Participants on the waiting list were assessed in person 6 and 12 weeks after condition allocation (mid- and post-waiting list, respectively). Following the post-waiting list assessment, they were offered the study treatment with full assessments at mid- and posttreatment, and at 6- and 12-month follow-up.

Therapists

Four graduate students with previous experience in CBT were responsible for the treatment of all participants. Two of the therapists completed their doctoral training and the other two completed their predoctoral internship during the study. The primary author ([MASKED]), a clinical psychologist with 30 years of experience in CBT, held 90-minute weekly group supervision meetings with all therapists throughout the study.

Data Analytic Approach

Baseline differences between the conditions on demographic and clinical variables were examined using the Statistical Package for the Social Sciences (SPSS version 25; IBM Corp.). Similarly, a series of t-tests were conducted in SPSS to compare therapists' treatment outcomes. We took a liberal approach in order to increase the chance of detecting any baseline differences or therapist effects if present and did not adjust for family-wise Type I error in these analyses.

We employed growth curve modeling to assess both short- and long-term outcomes. The multilevel modeling program *Hierarchical Linear Modeling* (HLM 7; Raudenbush et al., 2011)

was used to estimate growth curves for each participant, and to model the effect of condition (Treatment or Waiting List) on the rate of change. We also used HLM to estimate growth curves over follow-up. Separate analyses were performed for each of the measured outcomes. To better model outcome while taking account of the rate of attrition in this study, all participants who began treatment were retained in the analyses and all available data were used to estimate the expected intercepts and rates of change (slopes) for each individual over time. Effect sizes for the rates of change (slope) over time and the relative difference in the final outcome scores between the two conditions are presented as pseudo R^2, the percentage of variability in the outcome explained. We used model-adjusted least square means with *SEs* (unadjusted for multiple outcomes) to present the effect estimates of (1) the slope over time for the wait-list condition and (2) the difference between the slope of the treatment condition and the slope of the wait-list condition. Of note, the effect estimate of the difference between the slopes in the wait-list and treatment conditions is conceptually similar to an Group X Time interaction in a mixed-model ANOVA. To achieve this, the conditions were coded with the wait-list condition as the reference group (Condition = 0) relative to the treatment condition (Condition = 1), and the intercept (Time = 0) was set at posttest for the analyses of short-term outcomes and at the 12-month follow-up point for the analyses of long-term outcomes.

Clinically significant change was assessed in two ways. First, a reliable change index (Jacobson & Truax, 1991) was calculated for each participant on each outcome measure from pretreatment to posttreatment, and from pretreatment to 6-month and 12-month follow-ups. Second, endstate functioning was assessed for each participant on each outcome measure. Following guidelines set forth by Jacobson and Truax, and using existing population and clinical norms for each measure, a cutoff point was calculated to evaluate whether each participant's

posttreatment or follow-up score fell within the normal or clinical population range. Endstate functioning was defined as follows: participants who met criteria on 0 or 1 outcome measure were characterized as low, on 2 or 3 measures were deemed moderate, and on 4 or 5 measures were designated as high endstate functioning.

Power Analysis

Optimal Design Software program, which is a power program designed specifically for HLM, was used to calculate power. The study was powered *a priori* in order to detect a medium effect ($f = 0.25$) based on the effect sizes reported in prior studies with anxious symptomatology. With three assessment times, specifying a power of 0.90 and an alpha level of 0.05, and an estimated drop-out rate of 16.99% (mean dropout rate of CBT for GAD; Gersh et al., 2017), we sought to recruit at least 57 participants to achieve an N of 47 entering the study's follow-up phase. The data and all measures used in the current study can be obtained from the primary author ([MASKED]).

Results

Preliminary Analyses

Treatment adherence was assessed by two doctoral students who listened to audiotapes of all sessions for 8 participants or 14% of the sample (each student rated one randomly selected participant per therapist). The students used an intervention checklist to rate therapist interventions and noted whether every item was properly addressed. Treatment integrity was 95% for Therapist 1, 100% for Therapist 2, 99% for Therapist 3, and 99% for Therapist 4, suggesting that the study therapists closely adhered to the treatment manual.

We compared therapists in terms of the treatment outcomes attained by the participants under their care. Specifically, we compared pre-, mid- and posttreatment scores on the primary

outcome (CSR) for participants treated by each therapist. For each of the four therapists, participants' mean CSR scores significantly decreased from pre- to midtreatment and from mid- to posttreatment. Further, there were no significant differences between participants' mean CSR scores at each time point across the four therapists. Given the lack of significant differences between therapists at each time point and the similar patterns of improvement over time, we did not distinguish between therapists in the multilevel analyses of treatment outcome.

Final preliminary analyses compared the groups (Treatment vs. Waiting List) on pretest variables to see if they were comparable at the point of intake. We found no differences between the groups on demographic variables (age, sex, ethnic origin, education level, employment status, and marital status), clinical variables (duration of GAD, number of comorbid conditions, medication use, and previous experience with CBT), or pretest scores on the measures of dependent variables (CSR, PSWQ, BDI-II, BAI, and IUS). The multilevel analyses of outcome measures over time were therefore conducted without controlling for any demographic, clinical or pretest variables. See Supplemental Online Material for detailed demographic and clinical characteristics of participants in each condition.

Short-Term Outcomes: Treatment vs. Waiting List

Table 1 presents descriptive statistics for the measures of primary (CSR) and secondary outcomes (PSWQ, BDI-II, BAI, IUS) at pretest, midtest and posttest in the treatment and wait-list conditions. We evaluated the hypothesis that individuals receiving treatment would show greater improvement on all measures relative to those in the wait-list condition during the 12-week course of treatment by comparing the rates of change (slopes) in each measure pretest to posttest. For the CSR (primary outcome), though the slope decreased over time in both the

treatment and wait-list conditions, there was a significant effect of condition, with the treatment group showing significantly greater decreases in the severity of GAD over time.[1]

For the PSWQ (worry), the BDI-II (depressive symptoms) and the BAI (somatic anxiety), the pattern of results was similar to that of the CSR; although the slope decreased over time in both conditions, the rate of change was significantly greater in the treatment condition than in the control condition for each measure. As for the IUS (intolerance of uncertainty), we found no change in the slope from pretest to posttest in the wait-list condition, but a significant decrease in the slope in the treatment condition. In addition, the decrease in IUS scores was significantly greater in the treatment condition than in the wait-list condition. Effect estimates of the slope over time for all measures in the wait-list condition and in the treatment condition relative to wait-list, and the pretest to posttest effect size for this relative difference, are shown in Table 2.

Short-Term Outcomes: Combined Sample

After a 12-week delay, wait-listed participants were offered the study treatment, which resulted in a combined sample of 57 participants who started treatment (3 participants did not complete the wait-list period). A total of 48 completed treatment. Forty-five (45) individuals at least partially completed the assessment at 6-month follow-up, and 36 at least partially completed the 12-month follow-up. Descriptive statistics for all measures at each measurement time for the combined treatment sample are presented in Table 3.

Over the course of the 12-week treatment, the rate of change (slope) in the primary outcome measure in the combined treatment sample differed significantly from a slope of zero, evidencing a significant drop in GAD severity from pretreatment to posttreatment (the CSR slope coefficient = -1.39, SE = 0.11, $p < .001$, pseudo $R^2 = .80$). This linear change over time

[1] For all main analyses, the addition of a quadratic term to assess for non-linear change did not significantly contribute to the models that were tested. Therefore, non-linear results are not reported.

accounted for 80% of the within-participant variability in CSR scores from pretreatment to posttreatment. For the four secondary outcome measures, the rate of change (or slope) differed significantly from a slope of zero over the course of the 12-week treatment: PSWQ slope, coefficient = -11.08, SE = 0.92, $p < .001$, pseudo $R^2 = .69$; BDI-II slope, coefficient = -6.55, SE = 0.70, $p < .001$, pseudo $R^2 = .50$; BAI slope, coefficient = -8.09, SE = 0.80, $p < .001$, pseudo $R^2 = .62$; and IUS slope, coefficient = -15.16, SE = 1.49, $p < .001$, pseudo $R^2 = .50$. Thus, from pretest to posttest, there were large and significant decreases on all measures in the combined sample.

Long-Term Outcomes: Combined Sample

To test the hypothesis that treatment would lead to continued progress over follow-up, we compared the slope for GAD severity (CSR) with a slope of zero (a slope of zero denotes no change over time). There were further decreases in CSR scores over the 12 months following the end of treatment (CSR slope coefficient = -0.30, SE = 0.11, $p < .01$, pseudo $R^2 = .38$). In other words, further decreases in symptoms were observed over the follow-up period. This linear pattern of change over time accounted for 38% of the within-participant variability in CSR scores over the follow-up period.

We also compared the slope for each secondary outcome measure over the 12-month follow-up period with a slope of zero. For treatment completers in the combined sample (n = 48), the linear slopes for the four secondary outcome measures did not differ significantly from a slope of zero: PSWQ slope, coefficient = -0.26, SE = 0.89, $p > .05$, pseudo $R^2 = .36$; BDI-II slope, coefficient = -0.08, SE = 0.74, $p > .05$, pseudo $R^2 = .25$; BAI slope, coefficient = 0.85, SE = 0.73, $p > .05$, pseudo $R^2 = .15$; and IUS slope, coefficient = -0.96, SE = 1.58, $p > .05$, pseudo $R^2 = .11$. These findings suggest that treatment gains were maintained on each secondary outcome.

Clinically Significant Change

Frequencies and percentages of participants meeting criteria for reliable change and endstate functioning across study measures at posttreatment and follow-up are presented in Table 4. On the measure of worry (PSWQ), 47 of 57 (82.5%) participants fell within range of the normal population at posttreatment, as did 39 of 48 (81.3%) treatment completers (or 90.7% of 43 respondents) at 6-month follow-up, and 29 of 48 (60.4%) treatment completers (or 80.9% of 35 respondents) at 12-month follow-up. On the IUS, a measure of the primary target of the intervention, 40 of 57 (70.2%) participants met criteria for reliable change at posttreatment, as did 33 of 48 (68.8%) treatment completers (or 76.7% of 43 respondents) at 6-month follow-up, and 27 of 48 (56.3%) treatment completers (or 77.1% of 35 respondents) at 12-month follow-up.

Common Therapy Factors and Medication

Following the third treatment session, participants rated the quality of the therapeutic alliance (WAI-SF) as well as treatment credibility and expectations of change (CES-GAD). Mean scores were 74.13 (*SD* = 8.13) on the WAI-SF and 25.98 (*SD* = 2.86) on the CES-GAD for the 56 participants having completed three sessions. Scores on the WAI-SF were comparable to those typically reported in previous treatment studies (for a review, see Sturgiss et al., 2019). As for the CES-GAD, scores obtained in the current study were almost identical to those reported in earlier studies of CBT-IU (Dugas et al., 2010; Ladouceur et al., 2000). Thus, the quality of the therapeutic alliance as well as the credibility of treatment and expectations of change appear to be as strong as with a multicomponent treatment for GAD.

As mentioned previously, 24 of 60 participants (40%) were taking anxiolytic or antidepressant medication at the beginning of the study. For the combined sample at posttreatment, 20 of 49 participants (40.8%) continued to use anxiolytic or antidepressant medication. Of the 24 participants who were taking medication at intake, 3 did not complete

posttreatment assessments and one discontinued their medication. Thus, the treatment had a negligible impact on medication use in this study.

Discussion

The current study provides support for the use of behavioral experiments to increase tolerance of uncertainty in adults with GAD. The results show that, compared to a waiting list, the focused treatment led to greater decreases in the severity of GAD, associated psychopathology (worry, depressive symptoms, somatic anxiety), and cognitive vulnerability (intolerance of uncertainty). The findings also reveal large and significant decreases on all outcomes for the total sample, once participants in the control condition received treatment after a 12-week waiting period. Finally, we found that treatment gains on all outcomes were either maintained or increased (for the severity of GAD) over the 12-month follow-up period of the study. Thus, it appears that a single-component treatment, *Behavioral Experiments for Intolerance of Uncertainty*, represents a promising treatment option for individuals with GAD.

Although cross-study comparisons should be made with extreme caution, they can nonetheless be informative in terms of ruling out large differences between findings. Keeping this in mind, it appears that Behavioral Experiments for Intolerance of Uncertainty (a single-component treatment) and CBT with a Focus on IU (a multicomponent treatment) may lead to similar outcomes. Overall, the results of the current study appear to be (at least) comparable to those of previous clinical trials of the multicomponent treatment in a wait-list design. One notable difference, however, may be that the single-component treatment produces greater decreases in intolerance of uncertainty. In previous trials of the multicomponent treatment (Dugas et al., 2010; Gosselin et al., 2006; Ladouceur et al., 2000; van der Heiden, 2012), within-group effect sizes on the IUS (for the intent-to-treat sample) ranged from $d = 0.58$ to $d = 0.72$.

By comparison, the within-group effect sizes on the IUS (also intent-to-treat) were $d = 1.72$ in the Hebert and Dugas (2019) study and $d = 1.49$ in the current study. It may be that by focusing exclusively on intolerance of uncertainty using behavioral experiments, the single-component treatment produces larger improvements in cognitive vulnerability for GAD. This possibility, however, awaits testing in a clinical trial directly comparing both treatments.

It should be noted that, in addition to the encouraging findings of the current study, the new treatment presents a number of important advantages compared to more comprehensive treatments. Most importantly, an added benefit of using a single treatment strategy focusing exclusively on intolerance of uncertainty is that the resulting treatment is highly parsimonious. A number of authors have called for a move away from traditional assessments of the quality of therapy and towards newer assessment models that take into consideration factors such as treatment parsimony (e.g., Cougle, 2012; Mazzucchelli et al., 2009). As argued by these authors, given equivalent outcomes between two different treatments, the more parsimonious approach should be preferred. Why? First, more parsimonious treatment options may require less clinical training and may be easier to disseminate. For full-time therapists who have limited time to devote to learning new treatment protocols, the amount of time required to learn new treatment procedures is an important consideration. Second, more parsimonious treatments may increase treatment integrity; the lower the number of treatment components, the greater the likelihood that the treatment will be implemented as intended. Given that treatment integrity may predict positive outcomes, treatments that are more likely to be administered as intended can offer important advantages. Third, clients may adhere more easily to treatments that are parsimonious. Relatedly, treatment receipt and enactment may be limited if treatments are overly inclusive. Finally, from a theoretical point of view, the active ingredients of therapy can be more easily

identified when treatments have fewer components. Conversely, multicomponent treatment protocols often require dismantling studies to tease apart active and non-active ingredients. Thus, it appears that assessments of treatment quality that do not take into consideration the issue of parsimony may not be ideal; ease of training, ease of dissemination, treatment integrity, client adherence, and identification of active ingredients are important factors to consider in addition to treatment outcome.

Relatedly, previous attempts to increase the efficacy of CBT for GAD using more comprehensive methods have been met with limited success. For example, studies that have combined CBT with pharmacotherapy have produced inconclusive results (e.g., Bond et al., 2002). Other studies, which have attempted to increase the efficacy of CBT for GAD by offering more comprehensive CBT, have also produced inconclusive results (e.g., Newman et al., 2011). The current study, which tested a more focused treatment, not only holds the promise of contributing to our understanding and ability to treat GAD, but also of leading the way to increasing the dissemination and accessibility of psychological treatments for GAD.

Having acknowledged the advantages of parsimonious treatments, one should nonetheless keep in mind that comprehensive treatment options also have important advantages. Considering that individuals with GAD may be a particularly heterogenous group, multicomponent treatments can offer therapists greater flexibility in emphasizing interventions that are either highly valued by their clients or appear to be directly related to their clients' greatest vulnerabilities. For example, because intolerance of uncertainty is associated with ineffective problem solving (Clarke et al., 2017) and cognitive avoidance (Koerner & Dugas, 2006), some clients with GAD may prefer (or particularly benefit from) problem-solving training or imaginal exposure, both of which are not germane to the new treatment.

Strengths and Limitations

One of the main strengths of the current study is that the treatment intervention (behavioral experiments) and treatment target (intolerance of uncertainty) are solidly grounded in the anxiety disorders literature. As mentioned previously, behavioral experiments, which require clients to formulate specific predictions and to compare their predictions to actual outcomes following an exposure-type situation, are directly in line with contemporary learning theories of fear reduction (e.g., Craske et al., 2014). As for intolerance of uncertainty, it bears repeating that over 25 of research have supported its central role in the etiology and treatment of GAD (for a review, see Robichaud et al., 2019). Consequently, both the specific intervention and target of the new treatment constitute an important strength of the current study.

A second strength of the study is that the treatment is based on a model of GAD that is closely tied to a generic cognitive-behavioral model of psychopathology (see e.g., Tolin, 2016). Namely, individuals who are intolerant of uncertainty tend to make catastrophic misinterpretations of uncertainty when faced with situations that are novel, unpredictable or ambiguous. These misinterpretations then lead to emotional (anxiety), cognitive (worry), and behavioral (avoidance, safety behaviors) symptoms. Simply stated, the model of GAD and the generic cognitive-behavioral model of psychopathology both emphasize the activation of latent core beliefs by precipitating events, leading to biased information processing and emotional-cognitive-behavioral symptoms. Given the numerous and heterogenous cognitive-behavioral models of GAD currently under study, the development of treatments that are closely tied to the general theory of CBT may hold promise for moving the field forward in terms of a unified understanding of anxiety and other forms of psychopathology.

One important limitation of the current study is that it was conducted in a single site with a relatively homogenous (largely white, female) community sample, which restricts the generalizability of the findings to other settings and populations. Although the use of a single site is a tangible limitation of the study, it is lessened by the fact that we recruited participants from community medical clinic waiting lists. Further, compared to multisite studies, single-site studies offer important advantages in terms of internal validity (e.g., consistency of procedures, uniformity of care), particularly in the early stages of treatment validation. In terms of the sample, it can be argued that the disadvantages of having a homogenous sample is offset by the advantages of having a Francophone sample because the latter are noticeably under-represented in the clinical literature.

Another limitation of the study relates to the use of a wait-list design. By using such a control condition, we are not in a position to disentangle the effects of treatment specific factors (e.g., change in certainty-seeking safety behaviors, change in beliefs about uncertainty) from those of common therapy factors (e.g., time spent with therapist, therapeutic alliance, treatment motivation). Although we are aware that the use of a supportive therapy control condition would have allowed us to disentangle specific and common effects, we nonetheless opted for wait-list control to limit the costs of the trial. From a treatment development perspective, the use of a wait-list control condition at this point of treatment validation is appropriate. According to best practice guidelines for treatment development (Hayes et al., 2013), the validation of new treatment procedures should follow a cost-effective sequence starting with case replication (i.e., Hebert & Dugas, 2019), followed by graded control conditions requiring more and more participants (e.g., waiting list, supportive therapy, competing treatment). By using a wait-list

condition, we were in a position to collect controlled data on the new treatment in a cost-effective manner, which is a judicious and ethical choice at this stage of treatment development.

A third limitation of the study relates to the putative mechanism underlying the behavioral experiments. Although every effort was made to ensure that exposure to uncertainty was at the heart of each behavioral experiment, we are not in a position to rule out the possibility that other constructs related to GAD were simultaneously targeted. For example, an experiment that involves waiting 48 hours before attempting to solve a problem (exposure to the uncertainty of not knowing if the problem can be solved) may also foster greater acceptance of emotional distress (a central construct in acceptance-based behavior therapy). In future studies, it would therefore be important to test hypothesized mediators from competing models.

In conclusion, it appears that a focused, single-component treatment, *Behavioral Experiments for Intolerance of Uncertainty*, represents an interesting treatment option for adults with GAD. In addition to the encouraging findings reported above, the focused treatment has the advantage of being parsimonious and of being compatible with general cognitive-behavioral models of psychopathology. We suggest that treatment models that emphasize specific targets (e.g., intolerance of uncertainty) while recognizing common change processes (e.g., prioritizing behavioral change within a learning theory framework) may represent an interesting avenue for the further development and dissemination of CBT.

References

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Beck, A. T., Epstein, N., Brown, G. K., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897. http://doi.org/10.1037/0022-006X.56.6.893

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory Manual* (2nd ed.). San Antonio, TX: Psychological Corporation.

Beesdo-Baum, K., Jenjahn, E., Hofler, M, Lueken, U, Becker, E.S., & Hoyer, J. (2012). Avoidance, safety behavior, and reassurance seeking in generalized anxiety disorder. *Depression and Anxiety, 29,* 948-957. http://doi.org/10.1002/da.21955

Bond, A. J., Winegrove, J., Curran, H. V., & Lader, M. H. (2002). Treatment of generalized anxiety disorder with a short course of psychological therapy, combined with buspirone or placebo. *Journal of Affective Disorders, 72,* 267-271. http://doi.org/10.1016/S0165-0327(01)00469-4

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3,* 257-260. http://doi.org/10.1016/0005-7916(72)90045-6

Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders, 20,* 222-236. http://doi.org/10.1016/j.janxdis.2004.12.004

Carleton, R. N. (2012). The intolerance of uncertainty construct in the context of anxiety disorders: Theoretical and practical perspectives. *Expert Review of Neurotherapeutics, 12*, 937-947. http://doi.org/10.1586/ern.12.82

Clarke, J. B., Ford, M., Heary, S., Rodgers, J., & Freeston, M. H. (2017). The relationship between negative problem orientation and worry: A meta-analytic review. *Psychopathology Review, a4,* 319-340. https://doi.org/10.5127/pr.034313

Cougle, J. R. (2012). What makes a quality therapy? A consideration of parsimony, ease, and efficiency. *Behavior Therapy, 43,* 468-481. http://doi.org/10.1016/j.beth.2010.12.007

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T.,Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behavior Research and Therapy, 58,* 10-23. http://doi.org/10.1016/j.brat.2014.04.006

Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clinical Psychology Review, 34,*130-40. http://doi.org/10.1016/j.cpr.2014.01.002

Deacon, B., & Maack, D. J. (2008). The effects of safety behaviors on the fear of contamination: An experimental investigation. *Behaviour Research and Therapy, 46,* 537-47. http://doi.org/10.1016/j.brat.2008.01.010

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV). Psychological Corporation: San Antonio, TX.

Dugas, M. J., Brillon, P., Savard, P., Turcotte, J., Gaudet, A., Ladouceur, R., Leblanc, R., & Gervais, N. J. (2010). A randomized clinical trial of cognitive-behavioral therapy and applied relaxation for adults with generalized anxiety disorder. *Behavior Therapy, 41,* 46-58. http://doi.org/10.1016/j.beth.2008.12.004

Dugas, M. J., Laugesen, N., & Bukowski, W. M. (2012). Intolerance of uncertainty, fear of anxiety, and adolescent worry. *Journal of Abnormal Child Psychology, 40,* 863-870. http://doi.org/10.1007/s10802-012-9611-1

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17,* 791-802. http://doi.org/10.1016/0191-8869(94)90048-5

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*, 289-344. http://doi.org/10.1037/0033-295X.107.2.289

Gersh, E., Hallford, D. J., Rice, S. M., Kazantzis, N., Gersh, H., Gersh, B., & McCarty, C. A. (2017). Systematic review and meta-analysis of dropout rates in individual psychotherapy for generalized anxiety disorder. *Journal of Anxiety Disorders, 52*, 25-33. http://dx.doi.org/10.1016/j.janxdis.2017.10.001

Gosselin, P., Ladouceur, R., Morin, M., Dugas, M. J., & Baillargeon, L. (2006). Benzodiazepine discontinuation among adults with GAD: A randomized trial of cognitive-behavioral therapy. *Journal of Consulting and Clinical Therapy, 74*, 908-919. http://doi.org/10.1037/0022-006X.74.5.908

Grenier, S., & Ladouceur, R. (2004). Manipulation de l'intolérance à l'incertitude et inquiétudes [Manipulation of intolerance of uncertainty and worries]. *Canadian Journal of Behavioural Science, 36,* 56–65. https://doi.org/10.1037/h0087216

Hayes, S. C., Long, D. M., Levin, M. E., & Follette, W. C. (2013). Treatment development: Can we find a better way? *Clinical Psychology Review, 33*, 870–882. https://doi.org/10.1016/j.cpr.2012.09.009

Hebert, E. A., & Dugas, M. J. (2019). Behavioral experiments for intolerance of uncertainty: Challenging the unknown in the treatment of generalized anxiety disorder. *Cognitive and Behavioral Practice, 26,* 421-436. https://doi.org/10.1016/j.cbpra.2018.07.007

Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223-233. http://doi.org/10.1037/0022-0167.36.2.223

Hunot, V., Churchill, R., Teixeira, V., & Silva de Lima, M. (2010). Psychological therapies for generalised anxiety disorder (review). *Cochrane Database of Systematic Reviews 2007, Issue 1*, Art. No. CD001848. http://doi.org/10.1002/14651858.CD001848.pub4

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

Koerner, N., & Dugas, M. J. (2006). A cognitive model of generalized anxiety disorder: The role of intolerance of uncertainty. In G. C. L. Davey & A. Wells (Eds.), *Worry and its psychological disorders: Theory, assessment and treatment* (pp. 201-216). Chichester: John Wiley and Sons.

Ladouceur, R., Dugas, M. J., Freeston, M. H., Léger, E., Gagnon, F., & Thibodeau, N. (2000). Efficacy of a cognitive-behavioral treatment for generalized anxiety disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 68*, 957-964. http://doi.org/10.1037/0022-006X.68.6.957

Mazzucchelli, T., Kane, R., & Rees, C. (2009). Behavioral activation treatments for depression in adults: A meta-analysis and review. *Clinical Psychology: Science and Practice, 16*, 383-411. https://doi.org/10.1111/j.1468-2850.2009.01178.x

McMillan, D., & Lee, R. (2010). A systematic review of behavioral experiments vs. exposure alone in the treatment of anxiety disorders: A case of exposure while wearing the emperor's new clothes? *Clinical Psychology Review, 30,* 467-478. http://doi.org/10.1016/j.cpr.2010.01.003

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behavior Research and Therapy, 28,* 487-495. http://doi.org/10.1016/0005-7967(90)90135-6

Molina, S., & Borkovec, T. D. (1994). The Penn State Worry Questionnaire: Psychometric properties and associated characteristics. In G. C. L. Davey & F. Tallis (Eds.), Worrying: Perspectives on theory, assessment and treatment (pp. 265–283). New York: John Wiley & Sons.

Newman, M. G., Castonguay, L. G., Borkovec, T. D., Fisher, A. J., Boswell, J. F., Szkodny, L. E., & Nordberg, S. S. (2011). A randomized controlled trial of cognitive-behavioral therapy for generalized anxiety disorder with integrated techniques from emotion-focused and interpersonal therapies. *Journal of Consulting and Clinical Psychology, 79,* 171-181. http://doi.org/10.1037/a0022489

Norton, P. J., Sexton, K. A., Walker, J. R., & Norton, G. R. (2005). Hierarchical model of vulnerabilities for anxiety: Replication and extension with a clinical sample. *Cognitive Behavior Therapy, 34,* 50-63. http://doi.org/10.1080/16506070410005401

Olatunji, B. O., Etzel, E. N., Tomarken, A. J., Ciesielski, B. G., & Deacon, B. (2011). The effects of safety behaviors on health anxiety: An experimental investigation. *Behaviour Research and Therapy, 49,* 719-28. http://doi.org/10.1016/j.brat.2011.07.008

Rachman, S., Radomsky, A. S., & Shafran, R. (2008). Safety behaviour: A reconsideration. *Behaviour Research and Therapy, 46,* 163-173. http://doi.org/10.1016/j.brat.2007.11.008

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM: Linear and nonlinear modeling.* Skokie, IL: Scientific Software International.

Robichaud, M., Koerner, N., & Dugas, M. J. (2019). *Cognitive-behavioral treatment for generalized anxiety disorder: From science to practice (2nd ed.).* New York: Routledge.

Roemer, L., & Orsillo, S. M. (2007). An open trial of acceptance-based behavior therapy for generalized anxiety disorder. *Behavior Therapy, 38,* 72–85. http://doi.org/10.1016/j.beth.2006.04.004

Sturgiss, E. A., Rieger, E., Haesler, E., Ridd, M. J., Douglas, K., & Galvin, S. L. (2019). Adaption and validation of the Working Alliance Inventory for general practice: Qualitative review and cross-sectional surveys. *Family Practice, 36,* 516-522. http://doi.org/10.1093/fampra/cmy113

Tolin, D. F. (2016). *Doing CBT: A comprehensive guide to working with behaviors, thoughts, and emotions.* New York: The Guilford Press.

van der Heiden, C., Muris, P., & van der Molen, H. T. (2012). Randomized controlled trial of the effectiveness of metacognitive therapy and intolerance-of-uncertainty therapy for generalized anxiety disorder. *Behavior Research and Therapy, 50,* 100-109. http://doi.org/10.1016/j.brat.2011.12.005

Wells, A., & King, P. (2006). Metacognitive therapy for generalized anxiety disorder: An open trial. *Journal of Behavior Therapy and Experimental Psychiatry, 37,* 206-212. http://doi.org/10.1016/j.jbtep.2005.07.002

Table 1

Means and Standard Deviations on Outcome Measures in Both Conditions at Pretest, Midtest, and Posttest

Measure and condition	Pretest *n* = 60 *M (SD)*	Midpoint *n* = 53 *M (SD)*	Posttest *n* = 50 *M (SD)*
CSR			
TRT	5.48 (0.83)	4.35 (1.21)	2.46 (1.41)
WL	5.47 (0.91)	5.44 (0.75)	5.17 (0.94)
PSWQ			
TRT	67.43 (7.51)	55.77 (9.77)	44.13 (8.83)
WL	64.00 (6.48)	61.48 (7.80)	60.00 (8.33)
BDI-II			
TRT	21.70 (9.74)	13.46 (11.71)	7.96 (7.34)
WL	20.70 (7.97)	16.52 (9.60)	15.81 (8.70)
BAI			
TRT	25.23 (11.76)	14.31 (11.06)	8.30 (5.84)
WL	24.23 (9.68)	21.15 (10.90)	18.89 (9.96)
IUS			
TRT	85.07 (19.48)	68.15 (21.98)	50.48 (15.73)
WL	79.70 (15.45)	78.52 (20.48)	79.30 (20.29)

Note. CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; TRT = treatment; WL = waiting list; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

Table 2

Results of Multilevel Analysis of Outcome Measures, Including Random Slopes

Outcome	Fixed effects	Intercept (posttest)[a] coeff. (*SE*)	Pseudo R^2	Slope coeff. (*SE*)	Pseudo R^2
CSR	WL (ref grp)[b]	5.18*** (0.16)	.–	-0.16*** (0.06)	
	TRT vs. WL[c]	-2.46*** (0.32)	.63	-1.27*** (0.15)	.86
PSWQ	WL (ref grp)	59.95*** (1.63)	.–	-1.96* (0.77)	
	TRT vs. WL	-15.92*** (2.46)	.56	-9.73*** (1.46)	.69
BDI-II	WL (ref grp)	15.38*** (1.69)	.–	-7.79** (2.39)	
	TRT vs. WL	-2.41*** (0.59)	.24	-4.42*** (1.22)	.78
BAI	WL (ref grp)	18.93*** (1.93)	.–	-2.61** (0.86)	
	TRT vs. WL	-11.23*** (2.38)	.39	-5.81*** (1.49)	.39
IUS	WL (ref grp)	79.30*** (4.06)	.–	-0.07 (1.78)	
	TRT vs. WL	-28.83*** (5.27)	.47	-17.33*** (2.78)	.82

Note. coeff. = coefficient; *SE* = standard error; Pseudo R^2 = proportion of between-person (intercept) or within-person (slope) variance explained; ref grp = reference group; CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; WL = waiting list; TRT = treatment; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

[a] Pretest coded as -2, midpoint coded as -1, and posttest coded as 0, such that the intercept reflects the estimated fixed effects at posttest (Time = 0).

[b] WL condition coded as 0; this effect therefore reflects the effect of Time in the WL condition.

[c] TRT condition coded as 1; this effect therefore reflects the effect of Time in the TRT condition relative to the effect of Time in the reference group (the WL condition).

*p < .05 ** p < .01 *** p < .001

Table 3

Means and Standard Deviations on Outcome Measures for Combined Sample at all Measurement Times

Measure	Pretreatment (n = 57) *M (SD)*	Midtreatment (n = 52) *M (SD)*	Posttreatment (n = 48) *M (SD)*	6 months ($n \geq 43$[a]) *M (SD)*	12 months ($n \geq 35$[b]) *M (SD)*
CSR	5.33 (0.88)	4.35 (1.10)	2.63 (1.50)	2.16 (1.55)	2.03 (2.56)
PSWQ	63.91 (8.69)	53.44 (10.47)	43.59 (10.92)	44.63 (9.77)	43.43 (10.74)
BDI-II	18.91 (9.65)	13.44 (9.85)	8.02 (6.86)	8.26 (6.33)	7.80 (8.33)
BAI	22.22 (11.31)	14.23 (9.33)	8.48 (6.28)	10.14 (5.73)	9.94 (9.73)
IUS	82.33 (19.90)	67.92 (20.49)	51.93 (17.97)	51.52 (17.54)	50.46 (17.33)

Note. CSR = Clinician's Severity Rating from the Anxiety Disorders Interview Schedule for DSM-IV; PSWQ = Penn State Worry Questionnaire; BDI-II = Beck Depression Inventory II; BAI = Beck Anxiety Inventory; IUS = Intolerance of Uncertainty Scale.

[a] n = 45 completed the ADIS; n = 43 completed the self-report measures. [b] n = 36 completed the ADIS; n = 35 completed the self-report measures at 12-month follow-up.

Table 4

Frequency and Percentages of Participants (n = 57) Meeting Criteria for Reliable Change and Endstate Functioning at Posttreatment, 6-Month Follow-Up, and 12-Month Follow-Up

	Reliable change		Endstate functioning	
No. of measures	Frequency	%	Frequency	%
Posttreatment				
0-1	14	24.6	10	17.5
2-3	13	22.8	13	22.8
4-5	30	52.6	34	59.7
6-month follow-up (*n* = 43)				
0-1	6	14.0	3	7.0
2-3	14	32.6	14	32.6
4-5	23	53.5	26	60.5
12-month follow-up (*n* = 35)				
0-1	4	11.4	3	8.6
2-3	10	28.6	9	25.7
4-5	21	60.0	23	65.7

Note. At posttreatment, 25 participants (43.9%) met criteria on 4-5 measures for both reliable change and high endstate functioning. At 6-month follow-up, 18 participants (41.9%) met criteria on 4-5 measures for both reliable change and high endstate functioning. At 12-month follow-up, 18 participants (51.4%) met criteria on 4-5 measures for both reliable change and high endstate functioning.

Figure 1

Flow of Participants Through the Trial