Journal of Anxiety Disorders 118 (2026) 103112

Contents lists available at ScienceDirect

Journal of Anxiety Disorders

journal homepage: www.elsevier.com/locate/janxdis

Behavioural experiments for intolerance of uncertainty: A brief intervention delivered via videoconference for adults with generalised anxiety disorder

Emily J. Wilson [a,b], Maree J. Abbott [b,*], Alice R. Norton [b], Jessica Riley [b], David Berle [c]

[a] Graduate School of Health, University of Technology Sydney, Australia
[b] School of Psychology, The University of Sydney, Australia
[c] School of Medicine and Psychology, Australian National University, Australia

ARTICLE INFO

Keywords:
Generalised anxiety disorder
GAD
Intolerance of uncertainty
Cognitive behaviour therapy
Treatment outcomes

ABSTRACT

Accessing psychological treatment is often met with barriers of time, cost, and availability. Focused brief interventions delivered via videoconference can overcome some of these barriers. We sought to evaluate the efficacy and feasibility, as well as processes of exposure-based learning for a brief intervention for treating generalised anxiety disorder (GAD), utilising behavioural experiments delivered via videoconference. Participants ($N = 40$) with a primary diagnosis of GAD were assessed via clinical interview and randomised to either the treatment condition ($n = 20$) or waitlist condition ($n = 20$). Treatment consisted of two weekly 1-hour sessions where participants utilised behavioural experiments to test negative beliefs about uncertainty. The primary outcomes were worry, safety behaviours, avoidance, depression, anxiety, physiological tension, and intolerance of uncertainty (IU). Linear mixed models indicated that the treatment group was only superior to the waitlist group on change from pre- to post-treatment for worry. The combined group (once waitlisted participants received treatment) evidenced significant reduction across all outcomes from pre- to post-treatment, except for anxiety. Additionally, there was evidence that expectancy violation and habituation occurred, suggesting that behavioural experiments facilitate different processes of exposure-based learning. The intervention was also found to be acceptable, appropriate, and feasible by adults with GAD. Thus, the remotely delivered brief intervention shows promise as an effective option for individuals with GAD.

1. Introduction

Generalised anxiety disorder (GAD) is typified by excessive worry and anxiety, that is persistent and marked by a range of physiological symptoms including restlessness, irritability, muscle tension, concentration disturbance, and/or sleep difficulties (American Psychiatric Association, 2022). Individuals with GAD report levels of mild to moderate disability, and moderate to severe levels of psychological distress (Pelletier et al., 2017). A recent cross-sectional survey across 26 countries found GAD had a lifetime prevalence of 3.7% (Ruscio et al., 2017), consistent with surveys reporting a prevalence of 6.1% in the Australian (McEvoy et al., 2011) and 4.3% in the US population (Kessler et al., 2012).

Cognitive behavioural therapy (CBT) is the gold-standard treatment for GAD (Newman et al., 2022). Despite its established efficacy, a considerable proportion of adults with mental health difficulties do not engage in or complete CBT (Harris et al., 2015). This gap in treatment uptake can be attributed to several factors, including individuals' preference for managing symptoms independently, financial and time constraints, and the additional burden of coping with co-occurring physical and mental health conditions (Coombs et al., 2021). While low-intensity, internet-delivered CBT offers a potential solution to these barriers, recent research investigating obstacles to care for adults with GAD symptoms found the majority of participants preferred real-time contact with an individual therapist, through face-to-face or video-conference (Basile et al., 2024). Currently there is a lack of research regarding the efficacy and acceptability of briefer interventions delivered by a therapist for individuals with GAD that could overcome some of the barriers to treatment.

There are multiple cognitive-behavioural conceptualisations of GAD

* Correspondence to: Brain and Mind Centre, School of Psychology, The University of Sydney, 94 Mallett Street, Camperdown, New South Wales 2050, Australia
E-mail address: maree.abbott@sydney.edu.au (M.J. Abbott).

https://doi.org/10.1016/j.janxdis.2026.103112
Received 5 May 2025; Received in revised form 23 December 2025; Accepted 5 January 2026
Available online 6 January 2026

(Freeston, 2023), and as a result there are numerous CBT-based treatments for the disorder (Newman et al., 2022). The intolerance of uncertainty model (IUM) of GAD suggests that intolerance of uncertainty (IU) is the key maintaining factor of excessive worry, along with three secondary processes: negative problem orientation, cognitive avoidance, and positive beliefs about worry (Dugas et al., 1998). IU is a set of negative beliefs about uncertainty and its consequences (Robichaud et al., 2019). From the IUM, a treatment targeting each component of the model was developed, Cognitive Behaviour Therapy-Intolerance of Uncertainty (CBT-IU). A meta-analysis found that CBT-IU was significantly more effective than general CBT at reducing worry and IU at post-treatment (Wilson et al., 2023). These findings suggest that the reduction in worry and IU occurs faster in CBT-IU and may be of benefit for briefer interventions. In addition, a recent study found that CBT-IU was comparable in both face-to-face and videoconference settings (Bouchard et al., 2022).

Given the importance of IU in GAD, Hebert and Dugas (2019) proposed a model centring on IU as the driving factor leading to worry, anxiety, and safety behaviours. A treatment protocol was developed using one technique from CBT-IU, namely behavioural experiments targeting IU (Hebert & Dugas, 2019). The treatment protocol was developed to include 12-weekly, 50-minute sessions, with session one and two involving psychoeducation and uncertainty awareness training, sessions three to eleven focusing on behavioural experiments testing beliefs about uncertainty, and session twelve providing relapse prevention. A case series using the protocol found large effect sizes across worry, IU, and anxiety for seven participants with GAD (Hebert & Dugas, 2019). A randomised controlled trial (RCT) compared the behavioural experiment treatment to a 12-week waitlist condition, finding that the treatment condition had significantly greater reduction from pre-to post-treatment across GAD severity, worry, IU, anxiety and depression, compared to waitlist (Dugas et al., 2022). Thus, a brief intervention utilising behavioural experiments targeting beliefs about uncertainty could be an effective option.

Behavioural experiments are a technique that requires clients to formulate specific expectations about a feared situation and compare these expectations to the actual outcome following exposure (Bennett-Levy et al., 2004). Behavioural experiments have been found to instigate belief change, with research demonstrating that they are effective at reducing the endorsement of specific target beliefs in adults with OCD cognitions (McManus et al., 2012). Behavioural experiments utilise inhibitory-based learning as they are set up to facilitate expectancy violation regarding the feared situation (Craske et al., 2008). Specifically, reflections on the outcome of the behavioural experiment are centred around whether the expected negative outcome occurred or not, as well as exploring how 'bad' and how well they 'coped' if the negative outcome did occur are important in this type of learning (Craske et al., 2014).

Behavioural experiments targeting IU aim to reduce the endorsement of specific idiosyncratic beliefs about uncertainty. Individuals with anxiety disorders demonstrate an interpretation bias towards generating negative interpretations towards ambiguous or uncertain material (Hirsch et al., 2016). Specifically, individuals with anxiety disorders report significantly more concern for negative and ambiguous situations, compared to control participants (Anderson et al., 2012). As such, behavioural experiments may not only alter specific beliefs but also may facilitate broader cognitive change in the way individuals interpret ambiguous or uncertainty information their environment.

Behavioural experiments targeting IU may also assist individuals with habituation. Habituation is the reduction in anxiety in the presence of a feared situation, while minimising the use of safety behaviours (Benito & Walther, 2015; Rankin et al., 2009). Within GAD, research has largely focussed on identifying the cognitive and somatic features, rather than the maladaptive safety and avoidance behaviours (American Psychiatric Association, 2022). This is likely due to the current nosology within the DSM-5, which does not include behavioural components for GAD, despite behavioural avoidance being included in all other DSM-5 anxiety disorder classifications. Recent research has found that individuals with GAD use maladaptive behaviours, consisting of safety behaviours (checking, watching, planning, reassurance seeking and attempts to control others) and avoidance (avoidance of decision making, situations, people, and activities) (Mahoney et al., 2016, 2018). Thus, it is hypothesised that behavioural experiments instigate change through expectancy violation as well as habituation.

Considering this research, the current study aimed to assess the efficacy of a brief, videoconference intervention using behavioural experiments targeting IU for individuals with GAD. The study consists of a mixed factorial design, with 2 conditions (treatment and waitlist) x 2 time-points (pre-treatment and post-treatment). Forty participants were randomised to receive either 1) the treatment condition, consisting of a two-session behavioural experiment intervention, or alternatively 2) the waitlist condition, involving a two-week waitlist. Participants in the waitlist condition were offered the two-session behavioural experiment intervention following waitlist period.

It was hypothesised that participants would find the intervention acceptable, feasible, and appropriate. Regarding clinical outcomes, it was predicted that the combined treatment group (i.e., including waitlist participants once they received treatment) would show significant within-group reductions from pre- to post-treatment across all GAD-related outcomes (i.e., worry, safety behaviours, avoidance, depression, anxiety, physiological tension, and IU), with medium effect sizes. In addition, it was predicted that both worry, and IU would demonstrate clinically significant and reliable change for the combined group. In addition, in was predicted that between-groups analysis would show greater reduction for the treatment group in all GAD outcomes at post-treatment compared to waitlist. A medium within-group effect was expected for the treatment group, with no pre- to post-treatment change anticipated for the waitlist group.

Additionally, exposure-based processes were also explored. It was hypothesised that expectancy violation would occur, indexed by a significant reduction in the strength of target negative beliefs about uncertainty from pre- to post-experiment. It was hypothesised that concern regarding uncertain and negative situations would decrease from pre- to post-treatment, with no change in concern for positive situations. Further, it was hypothesised that habituation would occur, with subjective units of distress (SUDs) higher in anticipation of the behavioural experiment than during or after completing the behavioural experiment if they were to re-complete the experiment again (i.e., re-complete).

2. Methods

2.1. Participants

The University of Technology Sydney (UTS) Human Ethics Research Committee (HREC) approved all aspects of this research (HREC: ETH22–7185). The trial was registered on the Australian New Zealand Trials Registry under identifier [ACTRN12623000262606].

Participants were 40 individuals who met criteria for DSM-5-TR GAD (American Psychiatric Association, 2022). All participants were administered the Diagnostic Interview for Anxiety, Mood, and OCD Related Neuropsychatric Disorders (DIAMOND) version 1.2 (Tolin et al., 2018) by a registered psychologist in the clinical psychology registrar programme or a provisionally registered psychologist in a clinical psychology training program under the supervision of a senior clinical psychologist. Each diagnosis was provided with a clinician severity rating (CSR) based on distress and functional impairment in the past month, rated on a 7-point Likert scale ranging from 1 (Normal) to 7 (Extreme). The mean CSR was 5.25 (SD = 0.74). Participants demonstrated high rates of comorbidity, with 70% having at least one comorbid secondary diagnosis: including social anxiety disorder (20%), major depressive disorder or persistent depressive disorder (12.5%), adjustment disorder (7.5%), OCD (7.5%), attention deficit hyperactivity

disorder (5%), panic disorder (5%), specific phobia (2.5%), trichotillomania or excoriation disorder (5%), post-traumatic stress disorder (2.5%), and other specific eating disorder (2.5%).

Most participants (62.5%) were not taking medication for their mental health prescribed by a medical professional. For participants that were taking medications, these included selective serotonin reuptake inhibitors (15%), serotonin-norepinephrine reuptake inhibitors (7.5%), benzodiazepines (5%), stimulants (5%), tricyclic anti-depressants (2.5%), beta-antagonist (2.5%). In combination with the medications previously detailed participants were also prescribed tetrahydrocannabinol (THC: 5%), and an alpha agonist (2.5%).

Most participants identified as being from Oceania, were married or in a de facto relationship (45%), were in full-time employment (30%) and had completed an undergraduate (32.5%) or postgraduate degree (32.5%). See Table 1 for demographics for the combined sample, as well as separate groups. There was no significant difference in age between the treatment and waitlist group, t (38) = 0.61, p = 0.54. Chi-square tests revealed that there were no significant differences across groups for gender, χ^2 (2, N = 40) = 3.94, p = 0.14; relationship status, χ^2 (4, N = 40) = 2.28, p = 0.81; education, χ^2 (3, N = 40) = 0.65, p = .88; ethnicity, χ^2 (6, N = 40) = 5.44, p = 0.49; employment, χ^2 (6, N = 40) = 4.61, p = 0.60; taking of medication for mental health, χ^2 (6, N = 40) = 4.69, p = 0.58, or co-morbidity, χ^2 (12, N = 40) = 8.84, p = 0.64. A t test demonstrated no significant difference for the CSR between the immediate (M = 5.35, SD =.59) and waitlist group (M = 5.15, SD =.88) t (38) = .85, p = 0.20 at pre-treatment.

Table 1
Demographic characteristics for the immediate treatment (n = 20), waitlist (n = 20), and combined group (N = 40).

Demographics	Immediate Tx condition	Waitlist condition	Combined
Age			
Mean age in years (SD)	42.05 (16.36)	45.15 (15.57)	43.60 (15.86)
Age range in years	20–71	19–78	19–78
Gender – Frequency (%)			
Female	75	90	82.5
Male	25	5	15
Non-binary	0	5	2.5
Ethnicity – Frequency (%)			
Oceanian (e.g., Aboriginal, Australian, Maori, New Zealand)	50	50	50
North-East Asian (e.g., Chinese, Japanese, Korean)	25	20	22.5
South-East Asian (e.g., Thai, Vietnamese)	15	5	10
North African and Middle Eastern	5	10	7.5
Southern or Central Asia (e.g., Indian)	0	10	5
Americas (e.g., North, Central, and South American)	0	5	2.5
European (e.g., European, British, Irish)	5	0	2.5
Sub-Saharan African (e.g., Central and West African)	0	0	0
Relationship Status – Frequency (%)			
Married or De Facto	40	50	45
Single	35	35	35
Committed Relationship	15	15	15
Dating	5	0	2.5
Separated/Divorced	5	0	2.5
Employment – Frequency (%)			
Full-Time employment	40	20	30
Part-time employment	15	35	25
Full time student	15	20	17.5
Retired	10	15	12.5
Contractually employed	10	5	7.5
Casually employed	5	5	5
Home duties	5	0	2.5
Education* – Frequency (%)			
Post-Graduate degree	30	35	32.5
Undergraduate degree	30	35	32.5
Certificate/Technical and Further Education (TAFE)	25	15	20
High School	15	15	15

Note. * Highest attainment of education

2.2. Procedure

2.2.1. Inclusion criteria

Participants met inclusion criteria if they (1) had a principal diagnosis of GAD; (2) were at least 18 years or older; (3) living in Australia; (4) fluent in English; (5) had access to a computer-based device with stable internet connection; (6) no evidence of active suicidal intent (based on clinical judgement); (7) no evidence of active psychotic symptoms (based on DIAMOND self-report screener and clinical judgement); (8) no concurrent engagement in psychological treatment during the trial; and (9) if taking medication for their mental health, able to make a commitment to keep medication dose stable for the duration of the study.

2.2.2. Recruitment

Participants were recruited from The University of Sydney Psychology Clinic, through flyers placed around The University of Sydney, Australian National University, and UTS campuses, as well as online advertisements on the Meta platform Facebook between October 2023 and July 2024. See Fig. 1 for flow diagram.

2.2.3. Telephone screening interview

Participants who signed up for the study were contacted via telephone by a psychologist. The interview was approximately 15-minutes and the psychologist explained what would be involved in the study, as well as assessing if the participant met inclusion criteria. Participants where the study did not appear suitable were provided with referral options and appropriate follow-up.

2.2.4. Pre-treatment assessment session

All sessions were conducted via Zoom. Informed consent was obtained and if a primary diagnosis of GAD was indicated by the diagnostic interview, participants were accepted into the study and completed measures. Participants were told at the end of the assessment session whether they were randomised to treatment or waitlist condition.

2.2.5. Randomisation

Participants were randomised to the condition with a block randomisation method (block size = 2) using a computer-generated randomisation procedure.

2.2.6. Experimental conditions

Participants randomised to the treatment condition were offered the treatment immediately. Treatment consisted of two individual weekly 60-minute sessions, followed by a 30-minute post-treatment assessment one week after completing treatment. Participants randomised to the waitlist condition completed a 2-week waitlist period. After the waitlist period, participants re-completed the questionnaires and were offered the treatment with another post-treatment assessment one week after completing the treatment.

2.2.7. Treatment

Treatment was based on the behavioural experiment protocol developed to directly target IU, by Robichaud et al. (2019). In the first individual treatment session, the psychologist oriented the participant to a CBT model of GAD (adapted from Robichaud et al., 2019, p. 165) and introduced the concept of IU. The psychologist used Socratic questioning to explore the function of worry (i.e., mental attempt to reduce

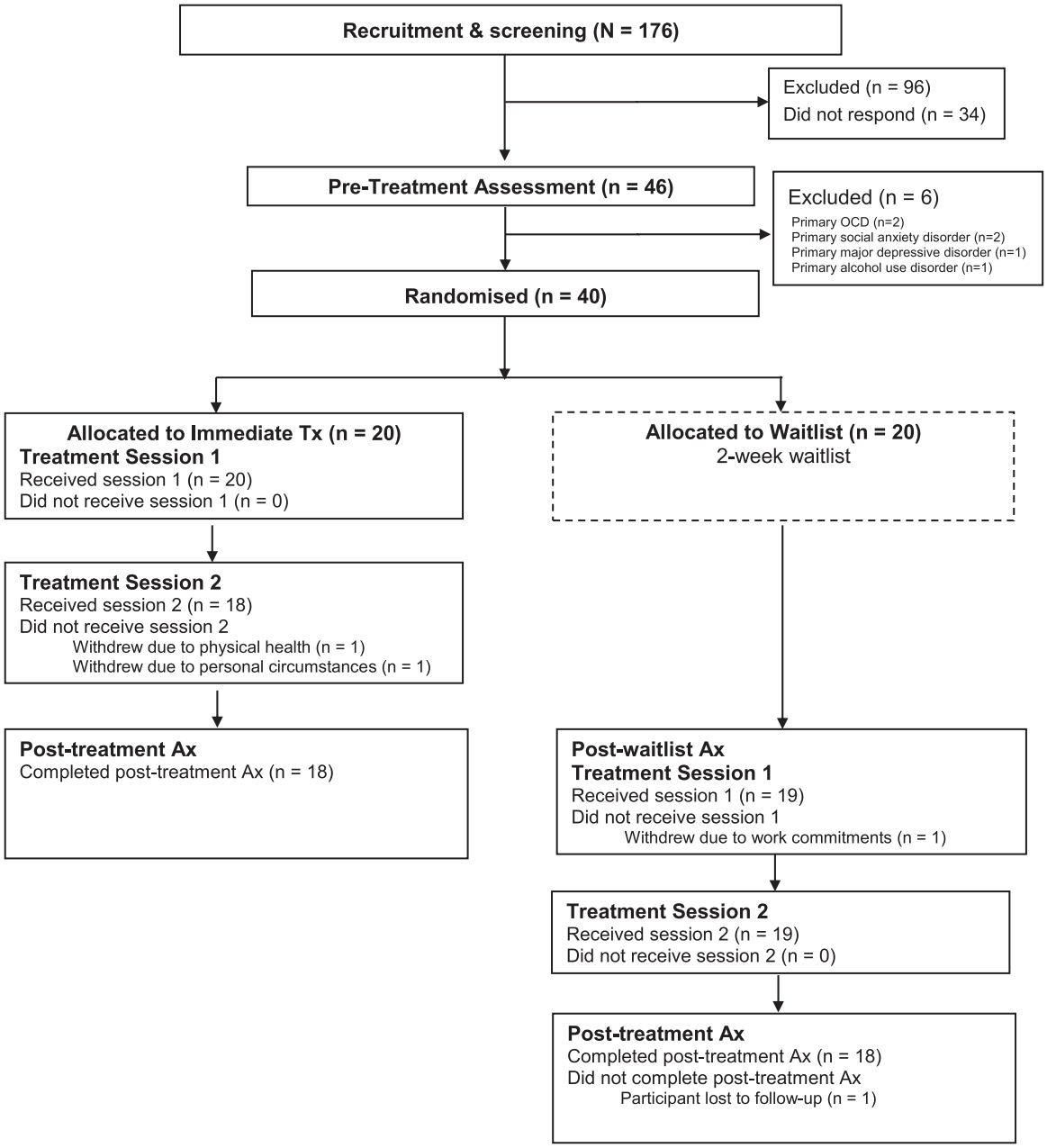

Fig. 1. Flow Diagram of Participant Recruitment, Randomisation, and Intervention Uptake.

uncertainty) with participants, and to provide psychoeducation that worry, anxiety, and safety behaviours/avoidance are all normal reactions to uncertainty. Participants then identified their specific types of safety behaviours (e.g., reassurance seeking, excessive re-checking, excessive information gathering, second guessing decisions) and avoidance (e.g., avoidance of situations and/or feelings, procrastinating). Socratic questioning was used to further understand the function of safety behaviours and avoidance (i.e., physical attempt to reduce uncertainty), as well as their short-term consequences (e.g., temporary anxiety reduction) and long-term consequences (e.g., interferes with new learning and prevents any violation of expectancy, thereby maintaining negative beliefs about uncertainty). The psychologist and participant collaboratively set-up one behavioural experiment to be attempted over the next week before the second treatment session. If time allowed, up to an additional two behavioural experiments could be set up with the participant (to a maximum of three). When setting up a behavioural experiment participants would first identify a situation that they are engaging in avoidance or using a safety behaviour (i.e., *I can't try a new restaurant, without excessive planning*). Participants were guided to identify a negative belief about uncertainty related to this situation (*i.e., If there's uncertainty, then something will always go wrong*) and then rate how strongly they think this belief is true from 0 to 100, as well as how distressed they felt in anticipation of attempting the behavioural experiment (i.e., SUDs anticipation).

The second individual session re-oriented the participant to the content covered in the first session and de-briefed using Socratic questioning regarding the outcome of the behavioural experiment. Debriefing included 1) whether the actual outcome matched the feared prediction, 2) whether the actual outcome was positive, neutral, or negative, 3) if the outcome was negative, A) was the outcome as negative as they anticipated (i.e., catastrophic) and B) how did they perceive their ability to cope with the negative outcome. Participants were asked to rate how distressed they felt during the experiment (i.e., SUDs during) and how distressed they would feel about re-completing the experiment

again (i.e., SUDs re-complete). In addition, participants were also asked to re-rate how strongly they think the negative belief about uncertainty (i.e., *If there is uncertainty, then something will always go wrong*) is true, and attempt to develop a reappraisal statement (i.e., *I can cope, even when there is uncertainty and things don't go to plan*).

2.2.8. Therapists

Treatment was delivered either by a registered psychologist in a clinical psychology registrar program or a provisionally registered psychologist in a clinical psychology training program under the supervision of a senior clinical psychologist. The same psychologist that conducted the pre-treatment assessment session, completed all treatment sessions and the post-treatment session with the participant. The therapist was therefore not blind to the treatment condition.

2.3. Measures

2.3.1. Diagnostic interview

2.3.1.1. Diagnostic interview for anxiety, mood, and OCD and related neuropsychiatric disorders (DIAMOND). The DIAMOND is a structured diagnostic interview based on DSM-5 criteria for assessment of a range of mental health disorders. Both psychologists that conducted the DIAMOND interviews completed the DIAMOND Training Course and Rater Certification through the Anxiety Disorders Centre at The Institute of Living, with training exam kappa coefficients of 0.84 and 0.76, indicating excellent and very good reliability. All interviews were video-recorded and 15% were randomly selected for coding. The independent rater was a registered psychologist and clinical psychology registrar and was blind to the diagnostic status of participants, showing 100% agreement, $\kappa = 1.0$.

2.3.2. GAD symptom and process outcome measures

2.3.2.1. Penn State worry questionnaire-abbreviated (PSWQ-A). The PSWQ-A is an 8-item measure of excessive worry (Hopko et al., 2003). The PSWQ-A has been found to have a stable one-factor structure as well as good construct validity and internal consistency in adults with GAD (Wilson et al., 2025). The Cronbach's α was adequate, $\alpha = 0.77$.

2.3.2.2. Intolerance of uncertainty scale −12 (IUS-12). IU was measured using the IUS-12, a 12-item questionnaire (Carleton et al., 2007; Freeston et al., 1994). The IUS-12 has demonstrated good construct validity and internal consistency in adults with GAD (Wilson et al., 2020). Cronbach's $\alpha = 0.91$ was excellent for the current study.

2.3.2.3. Depression anxiety stress scales – short form (DASS-21). The DASS-21 assess symptoms of depression, anxiety, and stress (Lovibond & Lovibond, 1995). The DASS-21 demonstrates a stable three-factor structure with adequate construct validity and internal consistency in a variety of clinical samples (Brown et al., 1997; Page et al., 2007). Internal consistency was good for depression (α=0.85) anxiety (α= 0.80) and stress (α= 0.85).

2.3.2.4. Worry behaviours inventory (WBI). The 10-item WBI aims to evaluate behaviours associated with GAD with two subscales: safety behaviours (7 items) and avoidance (3 items) (Mahoney et al., 2016). The WBI is rated on a 5-point Likert scale (ranging 0–4). Both subscales have good convergent validity as well as treatment sensitivity in adults with anxiety and/or depression (Mahoney et al., 2018). Internal consistency was good for safety behaviours ($\alpha = 0.88$) and avoidance ($\alpha = 0.81$) subscales.

2.3.3. Implementation outcome measures

Acceptability of Intervention Measure (AIM), the Intervention Appropriateness Measure (IAM), and the Feasibility of Intervention Measure (FIM) were used to assess implementation of treatment (Weiner et al., 2017). The AIM, IAM, and FIM each have four items rated on a 5-point scale. Each measure has a stable one-factor structure with adequate test re-test reliability (Weiner et al., 2017). A total score ≥ 16 on the AIM, IAM, and FIM were used as an indicator for effective implementation. Cronbach's α was good for the AIM ($\alpha = 0.88$) and excellent for the IAM ($\alpha = 0.96$) and FIM ($\alpha = 0.91$).

2.3.4. Treatment fidelity

All treatment sessions were video recorded to check treatment fidelity. Consistent with established practices in treatment fidelity, whereby 20–25% of sessions are commonly coded to assess adherence (McManus et al., 2012; Schoenwald & Garland, 2013), fidelity ratings were conducted on 30% of participants ($N = 12$) in the present study. An independent rater (who also completed the DIAMOND ratings) viewed all treatment sessions for 12 participants (30%). The independent rater was a registered psychologist and clinical psychology registrar, and used a checklist of 9-items that concerned key components from both intervention sessions (e.g., introduced IU, explored how worry behaviours maintain IU, setting up behavioural experiment, debriefing about the outcome). The checklist was dichotomous (Y/N), yielding a total score from 0 to 9.

2.3.5. Behavioural experiment variables

2.3.5.1. Amount completed. The number of behavioural experiments each participant completed was reported with between 1 and 3 behavioural experiments completed per participant.

2.3.5.2. Negative belief about uncertainty. Participants were asked to identify a specific negative belief about uncertainty to be tested in each behavioural experiment. Participants were asked to rate how strongly they think this belief is true from 0 (not at all true) to 100 (completely true) before completing the behavioural experiment (i.e., pre-experiment) as well as after completing the behavioural experiment (i.e., post-experiment).

2.3.5.3. Subjective units of distress (SUDs). Participants were asked to rate their level of distress from 0 (not at all distressed) to 100 (extremely distressed) before the experiment (i.e., anticipation), during behavioural experiment, and after the behavioural experiment if they were to re-complete the experiment again (i.e., re-complete).

2.3.5.4. The ambiguous/unambiguous situations diary (AUSD). The AUSD measures concern for ambiguous and unambiguous situations (Davey et al., 1992). The AUSD consists of fictitious diary entries for 28 consecutive days. There are 14 ambiguously framed entries and 14 unambiguously framed entries - with seven positively and seven negatively framed. Participants were asked to read each entry and decided whether they would be 'concerned' or 'unconcerned'. Total scores for ambiguous entries ranged from 0 to 14, with positive and negative entries ranging from 0 to 7, with higher scores indicating greater concern.

2.4. Data analysis

Statistical analyses were conducted using Statistical Package for the Social Sciences (SPSS) software 28.0.0.0. To determine if there were any differences across groups for demographic variables, an independent samples *t*-test and chi-square tests were conducted.

2.4.1. GAD symptom and process measures

A within-group linear mixed model (LMM) was estimated for each symptom and process measure using the combined treatment group to examine the effect of time (pre-treatment vs post-treatment). For each

analysis, time was entered as a fixed variable. Within-group Cohen's *d* effect sizes were also calculated for all measures, with effect sizes interpreted as small ($d = 0.20$), medium ($d = 0.50$) and large ($d = 0.80$) (Cohen, 1977). Clinical significance as well as a reliable change index (RCI) were calculated for worry and IU for the combined treatment group following guidelines by Jacobson and Truax (1991). Change was determined to be reliable if the RCI was less than −1.96 (Jacobson & Truax, 1991).

A between-group LMM was estimated for each symptom and process measure to examine the effects of condition (treatment vs waitlist), time-point (pre-treatment vs post-treatment) and their interaction (condition × time-point). Fixed effects included condition, time-point, and interaction. A between-group MANOVA was utilised to compare pre-treatment scores for outcome measures of worry, IU, safety behaviours, avoidance, depression, anxiety, and stress, and was found to have a significant effect $F(7, 69) = 2.72$, $p = .02$. Inspection indicated that there was a significant difference between the groups in avoidance $F(1, 69) = 7.41$, $p < .001$. Avoidance was therefore included as a covariate due to a significant pre-treatment difference between the groups. Within-group and between-group Cohen's *d* effect sizes were also calculated for all measures.

Missing data were handled using the LMMs, a robust method commonly used in longitudinal research. LMMs account for the hierarchical structure of repeated measures and can accommodate unequal numbers of observations across participants. They also provide unbiased estimates when data are missing at random (Huta, 2014). For each LMM, a restricted maximum likelihood estimation method was used, with an unstructured covariance structure.

2.4.2. Behavioural experiment variables

2.4.2.1. Negative beliefs about uncertainty. Two separate repeated-measures ANOVAs were conducted to determine if there were any significant differences for the pre-experiment belief ratings (i.e., first, second, and third experiments) and the post-experiment belief ratings (i.e., first, second and third experiments). Provided no differences were found between the pre-experiment ratings or post-experiment ratings, an average belief rating was calculated. Subsequently, a paired samples *t*-test was performed to determine if the mean pre-experiment belief rating was significantly higher than the mean post-experiment belief rating.

2.4.2.2. SUDs. Three separate repeated-measures ANOVAs were conducted to determine if there were any significant differences in SUDs ratings preceding the three experiments (i.e., SUDs anticipation), during the three experiments (i.e., SUDs during), and after the three experiments if they were to have to re-complete the experiment (i.e., SUDs re-complete). On the provision that there were no significant differences, an average was calculated for each of the SUDs ratings (i.e., mean SUDs anticipation, mean SUDs during, mean SUDs re-complete). Subsequently, a repeated-measures ANOVA was conducted to determine if there were any differences between mean SUDs ratings.

2.4.2.3. The ambiguous/unambiguous situations diary (AUSD). To determine if the intervention shifted appraisals, three paired *t*-tests (i.e., positive, negative and uncertain situations) were conducted for the combined treatment group.

2.5. Power analysis

G*Power version 3.1.9.7 (Faul et al., 2007) was used to conduct an *a priori* power analysis to calculate adequate sample size to test the within-between group interaction effect. Previous research has found a medium effect of IU for a one-session IU focused psychoeducation session (Shapiro et al., 2023). The study was powered to detect a medium effect (Cohen's $f = .25$), using two assessment time-points (i.e., pre-treatment vs post-treatment) and two groups (i.e., immediate treatment vs waitlist), as well as specifying a power of.80 and an alpha.05. Power analysis determined that a sample size of 34 was required to detect a significant within-between groups interaction effect. The total study sample size was 40 participants.

3. Results

3.1. GAD outcome measures – combined group

Results revealed that the majority of GAD symptom and process outcome measures significantly decreased over time for the combined treatment group. This included worry ($B = 2.71$, $SE = 0.77$, $p < .01$, 95% CI [1.15, 4.27]), IU ($B = 3.09$, $SE = 1.12$, $p < .01$, 95% CI [.81, 5.36]), safety behaviours ($B = 1.88$, $SE = 0.72$, $p < .01$, 95% CI [.42, 3.35]), avoidance ($B = 0.90$, $SE = 0.34$, $p < .01$, 95% CI [0.21, 1.59]), physiological tension/stress ($B = 1.33$, $SE = 0.58$, $p = .03$, 95% CI [.15, 2.51]) and depression ($B = 1.26$, $SE = .58$, $p = .04$, 95% CI [.09, 2.44]). Anxiety did not significantly decrease over time ($B = 0.77$, $SE = 0.50$, $p = 0.14$, 95% CI [-0.25,1.78]). See Table 2 for means, standard deviations and within-group Cohen's *d* effect sizes for the combined treatment group.

3.2. Clinical significance and RCI

In the combined treatment group, 27.5% of participants showed clinically significant change in worry and 80.5% demonstrated a clinically significant change for IU. Reliable improvement was observed in 50% for worry and 80.5% for IU (RCI < −1.96).

3.3. GAD outcome measures – treatment vs waitlist condition

See Table 3 for within-group effect sizes. All main effects for condition and time were non-significant. Results revealed a significant interaction between condition (treatment vs waitlist) x time (pre-treatment vs post-treatment) for worry ($B = 3.09$, $SE = 1.24$, $p = 0.02$, 95% CI [.57, 5.61]). There was a medium between-group effect ($d = 0.54$) at post-treatment, in favour of the treatment condition. All other GAD-related outcome variables had non-significant interaction effects: IU ($B = 3.61$, $SE = 1.87$, $p = 0.06$, 95% CI [-.022, 7.43]), safety behaviours ($B = 1.74$, $SE = 1.30$, $p = 0.19$, 95% CI [-0.89, 4.37]), avoidance ($B = 1.15$, $SE = 0.66$, $p = 0.09$, 95% CI [-0.19, 2.49]), physiological tension ($B = 0.46$, $SE = 1.10$, $p = 0.68$, 95% CI [-1.77, 2.68]), anxiety ($B = -0.18$, $SE = 1.05$, $p = 0.86$, 95% CI [-2.31,1.95]), and depression ($B = -0.15$, $SE = 1.17$, $p = 0.90$, 95% CI [-2.52, 2.22]). Between-groups effects at post-treatment ranged from no effect for IU ($d = 0.06$) and depression ($d = 0.06$), to a small effect for safety behaviours ($d = 0.34$), avoidance ($d = 0.34$), anxiety ($d = 0.34$) and physiological tension ($d = 0.24$).

Table 2
Within-group effect sizes for the combined treatment group.

Measure	Pre-treatment mean (SD) (N = 39)	Post-treatment mean (SD) (N = 38)	Cohen's *d*
PSWQ-A	33.59 (3.96)	30.56 (5.50)	0.63
IUS-12	38.69 (9.79)	35.55 (8.66)	0.34
WBI safety behaviours	17.87 (6.05)	15.94 (5.86)	0.32
WBI avoidance	7.13 (2.77)	6.11 (2.46)	0.39
DASS-21 depression	5.82 (3.82)	4.55 (4.00)	0.32
DASS-21 anxiety	5.41 (3.37)	4.64 (3.54)	0.22
DASS-21 stress	10.03 (3.92)	8.61 (4.08)	0.35

Note. *PSWQ-A* Penn State Worry Questionnaire – Abbreviated, *IUS – 12* Intolerance of Uncertainty Scale – 12, *WBI* Worry Behaviours Inventory, *DASS – 21* Depression Anxiety Stress Scales – 21, *CSR* Clinician severity rating.

Table 3
Within-group effect sizes for the immediate treatment and waitlist condition at pre-treatment and post-treatment.

Condition	Measure	Pre-treatment mean (SD) $N = 20$	Post-treatment mean (SD) $N = 18$	Cohen's *d*
Immediate Treatment	PSWQ-A	34.30 (3.25)	30.90 (5.91)	0.71
	IUS-12	40.20 (10.36)	36.61 (9.26)	0.37
	WBI safety behaviours	17.90 (6.29)	15.84 (5.92)	0.34
	WBI avoidance	8.30 (2.70)	6.95 (2.55)	0.51
	DASS-21 depression	6.70 (4.48)	5.18 (4.25)	0.35
	DASS-21 anxiety	5.35 (4.07)	4.46 (3.42)	0.24
	DASS-21 stress	10.25 (3.86)	8.84 (4.00)	0.36
Condition	**Measure**	**Pre-Treatment Mean (SD) *N*=20**	**Post-Treatment Mean (SD) N = 19**	**Cohen's *d***
Waitlist	PSWQ-A	33.10 (4.82)	32.79 (4.37)	0.07
	IUS-12	37.20 (10.05)	37.22 (9.54)	0.00
	WBI safety behaviours	18.15 (6.09)	17.84 (5.87)	0.05
	WBI avoidance	6.15 (2.53)	6.00 (2.33)	0.06
	DASS-21 depression	6.65 (4.31)	4.98 (2.87)	0.45
	DASS-21 anxiety	6.70 (4.07)	5.63 (3.46)	0.28
	DASS-21 stress	10.85 (4.81)	9.90 (4.94)	0.19

Note. *PSWQ-A* Penn State Worry Questionnaire – Abbreviated, *IUS – 12* Intolerance of Uncertainty Scale – 12, *WBI* Worry Behaviours Inventory, *DASS – 21* Depression Anxiety Stress Scale – 21.

3.4. Treatment Fidelity

Scores on the behavioural experiment adherence checklist indicated high treatment integrity, with 100% adherence for both psychologists that administered the intervention.

3.5. Implementation Outcome Measures

Implementation measures suggested that participants found the intervention acceptable (AIM: $M = 17.5$, SD = 1.97), feasible, (FIM: M = 17.61, SD = 1.97) and appropriate (IAM: M = 16.75, SD = 1.91).

3.6. Behavioural Experiment Variables

3.6.1. Number of Behavioural Experiments Completed

All participants ($N = 37$) who finished treatment completed at least two behavioural experiments (100%). Of these, a large percentage (81%) completed three total behavioural experiments ($N = 30$).

3.6.2. Negative Beliefs about Uncertainty

There were no significant differences between pre-experiment belief ratings for the first ($M = 77.68$, $SD = 17.18$), second ($M = 73.51$, $SD = 17.98$) and third behavioural experiments ($M = 75.33$, $SD = 16.29$), $F(2, 58) = .49$, $p = .61$. Similarly, there were no significant differences between post-experiment belief ratings for the first ($M = 49.92$, $SD = 24.34$), second ($M = 49.27$, $SD = 26.03$) and third experiment ($M = 52.43$, $SD = 21.73$), $F(2, 58) = .28$, $p = .76$. Average pre-experiment belief ratings ($M = 75.84$, $SD = 11.85$) were significantly higher than average post-experiment belief ratings ($M = 50.53$ SD = 16.33), $t(36) = 7.50$, $p < .001$. The mean difference was -25.31 (95% CI [-19.18, -33.07]), with a large within-group effect size ($d = 1.77$).

3.6.3. SUDs

There were no significant differences between SUDs ratings in anticipation of the first ($M = 56.87$, $SD = 26.40$), second ($M = 60.77$, $SD = 21.12$), and third ($M = 65.50$, $SD = 23.06$) experiment, $F(2, 58) = 2.44$, $p = .10$. There were no significant differences between SUDs ratings during the first ($M = 48.63$, $SD = 25.90$), second ($M = 53.67$, $SD = 23.33$), and third ($M = 48.10$, $SD = 28.51$) behavioural experiment, $F(2, 58) = .61$, $p = .55$. No significant differences were found for SUDs ratings after the first ($M = 34.90$, SD = 23.77), second ($M = 37.52$, $SD = 26.05$), and third ($M = 46.14$, $SD = 24.34$) behavioural experiment, $F(2, 58) = 2.92$, $p = .06$. Mean SUDs anticipation (60.58, $SD = 21.13$), mean SUDs during (49.36, $SD = 18.97$), and mean SUDs to re-complete (38.47, $SD = 19.36$) were significantly different $F(2, 72) = 29.42$, $p < 0.001$. Pairwise comparisons demonstrated a significant reduction in SUDs from anticipation of experiment to during the experiment (mean difference was -11.21, $p < 0.001$. 95% CI [-5.04, -17.38]) as well as from anticipation of experiment to if they were to re-complete the experiment (mean difference was -22.10, $p < 0.001$. 95% CI [-15.31, -28.90]). There was also a significant reduction for SUDs rated during the experiment to if they were to re-complete (mean difference was -10.89, $p < 0.001$. 95% CI [-15.17, -6.62]).

3.6.4. Ambiguous/Unambiguous Situations Diary (AUSD)

Appraisals of concern for ambiguous scenarios demonstrated a significant reduction from pre-treatment ($M = 8.52$, $SD = 2.13$) to post-treatment ($M = 7.19$, $SD = 2.52$), $t(35) = 3.65$, $p < 0.001$.There was no significant difference between appraisals of concern for positive scenarios at pre-treatment ($M = 0.86$, $SD = 1.02$) and at post-treatment ($M = 0.64$, $SD = 0.68$), $t(35) = 1.28$, $p = 0.21$. Similarly, there were no significant differences between appraisals of concern for negative scenarios at pre-treatment ($M = 6.03$, $SD = 1.18$) and at post-treatment ($M = 6.03$, $SD = 1.16$), $t(35) = 0.00$, $p = 0.99$.

4. Discussion

The current study supports the use of behavioural experiments in the treatment of GAD. Specifically, the brief two-session intervention delivered via videoconference significantly reduced most GAD-related symptoms and processes, with the exception of anxiety for the combined treatment group. The study also sought to understand the processes underlying exposure, and found evidence of belief change as well as habituation. In addition, treatment was also found to be acceptable, feasible, and appropriate by adults with GAD.

Findings for the combined treatment group revealed a significant decrease from pre- to post-treatment across worry, IU, safety behaviours, avoidance, physiological tension, and depression, in accordance with hypotheses. For the combined treatment group, worry demonstrated a medium within-group effect, with all remaining measures demonstrating a small within-group effect from pre- to post-treatment. Contrary to hypotheses, anxiety did not demonstrate a significant reduction from pre- to post-treatment for the combined treatment group after two-sessions. Previous research has found that 12-sessions utilising behavioural experiments has led to a significant reduction in anxiety (Dugas et al., 2022), suggesting that additional sessions are likely necessary to directly target autonomic bodily sensations and situational anxiety for individuals with GAD.

In addition, a proportion of participants in the combined treatment group demonstrated clinically significant change for worry (27.5%) and intolerance of uncertainty (80.5%). Reliable improvement was also observed in 50.0% of participants for worry and 80.5% for intolerance of uncertainty. Together, these findings indicate that although not all participants achieved clinically significant change, a substantial proportion showed meaningful and reliable reductions in worry and intolerance of uncertainty following the brief intervention, highlighting its potential utility in targeting a core mechanism of GAD.

To evaluate the efficacy of the intervention while controlling

potential confounding factors, a RCT design was used to examine whether participants in the treatment condition demonstrated greater reduction in outcomes compared to those in the waitlist. Worry demonstrated a significantly greater reduction from pre- to post-treatment for the treatment group in comparison to the waitlist group. In terms of within-group effect sizes from pre- to post-treatment, the immediate treatment group evidenced medium effect ($d > 0.50$), whereas the waitlist group showed no effect ($d < 0.20$). Furthermore, when comparing the post-treatment scores there was a medium between-group effect in favour of the immediate treatment group. In contrast, there was no significant differences between the treatment and waitlist conditions at post-treatment for the remaining measures: IU, safety behaviours, avoidance, physiological tension, anxiety, and depression. Interestingly, avoidance had a medium ($d > 0.50$) within-group effect for the immediate treatment condition, and the waitlist group had no effect ($d < 0.20$), there was still no significant interaction at post-treatment. Though this trend appears to be in favour of the treatment, one explanation pertains to the elevated pre-treatment avoidance scores for the immediate treatment group. For the remaining measures, the non-significant interaction may be due to initial assumptions about the effect size, as the study was powered for a medium effect size, however, the true effect size found for most of the immediate treatment measures was small. For example, IU, safety behaviours, and physiological tension had a small within-group effect ($d > 0.20$) from pre- to post-treatment for the immediate treatment condition, and despite the waitlist group having no effect ($d < 0.20$) there was no significant difference at post-treatment between the groups. Thus, given the study was adequately powered for a between-group effect, it may be the very design of the study that meant the effect was not detected. Two potential reasons that an effect was not detected could be in relation to treatment dose (i.e., too few sessions) or omission of follow-up time-point (i.e., not enough time to detect changes). With regard to the latter reason, ultra-brief interventions (e.g., one session) in anxiety and depression have found no difference between treatment and waitlist at 1-week post-baseline or at 5-weeks post-baseline, but instead find a significant difference at 9-weeks post-baseline (Bisby et al., 2024). Furthermore, Bisby et al. (2024) found no significant difference between the ultra-brief intervention (one week) and standard treatment (eight weeks) at the same 9-weeks post-baseline time point, suggesting that ultra-brief interventions can be as effective as standard treatment. Thus, it would be interesting in future research to see if the brief two-session behavioural experiment intervention demonstrates a significant interaction (in favour of the treatment) at an extended follow-up time-point (e.g., 12-week post-baseline) when compared to a waitlist. It is also possible that the observed effects diminish over time rather than strengthen. Previous research has demonstrated that a standard 12-week intervention targeting IU in individuals with GAD, which incorporated behavioural experiments, either maintained treatment gains or continued to reduce symptom severity at 6- and 12-month follow-up assessments (Dugas et al., 2022). Accordingly, further research is warranted to elucidate whether the effects of the brief behavioural experiment intervention are sustained, attenuated, or amplified over extended periods.

Implementation measures indicated that participants rated the intervention as acceptable, feasible, and appropriate, suggesting strong engagement and practicality in a brief videoconference format. In addition, a systematic review estimated mean dropout rate for CBT for adults with GAD to be 16.99% (Gersh et al., 2017). The current study operationalised drop-out as not completing the post-treatment assessment session, thus the study had a relatively low drop-out rate of 10% ($n = 4$). Furthermore, 100% of participants completed at least one behavioural experiment as per the protocol. Together these findings suggest that the intervention were well received by adults with GAD.

The study also aimed to better understand the processes of exposure-based learning hypothesised to underlie symptom change in behavioural experiments. The study found that behavioural experiments significantly reduced idiosyncratic negative beliefs about uncertainty, evidenced by large effect size. These results suggest that there was a mismatch between expectancy and outcome, suggesting that inhibitory learning occurred. In addition to idiosyncratic beliefs about uncertainty, *concern* regarding uncertain situations significantly reduced at post-treatment, whereas concern for positive situations had no change over treatment. Contrary to hypotheses, concern regarding negative situations did not change. One potential explanation for this finding is that individuals with GAD may maintain concern about negative situations to avoid sudden shifts from neutral or positive emotions to distress (Llera & Newman, 2017). By keeping themselves in a state of anticipatory worry, its hypothesised to reduce the impact of unexpected negative outcomes, consistent with the contrast avoidance model (Llera & Newman, 2017). From a contrast avoidance perspective, sustaining concern protects against shifts from neutral or positive states to negative affect, making behavioural experiments targeting uncertainty potentially less effective for concern about negative situations. Furthermore, evidence of habituation occurred with SUDs ratings significantly higher on average in anticipation of the experiment than 1) during and 2) after the experiment if they were to re-complete the experiment (i.e., SUDs re-complete). Thus, evidence emerged for both inhibitory learning and habituation as mechanisms of change during exposure. Together these findings suggest that exposure operates through multiple, complementary pathways.

The present study supports the clinical utility of a brief, two-session videoconference intervention incorporating behavioural experiments targeting intolerance of uncertainty. Given the high prevalence of GAD and common barriers to accessing traditional CBT (e.g., time, cost, and access), this remotely delivered protocol represents a promising and accessible treatment option. Additionally, behavioural experiments targeting IU may be readily integrated into existing treatment frameworks to reduce core GAD symptoms, including worry and maladaptive safety behaviours. Additionally, delivery via telehealth enhances reach and flexibility while maintaining acceptability, aligning with stepped-care models as a low-intensity, first-line or adjunctive intervention for GAD.

Several limitations need to also be acknowledged. First, the absence of follow-up time-point precluded evaluation of the durability of treatment effects. As noted earlier, symptom improvement may have continued beyond the post-treatment assessment however it is also possible that gains were not maintained or even diminished beyond post-treatment. Future studies should include follow-up assessments, even for brief interventions. Second, the study did not include an active control condition that involved psychological intervention, which limits the ability to account for non-specific treatment effects such as those associated with therapeutic alliance or participant expectations (Cuijpers et al., 2019). Thus, including an active control in future trials would allow for a more rigorous test of the intervention's specific efficacy. Finally, although GAD diagnoses were confirmed at baseline via clinical interview, post-treatment diagnostic interviews were not conducted, restricting evaluation of diagnostic change over time. Future research should incorporate repeated clinical interviews to strengthen outcome validity.

Together these results demonstrate that a two-session protocol can have a significant impact on GAD symptoms and processes, and that behavioural experiments targeting IU are a likely valuable component in the treatment of GAD. Despite the intervention being brief, it was rated as acceptable, feasible and appropriate by participants. Therefore, the findings support the continued exploration and implementation of brief, telehealth treatments for adults with GAD.

Funding

This research was supported by the Commonwealth through an Australian Government Research Training Program Scholarship [DOI: https://doi.org/10.82133/C42F-K220].

CRediT authorship contribution statement

Maree J. Abbott: Writing – review & editing, Supervision, Resources, Methodology, Investigation, Conceptualization. **Emily J. Wilson:** Writing – original draft, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jessica Riley:** Writing – review & editing, Project administration. **Alice R. Norton:** Writing – review & editing, Supervision, Methodology, Conceptualization. **David Berle:** Writing – review & editing, Supervision.

Declaration of Competing Interest

All authors declare they have no conflicts of interest.

Acknowledgements

Thank you to the participants, without you this study could not have been completed. Thank you to Wendy Smith for reviewing the videos for diagnositic validity and treatment fidelity. Thank you also to Associate Professor Kris Rogers for statistical support.

Data availability

The authors do not have permission to share data.

References

American Psychiatric Association. (2022). Anxiety disorders. *Diagnostic and statistical manual of mental disorders—text revision (5th ed.)*. American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787.x05_Anxiety_Disorders

Anderson, K. G., Dugas, M. J., Koerner, N., Radomsky, A. S., Savard, P., & Turcotte, J. (2012). Interpretive style and intolerance of uncertainty in individuals with anxiety disorders: A focus on generalized anxiety disorder. *Journal of Anxiety Disorders, 26*(8), 823–832. https://doi.org/10.1016/j.janxdis.2012.08.003

Basile, V. T., Newton-John, T., & Wootton, B. M. (2024). Treatment histories, barriers, and preferences for individuals with symptoms of generalized anxiety disorder. *Journal of Clinical Psychology, 80*(6), 1286–1305. https://doi.org/10.1002/jclp.23665

Benito, K. G., & Walther, M. (2015). Therapeutic process during exposure: Habituation model. *Journal of Obsessive-Compulsive and Related Disorders, 6*, 147–157. https://doi.org/10.1016/j.jocrd.2015.01.006

Bennett-Levy, J., Butler, G., Fennell, M., Hackmann, A., Mueller, M., & Westbrook, D. (Eds.). (2004). *Behavioural experiments: Historical and conceptual underpinnings*. Oxford University Press. ⟨https://academic.oup.com/book/1220/chapter/140089434⟩.

Bisby, M. A., Balakumar, T., Scott, A. J., Titov, N., & Dear, B. F. (2024). An online therapist-guided ultra-brief treatment for depression and anxiety: A randomized controlled trial. *Psychological Medicine, 54*(5), 902–913. https://doi.org/10.1017/s003329172300260x

Bouchard, S., Dugas, M. J., Belleville, G., Langlois, F., Gosselin, P., Robillard, G., Corno, G., & Marchand, A. (2022). A multisite non-inferiority randomized controlled trial of the efficacy of cognitive-behavior therapy for generalized anxiety disorder delivered by videoconference. *Journal of Clinical Medicine, 11*(19), 5924. https://doi.org/10.3390/jcm11195924

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the depression anxiety stress scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35*(1), 79–89. https://doi.org/10.1016/S0005-7967(96)00068-X

Carleton, R. N., Norton, M. A. P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: A short version of the Intolerance of Uncertainty Scale. *Journal of Anxiety Disorders, 21*(1), 105–117. https://doi.org/10.1016/j.janxdis.2006.03.014

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press, INC.

Coombs, N. C., Meriwether, W. E., Caringi, J., & Newcomer, S. R. (2021). Barriers to healthcare access among U.S. adults with mental health challenges: A population-based study. *SSM - Population Health, 15*, Article 100847. https://doi.org/10.1016/j.ssmph.2021.100847

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5–27. https://doi.org/10.1016/j.brat.2007.10.003

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

Cuijpers, P., Reijnders, M., & Huibers, M. J. (2019). The role of common factors in psychotherapy outcomes. *Annual Review of Clinical Psychology, 15*(1), 207–231.

Davey, G. C. L., Hampton, J., Farrell, J., & Davidson, S. (1992). Some characteristics of worrying: Evidence for worrying and anxiety as separate constructs. *Personality and Individual Differences, 13*(2), 133–147. https://doi.org/10.1016/0191-8869(92)90036-O

Dugas, M. J., Freeston, M. H., Ladouceur, R., Rhéaume, J., Provencher, M., & Boisvert, J.-M. (1998). Worry themes in primary GAD, secondary GAD, and other anxiety disorders. *Journal of Anxiety Disorders, 12*(3), 253–261. https://doi.org/10.1016/S0887-6185(98)00013-9

Dugas, M. J., Sexton, K. A., Hebert, E. A., Bouchard, S., Gouin, J.-P., & Shafran, R. (2022). Behavioral experiments for intolerance of uncertainty: A randomized clinical trial for adults with generalized anxiety disorder. *Behavior Therapy, 53*(6), 1147–1160. https://doi.org/10.1016/j.beth.2022.05.003

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175–191.

Freeston, M. H. (2023). What if we have too many models of worry and GAD? *Behavioural and Cognitive Psychotherapy, 51*(6), 559–578. https://doi.org/10.1017/S1352465822000649

Freeston, M. H., Rhaume, J., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*(6), 791–802.

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. https://doi.org/10.1016/j.janxdis.2017.10.001

Harris, M. G., Hobbs, M. J., Burgess, P. M., Pirkis, J. E., Diminic, S., Siskind, D. J., Andrews, G., & Whiteford, H. A. (2015). Frequency and quality of mental health treatment for affective and anxiety disorders among Australian adults. *Medical Journal of Australia, 202*(4), 185–189. https://doi.org/10.5694/mja14.00297

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*(2), 421–436. https://doi.org/10.1016/j.cbpra.2018.07.007

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology, 12*(1), 281–305. https://doi.org/10.1146/annurev-clinpsy-021815-093436

Hopko, D. R., Reas, D. L., Beck, J. G., Stanley, M. A., Wetherell, J. L., Novy, D. M., & Averill, P. M. (2003). Assessing worry in older adults: Confirmatory factor analysis of the Penn State Worry Questionnaire and psychometric properties of an abbreviated model. *Psychological Assessment, 15*(2), 173–183. https://doi.org/10.1037/1040-3590.15.2.173

Huta, V. (2014). When to use hierarchical linear modeling. *The Quantitative Methods for Psychology, 10*(1), 13–28.

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*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Llera, S. J., & Newman, M. G. (2017). Development and validation of two measures of emotional contrast avoidance: The contrast avoidance questionnaires. *Journal of Anxiety Disorders, 49*, 114–127. https://doi.org/10.1016/j.janxdis.2017.04.008

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Mahoney, A. E. J., Hobbs, M. J., Newby, J. M., Williams, A. D., Sunderland, M., & Andrews, G. (2016). The worry behaviors inventory: Assessing the behavioral avoidance associated with generalized anxiety disorder. *Journal of Affective Disorders, 203*, 256–264. https://doi.org/10.1016/j.jad.2016.06.020

Mahoney, A. E. J., Hobbs, M. J., Newby, J. M., Williams, A. D., & Andrews, G. (2018). Psychometric properties of the worry behaviors inventory: Replication and extension in a large clinical and community sample. *Behavioural and Cognitive Psychotherapy, 46*(1), 84–100. https://doi.org/10.1017/S1352465817000455

McEvoy, P. M., Grove, R., & Slade, T. (2011). Epidemiology of anxiety disorders in the australian general population: Findings of the 2007 Australian national survey of mental health and wellbeing. *Australian New Zealand Journal of Psychiatry, 45*(11), 957–967. https://doi.org/10.3109/00048674.2011.624083

McManus, F., Van Doorn, K., & Yiend, J. (2012). Examining the effects of thought records and behavioral experiments in instigating belief change. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(1), 540–547. https://doi.org/10.1016/j.jbtep.2011.07.003

Newman, M. G., Basterfield, C., Erickson, T. M., Caulley, E., Przeworski, A., & Llera, S. J. (2022). Psychotherapeutic treatments for generalized anxiety disorder: Cognitive and behavioral therapies, enhancement strategies, and emerging efforts. *Expert Review of Neurotherapeutics, 22*(9), 751–770. https://doi.org/10.1080/14737175.2022.2125800

Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the depression anxiety stress scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology, 46*(3), 283–297. https://doi.org/10.1348/014466506X158996

Pelletier, L., O'Donnell, S., McRae, L., & Grenier, J. (2017). The burden of generalized anxiety disorder in Canada-HPCDP. *Health Promotion and Chronic Disease Prevention in Canada, 37*(2), 54–62.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F., Wilson, D. A., Wu, C.-F., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory, 92*(2), 135–138. https://doi.org/10.1016/j.nlm.2008.09.012

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

Ruscio, A. M., Hallion, L. S., Lim, C. C. W., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L. H., Borges, G., Bromet, E. J., Bunting, B., Caldas de Almeida, J. M., Demyttenaere, K., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., Hinkov, H., Hu, C., & Scott, K. M. (2017). Cross-sectional comparison of the epidemiology of DSM-5 generalized anxiety disorder across the globe. *JAMA Psychiatry, 74*(5), 465–475. https://doi.org/10.1001/jamapsychiatry.2017.0056

Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychological Assessment, 25*(1), 146–156. https://doi.org/10.1037/a0029715

Shapiro, M. O., Allan, N. P., Raines, A. M., & Schmidt, N. B. (2023). A randomized control trial examining the initial efficacy of an intolerance of uncertainty focused psychoeducation intervention. *Journal of Psychopathology and Behavioral Assessment, 45*(2), 379–390.

Tolin, D. F., Gilliam, C., Wootton, B. M., Bowe, W., Bragdon, L. B., Davis, E., Hannan, S. E., Steinman, S. A., Worden, B., & Hallion, L. S. (2018). Psychometric properties of a structured diagnostic interview for DSM-5 anxiety, mood, and obsessive-compulsive and related disorders. *Assessment, 25*(1), 3–13. https://doi.org/10.1177/1073191116638410

Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science, 12*(1), 108. https://doi.org/10.1186/s13012-017-0635-3

Wilson, E. J., Abbott, M. J., & Norton, A. R. (2023). The impact of psychological treatment on intolerance of uncertainty in generalized anxiety disorder: A systematic review and meta-analysis. *Journal of Anxiety Disorders, 97*, Article 102729. https://doi.org/10.1016/j.janxdis.2023.102729

Wilson, E. J., Abbott, M. J., Norton, A. R., Berle, D., Stapinski, L., & Rapee, R. M. (2025). A comparison of the long and short forms of the Penn State Worry Questionnaire in adults with generalised anxiety disorder. *Journal of Psychopathology and Behavioral Assessment, 47*(24), 1–12. https://doi.org/10.1007/s10862-025-10200-4

Wilson, E. J., Stapinski, L., Dueber, D. M., Rapee, R. M., Burton, A. L., & Abbott, M. J. (2020). Psychometric properties of the Intolerance of Uncertainty Scale-12 in generalized anxiety disorder: Assessment of factor structure, measurement properties and clinical utility. *Journal of Anxiety Disorders, 76*, Article 102309. https://doi.org/10.1016/j.janxdis.2020.102309