Effectiveness of an internet-delivered intervention for generalized anxiety disorder in routine care: A randomised controlled trial in a student population

Derek Richards a,b,c, Ladislav Timulak a, Chuck Rashleigh c, Orla McLoughlin c, Amy Colla b, Ciara Joyce b, Gavin Doherty d, John Sharry e, Daniel Duffy b, Melissa Anderson-Gibbons b

a School of Psychology, University of Dublin, Trinity College Dublin, Dublin 2, Ireland
b School of Psychology, University of Dublin, Trinity College Dublin, Dublin 2, Ireland
c SilverCloud Health, The Priory, John’s Street West, Dublin 8, Ireland
d Student Counselling, Trinity College Dublin, Ireland
e School of Computer Science & Statistics, University of Dublin, Trinity College, Dublin 2, Ireland
f Mater Misericordiae University Hospital, Eccles St, Dublin 7, Ireland

A R T I C L E   I N F O
Article history:
Received 29 April 2016
Received in revised form 26 October 2016
Accepted 27 October 2016
Available online 29 October 2016

Keywords:
Generalized anxiety disorder
Internet-delivered interventions
Treatment
Students
CBT
Randomised control trial

A B S T R A C T
Background: Cognitive behavioural therapy is one of the main and preferred treatments for generalized anxiety disorder. Numerous barriers can hinder an individual from seeking or receiving appropriate treatment; internet-delivered CBT interventions offer a relatively new means of increasing access to treatment.
Methods: A service-based effectiveness randomised waiting list control trial examined the impact of an internet-delivered CBT intervention, Calming Anxiety, amongst Irish university students (N = 137). Primary outcome was self-reported GAD and secondary outcomes included depression and work and social functioning.
Results: Analyses returned inconclusive results. Both treatment and waiting list conditions displayed significant decreases in anxiety symptoms post-treatment, but we did not observe a significant between-group effect (p = 0.076). Significant within-group differences from pre to post time points were observed for depression (BDI-II) and work and social functioning (WASA), and between group differences were also significant for depression (d = 0.46) and functioning (d = 0.36). Both groups demonstrated cases of remission and recovery from anxiety, however differences in the number of cases reaching clinically meaningful change between conditions were non-significant.
Conclusions: Several explanations regarding the results are presented, examining issues related to active waiting lists, study limitations and treatment expectancies.

Trial registration: Current Controlled Trials ISRCTN16303842.
© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

GAD is characterised as a chronic condition, with a DSM-V diagnosis typically requiring the persistent symptom of excessive worry occurring over a period of 6 months (American Psychiatric Association, 2013). Further symptoms consist of hypervigilance and the typical somatic responses of anxiety, where the chronicity of these symptoms often results in long-term personal suffering and feelings of a loss of control (Wittchen, 2002). European statistics place the lifetime prevalence of GAD to be between 4.3 and 5.9% and a 12-month prevalence to be between 1.2 and 1.9% (Wittchen, 2000). However, a significant amount of individuals do not seek treatment for their disorder (Wittchen et al., 2011). GAD tends to present itself with comorbidity, most often with mood and other anxiety disorders (Alonso et al., 2004; Royal College of Psychiatrists, 2011). Furthermore, negative effects can extend from the direct effects the disorder has on the individual; economic, personal and social roles can be negatively impacted upon by symptoms of GAD, which in turn can decrease quality of life (Loebach Wetherell et al., 2004; Stein and Heimberg, 2004).

1.1. Anxiety and students

Several studies have examined the link between anxiety symptomatology and university students (Royal College of Psychiatrists, 2011). Major life events, such as the transition to university life and the responsibilities that come with it can be a large source of anxiety. Furthermore, societal trends such as the increasing financial cost of university places extra burden on students, where they often have to source employment in order to fund their college lives (Royal College of Psychiatrists, 2011).
The rising incidence of mental health issues was raised by Dooley and Fitzgerald (2012) in their survey of mental health and wellbeing in an Irish sample. Utilising an Irish student-based sample, the researchers placed levels of anxiety at 36% and 30% with symptoms of stress. Although this research did not include data on GAD in students specifically.

1.2. Treating anxiety disorders

Treatment of anxiety disorders, and more specifically GAD, has largely consisted of pharmacological treatments, psychological therapies or a mixture of both (Gould et al., 1997). However, completion and cessation of a pharmacological therapy without continuing to take maintenance drugs can result in relapse (Davidson et al., 2008; Allgulander et al., 2006). Psychological therapies are often preferred, especially where they have produced similar outcomes to pharmacological treatments, along with lower relapse rates (Sturmev and Hersen, 2012).

Cognitive-behaviour therapy (CBT) is routinely chosen as the psychological treatment of choice for GAD, and has been subject to several successful trials (Cuijpers et al., 2014). CBT for GAD is composed of many elements, such as cognitive restructuring, worry exposure, mindfulness, relaxation techniques and information on the disorder (Dugas et al., 2003; Orsillo et al., 2003). CBT for GAD can be performed in either a group or one-to-one setting, where the therapist helps the patient to understand their disorder, manage it and regain quality of life. However, it has been estimated that across the anxiety disorders a large percentage do not seek treatment, or have a significant delay in receiving treatment (Kohn et al., 2004). One significant barrier is the ability to access evidence-based treatments such as CBT. Internet-delivered CBT (iCBT) interventions are a relatively new development that could help with increasing access to treatments.

1.3. iCBT for GAD

Several studies have investigated the effectiveness of iCBT for GAD and results have shown positive outcomes in regards to post-treatment and follow-up symptom relief (Titov et al., 2009, 2011; Padling et al., 2011). Internet-delivered psychological interventions for GAD are a promising new intervention with a growing evidence base. Results from research trials are evidencing large effect sizes that can be sustained up to 2/3 years post-treatment (Titov et al., 2016). Preliminary results are encouraging, and speak to the potential of using therapist, or self-guided iCBT for GAD (Titov et al., 2016).

Richards et al. (2015a) conducted a systematic review and meta-analysis of internet-delivered psychological interventions for the treatment of GAD. Studies recruited primarily through websites and/or adverts in local newspapers; some used clinical samples referred from GPs or mental health practitioners. All studies used the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First et al., 1997), the Mini International Neuropsychiatric Interview Version 5.0.0 (MINI; Sheehan et al., 1998), or an interview based on the MINI (Johansson et al., 2013) to establish a formal diagnosis of GAD, over the phone or in person. Four of the eleven studies targeted a GAD sample through a GAD-specific programme; the other seven were “transdiagnostic” in nature, addressing multiple anxiety disorders or GAD and mood disorders. Most studies (9/11) used CBT-based content while two used a psychodynamic approach. Generally, content was delivered in 6–8 modules over 8–10 weeks; most interventions also included some form of support from a psychologist or therapist, typically by phone or e-mail.

There were statistically significant improvements for internet-delivered interventions compared to waiting-list controls on self-reported GAD symptoms ($d = -0.91$) and pathological worry ($d = -0.74$), both yielding what can be considered large effects (Cohen, 1988). Similar effects were found for the active treatments compared to waiting-list controls for comorbid anxiety ($d = -0.57$), depression ($d = -0.63$), distress ($d = -0.91$), disability ($d = -0.77$), and quality of life ($d = 0.38$). For GAD subjects, effect sizes were similar for GAD-specific ($d = -0.81$) and transdiagnostic ($d = -0.91$) interventions. Psychodynamic approaches (2 studies) had less favourable results; in fact, one study (Andersson et al., 2012) saw an unexpected improvement in the waiting list group. The authors advised caution in interpreting these results given limited and heterogenous data, and suggested what research is needed to strengthen the field.

The current study aimed to implement an iCBT (Calming Anxiety) for GAD. Based on previous literature in the area of online interventions for GAD (Richards et al., 2015a), it was hypothesised that the intervention would produce significant decreases in GAD symptoms at post-intervention for the treatment group compared to the control group.

2. Methods

2.1. Design

The current study utilised a service-based effectiveness, randomised controlled trial design in order to examine the delivery of an Internet intervention for the treatment of individuals with GAD symptoms. The trial was registered (Current Controlled Trials ISRCTN16303842) and the protocol published (Richards et al., 2014). Participants were randomised into two groups using; the internet-delivered intervention (iCBT, Calming Anxiety) with clinical support and a waiting list control group. Randomisation was achieved using a computer algorithm established by a programmer and independently executed. The randomisation took place at the individual level using a 1:1 format for distribution between the groups. Participants in the waiting list control group did not receive any treatment for the 6-week duration in which the treatment group was receiving the intervention. At week 7, participants in the waiting list control group were given access to the supported intervention. The current paper describes the main outcomes from the trial. Other research data related to this trial including quality of life, satisfaction with treatment and significant events data are being analysed and will be reported elsewhere (see trial protocol for details; Richards et al., 2014).

Ethical approval for the current study was obtained from the appropriate University Ethics Committee (25/11/2013). Prior to commencing the study, all participants received information that detailed the intervention in its entirety. Informed consent was then obtained and participants were made aware that their involvement was completely voluntary and that they were free to withdraw from the study at any stage without prejudice.

2.2. Participants and sample size

Participants for the current study were recruited through the University Counselling Service and were all registered students. Participants were contacted via e-mail (delivered college wide) in order to advertise the current study and request their participation. This e-mail detailed how the prospective participant could obtain further information and initiate the screening process. This e-mail was sent to all students two times at an interval of two weeks during the first three weeks of the second academic semester.

Primary eligibility criteria for participation in the study was based on whether the individual’s self-reported GAD symptoms were confirmed to reach an acceptable clinical threshold, defined for the study as a score of 10 or above on the GAD-7 measure. Further criteria for inclusion in the study were that participants were to be at least 18 years of age or older. Participants attending face-to-face counselling were excluded from the study. In order to target the trial as closely as possible to a GAD population, severe depression was used as a criterion for excluding possible comorbidities (BDI II scores >29). Suicidal ideation was also
used as an exclusion criterion to ensure adequate service response to a client group seeking online support (score of 1 or higher on BDI II item 9). It was deemed unethical to include clients experiencing suicidal ideation for an online intervention without any face-to-face needs assessment. The study recruited and screened \( N = 324 \), 187 were excluded and 137 were included into the RCT and randomised (Treatment \( n = 70 \), Waiting List \( n = 67 \)).

2.3. The intervention

The Calming Anxiety supported programme was delivered to participants over the course of the 6 week intervention, where each week participants completed an online CBT-based intervention module for GAD. The programme structure and content were based on CBT concepts and strategies and included cognitive, emotional and behavioural components such as self-monitoring, relaxation, cognitive restructuring and worry outcome monitoring (Behar et al., 2009).

The intervention was delivered on a web 2.0 platform using interactive content. Each module was identically structured, where introductory quizzes, videos, informational content, interactive content activities and homework suggestions formed a part of each module. Personal stories and accounts from other clients were also incorporated into the programme and this was presented to participants as a text piece or video. Further details on the individual models are presented in Table 1.

Because this was a naturalistic study and student users were encouraged to complete one module per week for the duration of six weeks, this was not guaranteed. It is the case that students progressed through the intervention as they would normally through service offered by the counselling centre, that is, they could pause their participation at any time, they did not necessarily adhere strictly to the one module per week, and many keep using the intervention after the initial 6-week period.

A key function of the supporters was to monitor participant progress, helping participants who chose to do more than one module per week to get the most out of each module, and encouraging participants who worked slower through the intervention to find the most suitable content for their needs. Many participants continued to use the intervention after the 6-week treatment period.

Participants could access all modules at once, and could complete them in any order. There were no mandatory aspects of the programme – all reading and homework tasks were suggested, but not required to progress through the programme. Participants could also choose to share the content from exercises with supporters so that they were free to complete them without having to discuss them if this was preferable.

If participants did not engage with content for two weeks in a row, supporters wrote a review that let participants know their account was “paused” – a new review date was set if and when participants chose to log back in. The system sent automatic e-mails to encourage participants to log back in and use the programme. Usage statistics and feedback from supporters indicate that a proportion of participants signed up to the programme, but did not access any of the psychoeducational content, had their accounts paused, and did not log back in (except to complete time 2 data).

2.4. Supporters

Upon commencing treatment, participants were assigned a supporter to monitor their performance throughout the 6-week intervention. There were six supporters involved in the trial. All supporters were psychologists with a master’s degree or higher. The training for supporters consisted of first working through the entire Calming Anxiety programme as a user. Additionally, 4 h of training followed on how to write messages and how to screen user content for disclosures of risk. All supporters received weekly supervision regarding their work within the trial. Each week, supporters logged in and evaluated the work completed by participants spending 10–15 min providing feedback to each participant at a scheduled time. All feedback was sent to participants through the Calming Anxiety messaging system.

Participants were assigned to a specific supporter, based on availability, who followed their progress throughout their time using the programme to ensure continuity of care. Supporters provided limited information about themselves through a “profile”, where they chose whether to include a profile picture and a brief summary of who they were. Supporters had a guide to help them give feedback. They could choose how closely they based their responses to participants on the examples given, as long as their feedback conformed with the psychoeducational content of the programme, and was responsive to any homework content shared by participants or messages sent to them. On the review day, supporters were able to see what pages of content participants had accessed, their ratings of different modules, and whether participants had commented on pages, created journal entries, or used certain activities. Supporters could only see the content of comments and activities if the participant chose to “share” these with them. Supporters could also see any messages left for them by participants.

2.5. Risk management

Upon signing into the hosting platform for the first time and creating their user account, participants were presented with a user contract that detailed Calming Anxiety and the study in its entirety. This user contract also stated that Calming Anxiety does not provide emergency support and that if at any stage a participant feels a change in their wellbeing, they should contact their general practitioner or a clinician at the counselling service. Further to this, Calming Anxiety also provided a page with information about specific services the individual could contact in this case. Upon digitally signing this user contract, participants were assumed to understand the information that was presented to them.

Table 1

| Module name                      | Brief description                                                                 |
|----------------------------------|-----------------------------------------------------------------------------------|
| Getting Started                  | Outlines the basic premise of CBT and provides some information about anxiety. Users are encouraged to explore their current difficulties with anxiety and to begin monitoring their anxiety levels. |
| Understanding Moods and Emotions | This module describes the behavioural, physical, and emotional aspects of the Thoughts-Feelings-Behavior (TFB) cycle. The user is introduced to relaxation practices. Users start to build their own anxiety-related TFB cycles. |
| Anxious thoughts and worry       | This module focuses on noticing anxious thoughts and worry, and ways of relating to these thoughts, including acceptance, distraction, and ‘worry time’ |
| Face your Anxiety, Step by Step  | This module outlines why avoidance is harmful, and breaks down the steps needed for successful graded exposure. Users are encouraged to build their own fear hierarchies and to begin working through them. |
| Challenge your Anxious Thoughts  | This module explains negative automatic thoughts, their role in anxiety, and how to challenge them. Users are encouraged to challenge the thoughts in their TFB cycles, and make use of helpful thoughts. |
| Bringing it all Together         | In this final module, users are encouraged to bring together all the skills and ideas they have gathered so far, note their personal warning signs, and make a plan for staying well. |
When students first registered and completed the battery of assessments below, they were “flagged up” by the system if they scored 29 or more on the Beck Depression Inventory (BDI-II) and/or indicated suicidal ideation in the relevant item. Students were informed that they were temporarily not able to complete sign-up because of their scores; the message also stated that they would be contacted within a working day, and referred to emergency services. A member of the counselling team then contacted the student to arrange a face-to-face or phone meeting. Of those excluded from the study, they were offered other appropriate sources of support.

2.6 Assessments

2.6.1 Socio-demographic data

The Sociodemographic & Clinical History questionnaire was based on that used in an earlier study (Richards et al., 2013), and was developed to collect demographic and clinical details of the participants. To this extent, the questionnaire accounts for any existing diagnosis of anxiety disorders and the length of time that participants had experienced anxiety symptoms. It collected data on participant’s experience of counselling, therapy, and medication for anxiety, and on whether participants had any existing diagnoses of mental health disorders. In addition, it contained items related to comorbid psychosis, alcohol and drug misuse, and/or any recent medical diagnosis.

2.6.2 Primary outcomes measure

The primary outcome measure for the current study was the Generalized Anxiety Disorder 7 (GAD-7; Spitzer et al., 2006) that comprises 7 items measuring symptoms and severity of GAD based on the DSM-IV diagnostic criteria for GAD. The GAD-7 has good internal consistency (α = 0.89) and convergent validity with other anxiety scales (Kroenke et al., 2010). Higher scores indicate greater severity of symptoms. The GAD-7 has increasingly been used in large-scale studies as a generic measure of change in anxiety symptomatology (Richards and Suckling, 2009). Using the threshold score of 10, the GAD-7 has good convergent validity with other measures of depression across clinical and nonclinical adult samples (Beck et al., 1996).

2.6.3 Secondary outcomes measures

The Work and Social Adjustment (WASA; Mundt et al., 2002) is a simple, reliable and valid measure of impaired functioning. It is a 5-item self-report measure which provides an experiential impact of the disorder on the patient’s point of view. The measure looks at how the disorder impairs the patient's ability to function in day to day life on five dimensions: work, social life, home life, private life and close relationships.

The Beck Depression Inventory 2nd edition (BDI II; Beck et al., 1996) assesses depressive symptoms that correspond to the criteria for a depressive disorder diagnosis as outlined in the DSM-IV (American Psychiatric Association [APA], 2000). The scale designates four levels of severity: minimal (0 to 13), mild (14 to 19), moderate (20 to 28) and severe (29 to 63) (Beck et al., 1996). The scale has demonstrated good convergent validity with other measures of depression across clinical and nonclinical adult samples (Beck et al., 1988).

The Penn State Worry Questionnaire (PSWQ) (Meyer et al., 1990) consists of 16 items and is considered a valid clinical measure of the worry characteristic of GAD. Each item is measured on a 5-point Likert scale (1 – not at all typical of me to 5 – very typical of me) and a total score ranging between 0 and 80 is calculated by summing all items. Psychometric evaluations have revealed a high internal consistency (α = 0.86 to 0.95) and test-retest reliability over four weeks (r = 0.74 to 0.93; Molina and Borkovec, 1994). The measure has been found to successfully differentiate between patients with GAD and those with other anxiety disorders (Brown et al., 1992).

2.7 Data analysis

The analysis of the main outcome measure (GAD-7) was based on an intention-to-treat principle, where data from those who began treatment and provided subsequent data were included in the data set irrespective of treatment compliance. Last observation carried forward (LOCF) was initially proposed to handle missing data where it occurred. However, expectation maximisation (EM; Dempster et al., 1977) was chosen over LOCF to be implemented. EM was implemented for missing data using SPSS 21. As a method for handling missing data, it is an accessible procedure that generates a single imputation of data for analysis.

Originally, data analysis was planned to span across three time points – pre, post and 3 month follow-up. However, participant response for the third time point in regards to both the treatment and waiting list group returned minimal numbers (n = 24: 21 in the treatment group, 3 in the waiting list group). As such, this time point was omitted due to lack of valid data available for analysis and due to risk of biasing the imputation procedure.

In order to test the main hypotheses, repeated measures ANOVA were proposed to be performed in the trial protocol. However, on further examination of the data analysis procedure, and based on reviewers comments, it was decided that univariate analyses of covariance (ANCOVA) would be performed, using pre-treatment scores as covariance in each respective analysis. Within group differences were analysed using paired samples t-tests. Participants were then compared across groups and time points. Effect sizes were established using Cohen’s d statistic (Cohen, 1988).

Following the main analyses, pre and post scores were examined in order to estimate the percentage of participants who made a clinically meaningful change at the end of treatment. An estimation of remission was made using data from those who scored above the established clinical cut offs at pre measurement and compared with their post data in order to examine whether they scored below the clinical cut off (GAD-7 cut-off < 10, BDI-II cut-off < 14, PSWQ cut-off < 45, WASA cut-off < 10). Estimates of reliable recovery were made by identifying the number of participants in each group who demonstrated a 50% pre-treatment reduction of GAD-7, BDI-II and PSWQ scores.

3. Results

3.1 Baseline characteristics

Three hundred and twenty-four students were recruited and screened for eligibility and of those 137 were randomised to either the initial or delayed treatment groups. Chi square tests revealed that at post-randomisation there were no significant differences on any variables between the treatment group and waiting list control group. Fig. 1 shows the participant flow through the trial and the reasons for exclusion. Participants (N = 137) were predominately female (77%), with a mean age of 23.82 (S.D. = 7.05) and the age range was 17–58 years. Participant characteristics are presented in Table 2.

3.2 Treatment response rate

As per the treatment protocol participants were offered 7 modules of content to complete. Participants were encouraged to progress through in a linear fashion, but were not excluded from using the modules as they choose. Treatment response is a difficult concept in online as it seems that individuals’ progression through content in various ways can all lead to positive outcomes (Richards et al., 2015b). Graph 1 displays the adherence to the modules as a simple understanding of adherence in online interventions. Of those who began the intervention > 50% completed 5 modules of content and almost 25% completed all 7 modules. The average number of logons was 8.31 and the average amount of time spent per session was 27.38 min.
3.3. Primary & secondary outcomes

As per the published protocol (Richards et al., 2014), analyses were conducted on the data using participants with a cut-off of ≥10 on GAD-7 scores. The means, standard deviations and effect sizes (d) associated with all outcome measures for each group and between groups are presented in Table 2. Correcting for multiple testing, a conservative p criteria of <0.01 was applied to both paired samples t-tests and ANCOVA.

Paired samples t-test for GAD-7 scores conducted on the treatment and waiting-list control group showed significant within group differences between pre and post time points, t(1, 69) = 10.2, p ≤ 0.01 and t(1, 66) = 6.72, p ≤ 0.01 respectively. Univariate ANCOVA procedures on post-treatment GAD-7 scores, using pre-treatment scores as a covariate, showed no significant difference in generalized anxiety disorder symptoms between the treatment and control group, F(1, 134) = 3.20, p = 0.076.

Using paired samples t-tests, significant within group differences from pre to post time points on the BDI-II and WASA were observed. No significant within groups differences, in either the treatment or control group, were observed for PSWQ scores from pre-post time points. These results are summarised and presented in Table 3.

Univariate ANCOVAs were also carried out for secondary outcome measures, using pre-treatment scores as a covariate. No significant difference was found between treatment and control groups for PSWQ worry scores, F(1, 134) = 1.03, p = 0.311. Significant differences were found for both BDI-II and WASA, F(1, 134) = 7.02, p ≤ 0.01 and, F(1, 134) = 8.35, p ≤ 0.01 respectively. Table 4 illustrates a summary of the ANCOVA results, including between groups effect sizes (Cohen’s d) with 95% confidence intervals.

3.4. Clinically significant change

As previously discussed, clinically significant change was defined in two ways. First, an estimation of remission was established using the criteria for a cut-off of below 10 on GAD-7. Using this cut-off in the treatment group, we observed remission of anxiety symptoms in 53 (75%) participants in the treatment group compared to 40 (59%) participants in the waiting list control group. When correcting for multiple testing, chi squared analyses showed no significant difference between the two groups in terms of cases of remission, Χ² (1, N = 137) = 4.03, p = 0.045. Second, we defined recovery as a 50% reduction in pre to post scores on the GAD-7. In the treatment group, we observed this change in 33 subjects (47.1%) compared to 23 (34%) subjects in the waiting list control group. Further chi squared analyses found no differences between groups exhibiting a 50% reduction in symptoms as per the GAD-7, Χ² (1, N = 137) = 2.33, p = 0.13. Cut-off data are presented in Table 5.

The number needed to treat (NNT) for a client to reach recovery from the treatment was calculated at 8 individuals and for remission this number would be 6 (Cook and Sackett, 1995; Wen et al., 2005).

4. Discussion

The current study sought to investigate the effectiveness of the internet-delivered iCBT Calming Anxiety programme as a low-intensity
intervention for the treatment of general anxiety disorder symptoms in a student population. Data analysis highlighted a main effect of time, where participants in both groups (including wait-list) displayed a reduction in GAD symptoms from pre to post measurement. Analysis of the secondary outcomes indicated similar results for the PSWQ. Contrary to our expectations we did not show a significant between group effects for the main outcome on anxiety.

Other studies of similar design and topic to that of the current piece of research have demonstrated statistically significant between group differences (Paxling et al., 2011; Titov et al., 2009; Robinson et al., 2010). All three studies included support elements, much like the current study, and specifically featured iCBT interventions targeted at GAD, as opposed to transdiagnostic approaches. In these studies, the improvements in the waiting list groups, when quantified in terms of Cohen’s $d$, were much smaller than those of the present study (Paxling et al., 2011, $d = 0.18$; Titov et al., 2009, $d = 0.34$; Robinson et al., 2010, $d = 0.36$). From an analysis of recent literature, the current study can be seen to contradict normative results in the field, such that the control group improved significantly in comparison to the active treatment group. Therefore, the results of the current study can be viewed as unusual. Below, several possible explanations for these results are provided.

![Graph 1. Adherence to modules over time.](image-url)
Perhaps, when interpreting these results, a factor of note is that 206 participants of the initial 324 in the current study were recruited in January. During this period in the university where the trial occurred, winter examinations take place over the course of several weeks. It is plausible that the symptoms of GAD experienced at baseline may have been inflated due to academic pressures and did not accurately reflect normal manifestations of the disorder. However, at post-treatment data collection and following exams and the subsequent decrease in workload, stress and anxiety levels may have decreased, thus reducing GAD symptoms across the entire sample.

The overall decreasing trend in self-reported symptoms of depression in the treatment group suggests that the intervention demonstrated some potential not only addressing anxiety and worry, but the comorbid depressive symptoms so closely related to anxiety disorders in general. In addition, the significant changes noted for participants on the WASA is a positive indication of improvement especially given the deleterious effects mental illness can have upon work and social functioning (Loebach Wetherell et al., 2004; Stein and Hemberg, 2004).

A point of interest in the discussion is that the current intervention was shorter than typical internet-delivered interventions for GAD and the time to complete was generally shorter too (Andersson et al., 2012; Berger et al., 2014; Paxling et al., 2011; Titov et al., 2011). This may be suggestive that first the intervention itself may require tweaking and second that its format for delivery may need to be different in natural settings.

The result also leaves us with some questions about the entire project and some of these we would like to consider further. Treatment expectancies have been found to play a significant role in the outcomes of psychological therapies (Ditto and Westra, 2005), where positive expectancies can be found to produce more positive outcomes (Snyder et al., 2000). The current study utilised a waiting list control where participants were administered psychometric measures and informed they would be receiving treatment after a short period of time. Through promotional materials for the study or through self-sought information about cognitive behavioural therapy, participants in the waiting list may have held positive expectations for the therapy they would receive. The increase in scores, may in itself, have been the product of a quasi-placebo effect. In addition to these considerations is the further possibility that students in the current study may have accessed additional treatments or interventions during their participation in the research which may have impacted on results. This was not explicitly controlled for in this study.

A number of other factors may have confounded the results of the current study. Due to the naturalistic setting of the research, the study was constricted to the use of a single measure of anxiety, the GAD-7, rather than a battery of anxiety measures. GAD, however, is one of the most difficult psychological disorders to diagnose accurately (Brown et al., 2001a). Despite revisions to the diagnostic criteria of GAD since its addition to the DSM in 1980 (DSM-III), it continues to demonstrate lower reliability than other disorders (Brown et al., 2001b). Inherent in the criteria of GAD include issues pertaining to criteria that overlap with aspects of other disorders and vagueness of the criteria; diagnosis relies on subjective report of internal processes and lacks behavioural markers (Brown et al, 2001a, 2001b; Robichaud, 2013). Moreover, the evolving nature of the GAD construct has somewhat impeded the development of research, theory and appropriate measures (Holaway et al., 2006; Turk and Wolanin, 2006). By using this as the primary outcome measure we may not have been accurately capturing the mechanisms of change that the programme was facilitating but were rather looking for specific GAD symptomology in a population who were not routinely receiving diagnoses of disorders but were instead experiencing generalised anxiety in more individualised ways.

Follow-up was largely unsuccessful in the current study. Again as we were implementing the trial in a natural setting we were not permitted to engage in a robust follow-up protocol, therefore any post treatment data that was collected was due mainly to a follow-up email requesting participants to complete that data; further confounded by time at follow-up which was holiday period for most students. The result was a very small follow-up data that we decided not to include in the analysis.

### Table 3
Results of paired samples t-tests for within group differences across time points, including M, SD, M differences and effect sizes.

| Measure   | Group       | Pre M (SD) | Post M (SD) | t-test | Pre-post M difference (95% CI) | Effect size d (95% CI) |
|-----------|-------------|------------|-------------|--------|--------------------------------|------------------------|
| GAD 7     | Treatment   | 12.84 (2.39) | 7.73 (4.44) | t(1, 69) = 10.21, p ≤ 0.01 | 5.11 (4.11–6.11) | 1.5 (0.94–2.54) |
|           | Control     | 13.19 (2.78) | 9.13 (4.13)  | t(1, 66) = 6.72, p ≤ 0.01 | 4.06 (2.85–5.27) | 1.18 (0.51–2.16) |
| BDI-II    | Treatment   | 17.67 (6.44) | 14.54 (8.07) | t(1, 69) = 3.26, p ≤ 0.01 | 3.12 (1.22–5.03) | 0.43 (1.08–2.32) |
|           | Control     | 17.87 (5.41) | 18.1 (8.99)  | t(1, 66) = –0.42, p ≥ 0.05 | –0.43 (–2.51–1.64) | 0.06 (–3.16–2.09) |
| PSWQ      | Treatment   | 63.04 (8.11) | 58.53 (10.97) | t(1, 69) = 3.69, p ≤ 0.01 | 4.51 (2.08–6.95) | 0.47 (–1.43–3.04) |
|           | Control     | 63.48 (6.95) | 60.33 (8.79)  | t(1, 66) = 3.25, p ≤ 0.01 | 3.15 (1.21–5.08) | 0.41 (1.26–2.5)  |
| WASA      | Treatment   | 11.93 (6.39) | 9.69 (6.16)  | t(1, 66) = 2.83, p ≤ 0.01 | 2.24 (0.66–3.82) | 0.36 (1.14–1.8)  |
|           | Control     | 14.48 (6.31) | 13.34 (5.75)  | t(1, 66) = 1.25, p ≥ 0.05 | 1.13 (–0.684–2.95) | 0.19 (–1.32–1.56) |

*Note. GAD-7: General Anxiety Disorder 7; BDI-II: Beck Depression Inventory, 2nd edition; PSWQ: Penn State Worry Questionnaire; WASA: Work and Social Adjustment questionnaire.*

### Table 4
Results from ANCOVA, with baseline scores as covariate.

| Measure   | ANCOVA         | d (95% CI) |
|-----------|----------------|------------|
| GAD 7     | F(1, 134) = 3.20, p ≤ 0.05 | 0.3 (–0.07–1.3) |
| PSWQ      | F(1, 134) = 1.03, p ≤ 0.05 | 0.18 (–1.92–2.23) |
| BDI-II    | F(1, 134) = 7.02, p ≤ 0.01 | 0.46 (–1.43–2.3) |
| WASA      | F(1, 134) = 8.35, p ≤ 0.01 | 0.36 (–1.169) |

*Note. GAD-7: Generalized Anxiety Disorder 7; BDI-II: Beck Depression Inventory, 2nd edition; PSWQ: Penn State Worry Questionnaire; WASA: Work and Social Adjustment questionnaire.*

### Table 5
Participants above and below the designated cut-off scores from pre to post.

| Measure   | Treatment (n = 70) | Waiting list (n = 67) | Total sample (N = 137) |
|-----------|-------------------|----------------------|------------------------|
| GAD-7     | Above Below Above Below | Above Below Above Below | 137 0 |
|           | Post 17 53 40 44 | 109 28 |
|           | Pre 17 53 27 40 | 44 93 |
| BDI       | Pre 53 17 11 61 | 109 28 |
|           | Post 41 29 16 62 | 92 43 |
| PSWQ      | Pre 70 0 66 1 | 136 1 |
|           | Post 62 8 64 3 | 126 11 |
| WASA      | Pre 43 27 55 12 | 98 39 |
|           | Post 37 33 53 14 | 90 47 |

*Note. Cut-off scores applied: GAD-7 cut-off = 10; BDI-II cut-off = 14; PSWQ cut-off = 45; WASA cut-off = 10.*
4.1. Limitations and future research

Given the above difficulties, it can be proposed that these may have had a number of implications for the current study in terms of both measurement and intervention. The use of the GAD-7 may have proved limiting and problematic in several respects. Firstly, while this measure has been demonstrated to be congruent with the DSM-IV, only one study has examined and reported this congruency (Spitzer et al., 2006). This study utilised a relatively older student sample (mean age 23), posing questions about the generalisability of such results to the student population in the current research. Furthermore, the current study adopted the cut-off point of GAD-7 score ≥10 used in the congruency study, to indicate the presence of GAD. Healthcare services within the UK (NHS; IAPT services) and related research projects (e.g. Clark et al., 2009), however, use a GAD-7 cut-off point of ≥8 to indicate caseness; future research in naturalistic settings might then be recommended to follow suit using a cut-point as applied in practice. Indeed our analysis would show significant movement from caseness to non-caseness under these criteria.

These difficulties inherent in the accurate assessment and measurement of GAD thus raise questions about the nature of the population who participated in this research. Whether the students identified in this study were indeed a homogeneous group of adults who met the appropriate criteria for GAD (i.e. a GAD specific population) or a mixed group presenting with various anxiety related symptoms appears unclear and presents as a substantial limitation to this study. Given that the current intervention was targeted specifically at GAD, one could speculate whether treatment effects were concealed by this possible heterogeneity. Future research would recommend use of a wider range of assessment measures to enable increased accuracy of GAD identification and avoid the occurrence of false positives, as a single measure such as the GAD-7 is insufficient. Indeed, the GAD-7 only captures respondents’ answers from a two week time period whereas formal diagnosis requires an evaluation that considers a period of six months minimum. Furthermore, given the difficulty of GAD assessment, some researchers have put forward the suggestion that a diagnosis of GAD should ideally be confirmed by two separate assessment interviews (Borkovec and Costello, 1993; Brown et al., 2001b). Such recommendations, however, may be difficult to implement in naturalistic research settings.

This trial puts forward two main suggestions for the design of a trial concerning iCBT and GAD. Firstly, when utilising a university sample and targeting an illness of a specific nature, dates for the implementation of interventions and symptom recording/assessment should be scheduled appropriately. A large number of participants in the current study (206 out of 324) were recruited and assessed during the winter examination period. This may have had an effect on initial anxiety scores in both groups, where individuals presented with inflated anxiety during exams, and then both groups returned to sub-threshold scores post examinations. This may also serve as a valid explanation for the absence of interaction effects in the current study.

The second recommendation is informed by a further limitation of the study. As we are unable to determine whether participants accessed other treatments during the study period we could not control for this in the analysis, but it may be a factor of some importance in explaining the results. Future research in naturalistic setting should plan for recording what other, if any, treatments or supports that students might engage with that could potentially impact on outcomes. For instance, in addition to counselling and psychological interventions some universities have learning and academic supports for students and these may have a positive impact on anxiety symptoms if they are related to academic stressors.

4.2. Conclusion

The current study reports a main effect of time, but no between group differences following an internet-delivered intervention targeted at GAD in college students. In light of the potential difficulties inherent in the concept, measurement and nature of GAD, it seems reasonable to question whether a number of confounding factors related to these may have masked potential effects of the intervention. Other aspects that may have confounded the potential of the intervention include the implementation of the intervention, the time required to complete the intervention and the fact that the intervention content itself may benefit from appropriate revision. This study highlights some further considerations for implementing and researching iCBT interventions in naturalistic settings.

Conflict of interest

GD and JS declare a minority interest in the commercialization of the software hosting the current calming anxiety programme.

Acknowledgements

Thanks to all the people who participated in the study. Thanks to the student counselling service for hosting the study. Thanks to the reviewers for their helpful comments and direction.

References

Allgulander, C., Florea, I., Huusom, A.K.T., 2006. Prevention of relapse in generalized anxiety disorder by escitalopram treatment. J. Neuropsychopharmacol. 9 (5), 495.
Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., ... Vollebergh, W.A.M., 2004. 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr. Scand. 109 (420), 28–37.
American Psychiatric Association [APA], 2000. Diagnostic and statistical manual of mental disorders (DSM-IV-TR). 4th ed. American Psychiatric Association, Washington D.C.
American Psychiatric Association [APA], 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). fifth ed. American Psychiatric Association, Washington D.C.
Andersson, C., Paxling, B., Roos-Norlund, P., Östman, G., Norgren, A., Ahlström, J., ... Carlbring, P., 2012. Internet-based psychotherapy versus cognitive behavioral guided self-help for generalized anxiety disorder: a randomized controlled trial. Psychother. Psychosom. 81 (6), 344–355.
Beck, A.T., Steer, R.A., Carbin, M.G., 1986. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin. Psychol. Rev. 8 (1), 77–100.
Beck, A.T., Steer, R.A., Brown, G.K., 1996. Manual for the BDI-II. Psychological Corporation, San Antonio, TX.
Behar, E., DiMarco, I.D., Hekler, E.B., Mushlin, J., Staples, A.M., 2009. Current theoretical models of generalized anxiety disorder (GAD): conceptual review and treatment implications. J. Anxiety Disorder. 23 (8), 1011–1023.
Berger, T., Boeticher, J., Caspar, F., 2014. Internet-based guided self-help for several anxiety disorders: a randomized controlled trial comparing a tailored with a standardized disorder-specific approach. Psychotherapy 51 (2), 207.
Borkovec, T.D., Costello, E., 1993. Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. J. Consult. Clin. Psychol. 61 (6), 609–617.
Brown, T.A., Antony, M.M., Barlow, D.H., 1992. Psychometric properties of the Penn State Worry Questionnaire in a clinical anxiety disorders sample. Behav. Res. Ther. 30 (1), 33–37.
Brown, T.A., Di Nardo, P.A., Lehman, C.L., Campbell, L.A., 2001a. Reliability of DSM-IV anxiety and mood disorders: implications for the classification of emotional disorders. J. Abnorm. Psychol. 110 (1), 49.
Brown, T.A., O’Leary, T.A., Barlow, D.H., 2001b. Generalised Anxiety Disorder. In: Barlow, D.H. (Ed.), Clinical Handbook of Psychological Disorders (Third Edition): A Step-by-step Treatment Manual. Guildford publications, New York, pp. 154–208.
Clark, D.M., Layard, R., Smithies, R., Richards, D.A., Suckling, R., Wright, B., 2009. Improving access to psychological therapy: initial evaluation of two UK demonstration sites. Behav. Res. Ther. 47 (11), 910–920.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. 2 ed. Lawrence Earlbaum Associates, Hillsdale, NJ.
Cook, R.J., Sackett, D.L., 1995. The number needed to treat: a clinically useful measure of treatment effect. BMJ Br. Med. J. 310 (6977), 452–454.
Cuypers, P., Sijbrandij, M., Koole, S., Huibers, M., Berkling, M., Andersson, G., 2014. Psychol- ogy treatment of generalized anxiety disorder: a meta-analysis. Clin. Psychol. Rev. 34 (2), 130–140.
Davidson, J.R., Wittchen, H.U., Llorca, P.M., Erickson, J., Detke, M., Ball, S.G., Russell, J.M., 2008. Duloxetine treatment for relapse prevention in adults with generalized anxiety disorder: a double-blind placebo-controlled trial. Eur. Neuropsychopharmacol. 18 (9), 673–681.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. B Methodol. 1–38.
Dooley, B.A., Fitzgerald, A., 2012. My World Survey: National Study of Youth Mental Health in Ireland. Headstrong and UCD School of Psychology.
