The efficacy of interventions for test anxiety in university students: A protocol for a systematic review and meta-analysis

Christopher D. Huntley[a]*, Bridget Young[b], Vikram Jha[a], & Peter L. Fisher[b]

[a] School of Medicine, Faculty of Health and Life Sciences, University of Liverpool, UK

[b] School of Psychological Sciences, Faculty of Health and Life Sciences, University of Liverpool, UK

* Corresponding author

School of Medicine, Whelan Building, Brownlow Hill, University of Liverpool, Liverpool, UK, L69 3GB

Email: C.Huntley@liverpool.ac.uk

Tel: +44 (0)151 7945279

Fax: +44 (0)151 7945377

Funding: None

Conflict of Interest: None

Contributions: CH is the guarantor. CH drafted the manuscript with substantive contributions from BY, VJ, and PF. All authors contributed to the development of selection criteria, and search, risk of bias, data extraction and data analytic strategies. All authors read, provided feedback, and approved the final manuscript.

Abstract

The relative efficacy of interventions for test anxiety in university students is unknown. Previous meta-analyses have reported treatment outcomes across all age groups rather than specifically reporting on the effects for university students. Therefore, a new review is needed to determine the efficacy of psychological, educational, and pharmacological interventions for university students and the potential moderators of intervention efficacy. We present a protocol for a new review, which will enable a balanced and up-to-date appraisal of interventions in this field. Results from this review will therefore be of considerable interest to both students and those involved in the provision of interventions. Systematic review registration: PROSPERO CRD42016035859.

Keywords: Test anxiety, treatment, intervention, review, meta-analysis, university students

1. Introduction

Anxiety about examinations – test anxiety – has long been recognised as a pervasive and serious problem (Sarason, 1984, 1988; Spielberger, Anton, & Bedell, 1976). It is estimated that 20-25% of university students are highly test-anxious (Hahne, 1999, cf. Neuderth, Jabs, & Schmidtke, 2009; Hill & Wigfield, 1984; Naveh-Benjamin, Lavi, McKeachie, & Lin, 1997; Saravanan, Kingston, & Gin, 2014). Test anxiety is conceptualised as a situation-specific personality trait in which an individual experiences anxiety before, during, and after a performance-evaluative situation (Spielberger & Vagg, 1995). Test anxiety consists of two dimensions; a cognitive dimension labelled 'Worry' and an affective dimension labelled 'Emotionality' (Liebert & Morris, 1967; Spielberger, 1980). Worry consists of perseverative thinking about the personal and social consequences of failing to obtain one's performance goals and attentional distraction from the task-at-hand (e.g. test-irrelevant thinking). Emotionality refers to the perception of affective and physiological arousal in performance-evaluative situations.

Test anxiety is associated with poorer examination performance (Hembree, 1988; Seipp, 1991; Richardson, Abarham, & Bond, 2012). High-test-anxious (HTA) individuals typically score a half a standard deviation less than their low-test-anxious (LTA) peers in examinations (Seipp, 1991). Worry is more strongly associated with poorer exam performance than Emotionality (r = -.30 vs. -.15) (Hembree, 1988; Seipp, 1991). HTA students also report poorer mental health (Depreeuw & DeNeve, 1992) and are more likely to repeat years of study or dropout than their LTA peers (Schaefer, Mattheß, Pfitzer, & Köhle, 2007 cf. Neuderth, Jabs, & Schmidtke, 2009). Moreover, in a study investigating clinical manifestations of test anxiety, over 50% of participants had comorbid depression or a specific phobia (Herzer, Wendt, & Hamm, 2014).

Several psychological and educational interventions have been used to reduce test anxiety, and in turn improve academic performance. Psychological interventions aim to reduce anxiety based on the premise that anxiety is the major contributor to poor test performance. In contrast, educational interventions state that HTA students have limited study and test-taking skills and it is the failure to adequately prepare for, and undertake, a test that gives rise to decreased test performance. Here, test anxiety arises from a student's awareness that they are ill prepared to meet the demands of the test. The aim of educational-based interventions then is to improve test-taking and learning skills of students.

Three main psychological approaches have been used to treat test anxiety: behavioural therapy (BT), cognitive therapy (CT), and cognitive-behavioural therapy (CBT). BT, such as systematic desensitization (Wolpe, 1958) or applied relaxation (Bernstein & Borkovec, 1973), principally targets the somatic symptoms of test anxiety, whilst CT, such as rational-emotive therapy (Ellis, 1962) or schema-based therapy (Beck, Emery, & Greenberg, 1985), targets cognitive beliefs and structures for modification. CBTs combine treatment principles from both BT and CT approaches.

There are two main forms of educational intervention for test anxiety; 'Test-wiseness' training that focuses on the skills needed to take an examination (e.g. surveying a written test to ascertain which questions are given greater marks and so require more time and attention) and Study Skills Training that focuses on ways of learning and encoding material in the most optimal way (e.g. more emphasis on deeper levels of processing of to-be-learned material and less emphasis on rote memorization) (e.g. Culler & Holahan, 1980; Kirkland & Hollandsworth, 1979). Both forms are frequently combined into one intervention package.

The interventions discussed above have all been examined in controlled research studies. Randomized controlled trials (RCTs) represent the gold standard in determining

treatment efficacy as randomization, given sufficient numbers, ensures that both known and unknown prognostic factors are evenly distributed across treatment conditions, and therefore enable causal inferences to be drawn. However, many of the trials conducted in the test anxiety literature have had low statistical power to detect treatment effects as a result of using relatively small sample sizes – frequently less than 20 participants per treatment arm. This raises the risk of rejecting potentially useful treatments (a Type II error). Meta-analysis represents the best methodology for overcoming this problem; by systematically synthesizing trial outcome data, it increases statistical power and permits the calculation of more precise estimates of treatment effects.

1.1 Previous meta-analytic reviews

Two meta-analytic reviews of psychological and educational interventions for test anxiety have been conducted that included controlled trials with university student cohorts. In reporting the findings from these reviews, we focus on results pertaining to university students.

The first meta-analytic review (Hembee, 1988), analysed psychological and educational test anxiety outcome studies published between 1952 and 1986. Outcome studies had to include one or more active treatment conditions compared against a control condition. One hundred and thirty seven studies were included in the analyses (n = 7,641), though it is not known how many of these studies had university student samples. Between-groups effects sizes comparing active treatments versus control conditions were calculated, using Cohen's *d*, and effects were synthesized using a fixed-effect model. No overall mean effect was produced for all psychological and educational interventions, but a total of 14 summary effects were produced for differing treatment approaches. However, only four effect sizes

were produced that included studies whose samples consisted exclusively of university students: cognitive-behavioural therapy (CBT) (d = 0.87), 'taped individual' systematic desensitization (d = 0.59), 'other styles' of systematic desensitization (d = 1.08), and SST (d = 0.14). Examination of follow-up data revealed that BT (d = 1.21) and CBT (d = 0.96) were superior to untreated control groups. However, the median follow-up period for studies in this analysis was just 6 weeks. Finally, it was also found that participants receiving an intervention showed improved academic performance at post-treatment (d = 0.42) and at follow-up (d = 0.45) relative to those in the untreated control conditions, though it is not known which studies, and therefore samples, were included in these analyses.

The most recent meta-analytic review (Ergene, 2003) of interventions for test anxiety included psychological and educational outcome studies published between 1966 and 1998. Outcome studies had to include one or more active treatment conditions compared against a control condition. However, in contrast to Hembree (1988), random assignment of participants to conditions was explicitly stated in the eligibility criteria. Fifty-six studies were included in the analyses (n = 2,428), but again it is not known how many of these studies focused solely on university samples. Between-groups effects sizes comparing active treatment against a control condition were calculated using Cohen's d, and summary effects were synthesized using a fixed-effect model. An overall mean effect size of d = 0.68 (95% CIs 0.59 – 0.77) was computed across all interventions for university students but it is not reported how many studies – and what active conditions – contribute to this analysis. Though summary effect sizes were calculated for different treatment approaches (e.g. BT, CT), these were produced by synthesizing data from all age groups (from primary school to university student samples). Durability of treatment effects was not examined as no follow-up data was reported nor was the effect of the interventions on academic performance examined.

1.1.1 A Brief Critique Previous Reviews

We have identified several reporting and methodological problems with the previous meta-analytic reviews that limit the conclusions that can be drawn from them.

Firstly, the methods and results of both reviews are inadequately reported when viewed through the prism of modern reporting standards, such as the Preferred Reporting Standards for Systematic Reviews and Meta-analyses (PRISMA; Moher et al., 2009). The most salient example is that no details of the included studies were presented in either review. Thus, basic descriptive details about included studies, such as who conducted them or what the primary outcome measure was, were not presented. Similarly, neither review provides sufficient detail about how studies were located and selected, with eligibility criteria not presented in the PICOS (Participants, Interventions, Comparators, Outcomes, and Study design) format and no complete search protocol for at least one electronic database available for inspection. This latter point is now considered standard practice in order to enable replication of the search protocol. Finally, neither review presented a flowchart showing each stage of the screening process so readers are not informed about the flow of information through the screening phases and have no information about why studies may have been excluded.

Secondly, when it came to the pooling of study data, both reviews are inconsistent with their data-analytic strategy. For example, Hembree (1988) presented data on post-treatment efficacy for university students for four of the 11 interventions included, but the other seven intervention effect sizes synthesized data from studies across all age groups, from elementary school children all the way up to university students. Similarly, whilst Ergene (2003) provided an overall effect size for all interventions targeted at university students, no data was presented for specific interventions targeting university students; rather data was

synthesized across all age groups. As a result of this inconsistent pooling, it is not possible to make a balanced appraisal of interventions for each age group.

Thirdly, both reviews used a fixed-effect model that assumes there is a common 'true' effect size underlying all studies and that any variation between observed study effects is due to sampling error *only*. In essence, both reviews are making the assumption that the studies they pool together for synthesis are functionally identical. However, this assumption is untenable as studies differed across a whole range of domains, such as the setting, the baseline severity of the sample, the implementation of interventions, and the constitution of the samples as discussed above. A random-effects model would be preferable as it assumes included studies represent a random sample of potential studies and the goal is to estimate the mean of this (normal) distribution of effects. This model permits one to generalize from the observed study samples to similar populations, whereas a fixed-effect model speaks only to the efficacy of included studies.

Finally, there are several other methodological issues which impact on the internal and external validity of the results presented in both reviews. Firstly, no protocols were available prior to the conduct of the reviews, so it is not known which decisions were made *a priori* and which were *post hoc*. Secondly, risk of bias was not assessed by Ergene (2003), and whilst Hembree (1988) did examine 'study quality' across eight domains using an idiosyncratic scale from 1 ('poor') to 3 ('excellent'), no sensitivity analyses were conducted to examine the influence of poor quality studies on summary effects. This last point is particularly pertinent to the Hembree (1988) review as 'random assignment' of participants to interventions or control conditions was not explicitly stated in the inclusion criteria. Therefore, a sensitivity analysis could not be conducted to examine whether omitting studies without random assignment changes estimates of efficacy.

Clearly, both reviews have significant limitations, with neither meeting modern reporting standards. As such, the relative efficacy of psychological interventions for test anxious university students is currently unknown. Additionally, though it is clear that psychological and educational interventions exist for test anxiety, it is also not known if the efficacy of pharmacological intervention(s) for test anxious students has been examined. Therefore, a new review is required to ascertain the relative efficacy of interventions specifically for university students and present an up-to-date summary of the current best evidence. We present a protocol for such a review. To ensure our efforts are not duplicating that of another similar review, we checked the PROSPERO and Cochrane Database of Systematic Reviews (CDSR) databases first, finding no other review for test anxiety registered (correct as of 26th February 2016).

1.2 Review Objectives

Our main objective is to conduct a systematic and meta-analytic review that examines the relative efficacy of psychological, educational, and pharmacological interventions for (i) reducing test anxiety at post-treatment and follow-up, and (ii) improving examination performance in university students. Our secondary objective is to investigate potential moderators of intervention response.

2. METHODS

The proposed systematic review will be conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009) statement. The review protocol has been registered in the

PROSPERO database (CRD42016035859). This protocol follows the PRISMA-P reporting guidelines (Shameer et al., 2015). Please see Table 1 for an overview of the review tasks and author involvement.

2.1 Study Eligibility Criteria

2.1.1 Types of participants

We will include studies whose samples comprise solely of test anxious university undergraduate students. No structured diagnostic schedule is available to identify this cohort, so we will rely on self-report measures of sound psychometric properties for assessing test anxiety severity.

2.1.2 Types of interventions (and comparators)

Psychological, educational, and pharmacological interventions for test anxiety will be eligible for inclusion. Alternative therapy interventions, such as hypnosis and homeopathy, will be excluded. Comparators can include other psychological, educational, pharmacological or control conditions.

2.1.3 Types of Outcomes

The primary outcome is test anxiety severity, as measured by scores on any psychometrically reliable and valid self-report measure (continuous data). Self-rated measures were chosen as no structured diagnostic schedule or other standardized clinician-rated tool is available to assess test anxiety.

The secondary outcome is academic performance, as measured by scores on 'real world' examinations (continuous data). This outcome was selected as we wish to know if functional improvement (i.e. improved test performance) is demonstrated following an intervention. As our focus is on functional improvement related specifically to test anxiety, we are only interested in the effects of interventions on students' actual examination performance rather than laboratory tasks or other indices of academic performance.

2.1.4 Types of studies

Only studies that use random assignment to active treatment(s) and/or control condition(s) are eligible.

2.2 Search and Selection Methods for Identification of Studies

2.2.1 Search Strategy and Information sources

We will search the following databases: the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, PsycINFO, Scopus, Web of Science (Science and Social Science Index), and Educational Resources Information Center (ERIC). Additionally, we will search through the reference lists of the previous reviews of test anxiety and also search the reference lists of the articles obtained from our search. We will also contact the authors of the previous reviews to ensure we obtain as many studies as possible. Only published studies in peer-review journals will be eligible for inclusion. No date limits will be placed on the search. Searches will be carried out by CH. Titles, abstracts, and keywords will be searched in the databases using the search strategy exemplified in Table 2.

2.2.2 Study Selection

All article references will be combined into a single database. Duplicates will be removed first. CH and a postgraduate student will screen titles and abstracts for relevance. Full-text reports of all potential studies will then be accessed and their eligibility assessed independently by CH and PF. Discrepancies will be discussed, with BY or VJ arbitrating if consensus is lacking. All decision-making will be documented.

2.3 Data Extraction and Risk of Bias Methods

2.3.1 Data extraction

Data (e.g. means, SDs) will be extracted into predefined tables by CH. Study characteristics will also be extracted by CH, including: the type of treatment offered (including control conditions), the treatment manual (or its equivalent) used, treatment duration, number of sessions and numbers of contact hours (psychological, educational interventions)/medication dosage (pharmacological interventions), the primary outcome measure(s), follow-up points, baseline demographics (age, gender, pre-treatment test anxiety score), details of the site of study (i.e. what country the study was conducted in), and study attrition rates.

We will also report if studies have attempted to quantify the clinical significance of their trial results, including what methodology was used and the outcomes of this analysis.

Finally, we will document if studies have reported any adverse events. We expect data will be limited in this regard, so we will document how an adverse event has been defined and the numbers reported.

2.3.2 Risk of Bias Assessment

Risk of bias will be independently assessed by CH and a postgraduate student using the Cochrane Collaboration's Risk of Bias tool (Higgins et al., 2011). This tool assesses the following domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, attrition bias, reporting bias. Each domain is rated as having a low, unclear, or high risk of bias. Any additional threats to study validity will be noted.

2.4. Data Synthesis

Provided there are sufficient studies meeting our *a priori* selection criteria, data will be synthesized meta-analytically. Our analytic strategy is detailed below.

2.4.1 Meta-analytic Strategy

We will calculate the relative efficacy of psychological, educational, and pharmacological interventions for test anxiety as compared to control conditions.

We will combine continuous data from different outcome measures to allow the calculation of a standardized mean difference metric, Hedges' g and 95% confidence intervals for each intervention approach. A random-effects model will be used as we assume

that the studies included in our analyses merely represent a random sample of the distribution of effects, so that our combined summary effect represents an estimate of the mean effect of this distribution.

We plan to conduct our analyses on intention-to-treat data. We will assume those leaving an intervention had an unchanged outcome.

2.4.2 Subgroup Analyses

We will examine the relative efficacy of specific intervention approaches against control conditions, and compare their relative efficacy, if there are sufficient studies to allow this. This will permit a more fine-grained analysis of intervention efficacy. We will also examine the relative efficacy of interventions against other active interventions.

2.4.3 Heterogeneity, Sensitivity Analyses, Moderators, and Publication Bias

Heterogeneity will be assessed using the I^2 statistic. We consider values above 50% indicative of moderate or more heterogeneity (Higgins et al., 2003).

We will conduct a sensitivity analysis by examining whether the exclusion of studies with greater risk of bias affects the effect sizes and comparisons between groups.

Meta-regression will be used to explore moderators of treatment efficacy, provided there are sufficient studies to do so. The following moderators will be investigated: pre-treatment severity (self-report scores on measures of test anxiety), treatment dosage (for pharmacological interventions) number of hours contact time (for psychological and educational interventions), and intervention modality (individual vs. group). Additionally, the

use of treatment manual or its logical equivalent (manual vs. no manual) will be investigated as a moderator of treatment response for psychological and educational interventions.

Publication bias will be assessed using Egger's test (Egger et al., 1997) and funnel plots if there are a sufficient number of studies (≥ 10).

2.5. Quality of evidence

Ratings of evidence quality and our recommendations based on this will be informed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE; Guyatt et. al., 2008) system. All members of the review team will complete the GRADE online learning modules (http://cebgrade.mcmaster.ca/training/). Rationales for final decisions and recommendations will be documented. It is intended that this more detailed information will be made available as supplementary material to increase the transparency of our workings.

3. Discussion

The proposed review will elucidate the efficacy of psychological interventions for test anxiety for university students, and will be the first review to do so in a systematic way. Many trials have been conducted in the 13 years since the last meta-analytic review (Ergene, 2003), so this review will highlight current best evidence for interventions for test anxiety in university students. This review is therefore important for students and those involved in the provision of interventions (i.e. academic staff, clinicians), and will also highlight future research opportunities. Additionally, we hope by publishing this protocol and by using the

methods outlined here, the risk of bias to our analyses and recommendations will be minimized.

REFERENCES

Beck, A. T. (1975). *Cognitive therapy and the emotional disorders*. Madison, CA: International Universities Press Inc.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York, NY: Basic Books.

Bernstein, A. D., & Borkovec, T. D. (1973). *Progressive relaxation training*. Champaign, IL: Research Press.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons, Ltd.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7-18.

Chapell, M. S., Benjamin Blanding, Z., Takahashi, M., Silverstein, M. E., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268-274.

Culler, R. E., & Holahan, C. J. (1980). Test anxiety and academic performance: The effects of study-related behaviors *Journal of Educational Psychology, 72*(1), 16-20.Curtin, F., Elbourne, D., & Altman, D. G. (2002). Meta-analysis combining parallel and cross-over clinical trials. III: The issue of carry-over. *Statistics in Medicine, 21*(15), 2161-2173.

Depreeuw, E., & De-Neve, H. (1992). Test anxiety can harm your health: Some conclusions based on a student typology. In D. G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.), *Anxiety: Recent developments in cognitive, psychophysiological, and health research* (pp. 211–228). Washington: Hemisphere.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal*(315), 629-634.

Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York, NY: Lyle Stuart.

Ergene, T. (2003). Effective interventions on test anxiety reduction - A meta-analysis. *School Psychology International, 24*(3), 313-328.

Eysenck, M. W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences, 50*(7), 955-960.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353.

Guyatt, G. H., Oxman, A. D., Kunz, R., Falck-Ytter, Y., Vist, G. E., Liberati, A., & Schünemann, H. J. (2008). GRADE: Going from evidence to recommendations. *BMJ, 336*(7652), 1049-1051.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety *Review of Educational Research, 58*(1), 47-77.

Herzer, F., Wendt, J., & Hamm, A. O. (2014). Discriminating Clinical From Nonclinical Manifestations of Test Anxiety: A Validation Study. *Behavior Therapy, 45*(2), 222-231.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ (Online), 343*(7829).

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560.

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it *Elementary School Journal, 85*(1), 105-126.

Kirkland, K., & Hollandsworth, J. G. (1979). Test anxiety, study skills, and academic performance. *Journal of College Student Development, 20*(5), 431-436.

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety - a distinction and some initial data *Psychological Reports, 20*(3P1), 975-985.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery, 8*(5), 336-341.

Neuderth, S., Jabs, B., & Schmidtke, A. (2009). Strategies for reducing test anxiety and optimizing exam preparation in German university students: a prevention-oriented pilot project of the University of Wurzburg. *Journal of Neural Transmission, 116*(6), 785-790.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin, 138*(2), 353-387.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference - reactions to tests *Journal of Personality and Social Psychology, 46*(4), 929-938.

Sarason, I. G. (1988). Anxiety, self-preoccupation and attention. *Anxiety Research, 1*, 3-7.

Saravanan, C., Kingston, R., & Gin, M. (2014). Is test anxiety a problem among medical students: A cross sectional study on outcome of test anxiety among medical students? *International Journal of Psychological Studies, 6*(3), 24-31.

Schaefer, A., Mattheß, H., Pfitzer, G., & Köhle, K. (2007). Mental health and performance of medical students with high and low test anxiety. *Psychotherapie, Psychosomatik, Medizinische, Psychologie, 57*(7), 289-297.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27-41.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., . . . Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *Systematic Reviews, 4:1*.

Spielberger, C. D. (1980). *Preliminary professional manual for the Test Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., & Vagg, P. R. (1995). Test anxiety: A transactional process model. In C. D. Spielberger & P. R. Vagg. (Eds), *Test anxiety: Theory, assessment, and treatment* (pp.3-14). London: Taylor & Francis Ltd.

Szafranski, D. D., Barrera, T. L., & Norton, P. J. (2012). Test anxiety inventory: 30 years later. *Anxiety Stress and Coping, 25*(6), 667-677.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*(2), 92-104.

Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Standard University Press.

Zeidner, M. (1998). *Test anxiety: State of the art*. London: Kluwer Academic Publishers.

Table 1. Proposed timeline and author involvement

Stage	**Month**						**Author Involvement**
	1	**2**	**3**	**4**	**5**	**6**	
Search							CH
Screening							CH, PF, BY, VJ, Postgraduate
Analysis							CH, PF, BY, VJ
Write-up							CH, BY, VJ, PF

Table 2. Example of search strategy for Scopus database

Step	Search Terms
1	"Test anxiety" OR "Examination anxiety" OR "Test stress" OR "Examination stress" OR "Test performance" OR "Examination performance" OR "Evaluation Stress" OR "Evaluative Stress" OR "Evaluative Anxiety" OR "Evaluation Anxiety" OR "Test fear" OR "Examination fear"
2	College student OR University student OR undergraduate
3	Randomi?ed controlled trial OR Controlled clinical trial OR Clinical trial OR PHASE* Trial OR RCT OR randomi?ed OR random allocation OR randomly allocated OR Controlled trial OR placebo OR outcome study
4	Psycho* or behavi* or cogniti* or metacogniti* or implosive or desensiti* or interpersonal or gestalt or attenti* or activation* or rational or stress?innoculation or bibliotherapy* or counsel* or supportive or study skill* or skill?focus* or testwise* or relax* or training* or non?directive or guided or imag* or computer* or CBT
5	#1 AND #2 AND #3 AND #4
6	Limit #5: English Language

Note: Scopus search is to be limited to the Subject Areas: 'Life Sciences, Health Sciences, Social Sciences and Humanities' subject areas, with all document types permissible. Each terms is searched under the 'Article, Title, Abstract, Keywords' heading.