Improving academic performance and mental health through a stress management intervention: Outcomes and mediators of change

Edmund Keogh\textsuperscript{a,*}, Frank W. Bond\textsuperscript{b}, Paul E. Flaxman\textsuperscript{b}

\textsuperscript{a}Department of Psychology, University of Bath, Bath, BA2 7AY, UK
\textsuperscript{b}Department of Psychology, Goldsmiths College, London, SE14 6NW, UK

Received 22 October 2004; received in revised form 1 February 2005; accepted 2 March 2005

Abstract

Two hundred and nine pupils were randomly allocated to either a cognitive behaviourally based stress management intervention (SMI) group, or a non-intervention control group. Mood and motivation measures were administered pre and post intervention. Standardized examinations were taken 8–10 weeks later. As hypothesized, results indicated that an increase in the functionality of pupils’ cognitions served as the mechanism by which mental health improved in the SMI group. In contrast, the control group demonstrated no such improvements. Also, as predicted, an increase in motivation accounted for the SMI group’s significantly better performance on the standardized, academic assessments that comprise the United Kingdom’s General Certificate of Secondary Education. Indeed, the magnitude of this enhanced performance was, on average, one-letter grade. Discussion focuses on the theoretical and practical implications of these findings.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Stress intervention; Examinations; Anxiety; Cognitive behaviour therapy
Introduction

The majority of children in the United Kingdom (UK) take the General Certificate of Secondary Education (GCSE) examinations around the age of 15–16 years. They are comprised of tests, and classroom-based assignments, both of which are nationally standardized. Potential employers use students’ overall GCSE grades as an important indicator of basic competence and these scores also largely determine whether or not children go onto higher education. Given the importance placed on GCSE results, it is perhaps not surprising to discover that many children report a high level of mental strain and worry when preparing for, and taking, GCSE exams (e.g., Denscombe, 2000). UK children are, of course, not the only ones to suffer such concerns. Studies in the United States (US) estimate that a large proportion of children report examination stress as well (e.g., Hill, 1984; Hill & Wigfield, 1984). For example, Hill (1984) estimates that between 25% and 30% of US students suffer adverse effects from examination stress and as many as 10 million school students underachieve because of anxiety-related performance impairments. Given such stark figures, it is perhaps not surprising that a branch of psychological research, known as test anxiety (TA), has developed to investigate this phenomenon (e.g., Sarason, 1980, 1984; Spielberger & Vagg, 1995; Zeidner, 1998).

The nature of test anxiety

TA is viewed as a ‘situation-specific’ form of anxiety that predisposes people to perceive evaluative situations as threatening (Spielberger, 1972). Individuals high in TA are more likely to experience frequent and intense levels of state anxiety, which are in turn accompanied by worry and other test irrelevant thoughts, when being examined (Spielberger, Gonzalez, Taylor, Algaze, & Anton, 1978). Thus, TA consists of two distinct response sets: worry and emotionality. The former describes a cognitive response, where attention is focused on concerns about one’s performance, ability or adequacy (Deffenbacher, 1980), and the latter refers to an affective response involving an over-awareness of bodily arousal and tension in the face of evaluative situations (Sarason, 1984).

Research indicates that individuals high in TA frequently experience performance decrements in evaluative situations, and that the cognitive component of TA (worry) contributes most to these deficiencies (Eysenck & Calvo, 1992; Hembree, 1988; Sarason, 1980, 1984, 1988; Wine, 1971). For example, when Deffenbacher (1980) controlled for common variance between worry and emotionality, only worry was negatively associated with academic performance. Furthermore, we have recently shown that the worry component of TA best predicts examination performance amongst undergraduates (Keogh, Bond, French, Richards, & Davis, 2004).

Others argue that the cognitive interpretation, or appraisal, of emotionality determines the extent to which emotional arousal facilitates or debilitates performance (Hollandsworth, Glazeski, Kirkland, Jones, & Van Norman, 1979; Sarason, 1984, 1988). For example, Hollandsworth et al. (1979) found that during a mental ability task, arousal seemed to trigger distracting, task-irrelevant thinking (i.e., worry) in high TA individuals. In contrast, low test-anxious individuals spoke of getting ‘psyched up’ or of getting ‘in stride’ (i.e., engaging in task-relevant thinking). The importance of cognitive appraisal in evaluative contexts is further
supported by research that shows that objective measures of physiological arousal cannot distinguish high and low TA individuals (Holroyd & Appel, 1980).

Progress in conceptualizing test anxiety is reflected in measurement developments of this construct. Contemporary measures such as the Revised Test Anxiety Scale view test anxiety as a hierarchical construct that consists of lower order domains such as tension, worry, bodily symptoms and test-irrelevant thinking, as well as a single higher-order factor (Benson & Bandalos, 1992; Benson & El-Zahhar, 1994; Hodapp & Benson, 1997).

**Interventions for test anxiety**

Consistent with the apparent importance of cognition in TA, many interventions over the last 30 years have tried to change people’s cognitive appraisals, as a means of reducing or preventing this problem (Flaxman et al., 2002). These “cognitive change” interventions are typically referred to as cognitive-behavior therapies (CBTs), and research indicates that they can improve general study skills and exam performance in university students (e.g., Algaze, 1995; Vagg & Papsdorf, 1995). Furthermore, these CBT strategies seem to be equally effective in reducing the cognitive and emotionality symptoms of TA. Unfortunately, most interventions have tended to address either the cognitive or emotional components of TA, so it is unclear whether or not an intervention that targeted both components would be more successful (Zeidner, 1998).

There are also a number of methodological and theoretical limitations with many of the studies conducted to date. For example, Algaze (1995) had eight participants per group, which provided only a small chance of finding a positive impact for the SMI on examination scores, if, indeed, there was one to find. A second limitation is that the effectiveness of such interventions can be dependent on the time clients are allowed to practice and improve their new skills. Less practice may result in weaker changes in TA (Zeidner, 1998). Third, few studies examine whether SMIs, including CBTs, impact school children’s examination scores. Furthermore, of those that do, many use cognitive ability tests (see Zeidner, 1998), which may not necessarily determine future academic success. An improvement in assessing the impact of SMIs on school performance would be to use student grades based on reliable, nationally standardized examinations that have a large impact upon students’ academic progression. Examples of such examinations in the UK would be GCSE or Advanced-level exams. We are not aware of any SMI outcome study that has used such a rigorous, and potentially important, performance criterion.

Apart from small drops in TA, it also remains unclear whether such SMIs have a general impact on the mental health of students. Most studies simply do not consider general mental health. However, SMIs in adult workers have led to improvements in performance and mental health. Specifically, CBT-based SMIs have significantly improved mental health, absenteeism levels, motivation, and performance, in the adult work environment. They also seem particularly effective when these SMIs have used different techniques to promote cognitive change (e.g., problem solving, relaxation, belief modification) (Bond & Bunce, 2000; Bunce & West, 1996; van der Klink, Blonk, Schene, & van Dijk, 2001). These multi-method CBT interventions do not appear to have been systematically investigated using students in a school setting; there are no theoretical reasons, however, as to why they would be ineffective in this context.

Importantly, we are not merely interested in determining whether or not a SMI is effective in improving mental health and maximizing GCSE performance. We are also interested in...
identifying the mechanisms (or mediators) by which such benefits occur, if indeed they do so. Such assessment of potential mechanisms of change is necessary, as interventions of all kinds can be effective for reasons different from those hypothesized (Bond & Bunce, 2000). Moreover, by identifying mediators of change, one can improve interventions by trying to manipulate these mechanisms more effectively and, thus, better help people (Bond & Bunce, 2001).

The SMI and its mediators of change

Cognitive therapy theory (e.g., Beck, 1976) and achievement goal theory (e.g., Ames, 1992; Elliot, 1997) are used as the theoretical underpinnings of the current CBT SMI. As such, they inform our predictions as to what variables will mediate the proposed effects that the SMI has on these two outcomes. Flaxman et al. (2002) detail the intervention, which is based largely on Beck’s (1976; Beck & Emery, 1985) version of CBT. It maintains that an intervention can best protect and promote people’s mental health if it changes their dysfunctional cognitions, or beliefs (e.g., “I am bad at taking tests”) to functional ones (e.g., “I can take tests, if I prepare appropriately”). This multi-method SMI attempts to produce such cognitive change by showing people how their feelings and actions are linked to their cognitions, and most importantly, how they can intervene to make these beliefs more functional. Thus, we predict that the SMI will improve students’ mental health because it makes their cognitions more functional.

In addition to making beliefs more functional, the present multi-method SMI also uses didactic and experiential learning exercises in order to teach participants how to move their attention from their worries, self-defeating thoughts and images to their course material (see Flaxman, et al., 2002). These attention-directing exercises (see Borkovec, Hazlett-Stevens, & Diaz, 1999), in combination with problem solving, imagery, and progressive relaxation training, aim to develop, within participants, a learning goal set. Achievement Goal Theory maintains that people can improve their motivation to learn and gain knowledge by working under such a “learning-goal” set (Covington, 2000). This is a learning strategy in which the aim of knowledge acquisition is to increase one’s competency, understanding, and appreciation of a subject area, as opposed to, for example, learning in order merely to outperform one’s peers (e.g., Elliot & Church, 1997). Working under a learning goal set is thought to promote excitement and task absorption, which, in turn, helps people to process relevant information strategically and to a deep-level thus enabling increased school achievement (e.g., Covington, 2000; Elliot & Church, 1997). Research has generally supported this motivational theory of performance (e.g., Elliot, McGregor, & Gable, 1999; Meece & Holt, 1993; Schunk, 1996). Such a learning set, according to achievement goal theory, should function to motivate students to focus on, retain, and become more competent and knowledgeable about their course-related content. We predict, therefore, that an increase in motivation will mediate the relationship between our CBT SMI and GCSE performance.

The first aim of the current study was, therefore, to examine the effectiveness of an SMI administered during an academic school year on GCSE performance and the general mental health of 15–16 year old school children. We also sought to answer theoretically important, and to date unanswered, questions regarding the potential mechanisms that may explain any change found. We specifically hypothesized that this randomized controlled outcome experiment will show that:
1. school children who received a multi-method CBT SMI will have better mental health (including reductions in test anxiety) and GCSE scores than those in a control group that did not receive this intervention;
2. the increase in the functionality of students’ beliefs will mediate improvements in their mental health that the SMI produced;
3. higher levels of motivation will account for the difference in GCSE results that will be seen between the SMI and control groups.

**Method**

**Design**

To investigate the efficacy of our multi-method CBT SMI we used a repeated measures, randomized, matched-pairs design. Intervention group (intervention vs. no intervention) served as the between-groups variable, whereas time of testing served as the within-groups variable (before vs. after intervention). The dependent variables were various self-report measures of stress and exam performance. The latter was measured by students’ total GCSE grade points. Finally, in order to identify the mechanism of any differences in examination performance between the two groups, mediation analysis through regression was conducted. Specifically, we followed the recommendations made by Baron and Kenny (1986) and used multiple regression to determine if dysfunctional attitudes and motivation mediated any differences between students in the intervention and control group on the core outcomes, mental health and examination performance, respectively.

**Participants**

Participants came from the same school year of a large, co-ed, secondary school in the UK. A total of 209 pupils (54% male), aged between 15 and 16 (mean = 15.57; S.D. = .50), completed a set of questionnaires (described below). Due to financial constraints, we could only administer the SMI to a limited number of students; therefore, we used a systematic sampling procedure (Pedhazur & Schmelkin, 1991) to select 160 of these students, matching on intelligence. Also, given that national GCSE statistics indicate that females score slightly higher than males (Department for Education and Skills, 2005), we felt it was important to match groups on gender, as well. We then randomly allocated this sample either to the SMI or the control group. We divided the eighty members of the SMI group (40 males and 40 females) into eight groups of ten participants. These SMI training groups met once a week for ten weeks, with a one-week break after week four for a school holiday.

**The examination SMI group**

Full details of the SMI have been described elsewhere and so only a brief outline will be provided here (Flaxman et al., 2002). The SMI involved didactic teaching components as well as learning exercises and homework assignments. Table 1 contains an overview of the weekly topics,
all of which were facilitated by the same therapist (PF). Participants attended 10 weekly, 1-hour workshops during school hours, missing one of their timetabled lessons each week. Each workshop had a maximum of 10 students, and the days and times of the weekly sessions were alternated to ensure that participants would not be absent from the same curriculum subject each week.

The SMI is similar to existing technically eclectic CBT approaches such as Stress Inoculation Training (SIT; Meichenbaum, 1985) and Cognitive Behaviour Modification (CBM; Meichenbaum, 1977); the first few sessions involved a conceptualization of evaluation strain within a cognitive behavioural model, and similar to SIT, included a combination of relaxation training and cognitive change strategies. Our training protocol differs, however, from these approaches in that we also included a major focus on worry, and its management, in later sessions. Specifically:

Session 1 & 2: defined stress and anxiety, and introduced the idea that anxiety can have bodily, emotional, cognitive, and behavioural manifestations. These first two sessions also attempted to socialize students into the cognitive-behavioural model (Wells, 1997), which in turn serves as an explanatory model for understanding the nature and treatment of stress and anxiety, and how such responses can become maladaptive.

Session 3: focused on the nature of the interplay between stress and response systems, with particular emphasis on the importance of thoughts, beliefs and interpretations (i.e., the cognitive response modality). Participants are taught methods for challenging problematic modes of thinking and are given the opportunity to practice these methods in session. The third session, therefore, aimed to develop an awareness of dysfunctional cognitions, and the negative impact that they can have, as well as help participants learn how to assess the ‘usefulness’ of such cognitions.

Session 4: introduced full progressive muscle relaxation (PMR) techniques as a means of reducing the unhelpful aspects of the bodily response, such as excessive tension, light-headedness and trembling. An important issue in this respect is that the students’ attention can be drawn to such bodily manifestations of stress at the expense of focusing on the task at hand. Thus, the relaxation exercises are discussed alongside instruction on diverting attention back toward dealing with evaluative situations. A full discussion of the relaxation techniques used in these sessions can be found elsewhere (e.g., Bernstein & Borkovec, 1973; Öst, 1987).

<table>
<thead>
<tr>
<th>Session</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EDUCATION PHASE I—A definition of stress</td>
</tr>
<tr>
<td>2</td>
<td>EDUCATION PHASE II—When stress becomes a problem</td>
</tr>
<tr>
<td>3</td>
<td>THOUGHTS, EMOTION, AND BEHAVIOUR</td>
</tr>
<tr>
<td>4</td>
<td>RELAXATION I—Progressive Muscular Relaxation (PMR)</td>
</tr>
<tr>
<td>5</td>
<td>RELAXATION II—Shortened relaxation techniques</td>
</tr>
<tr>
<td>6</td>
<td>WORRY I—The nature and function of worry</td>
</tr>
<tr>
<td>7</td>
<td>WORRY II—Modifying meta-beliefs</td>
</tr>
<tr>
<td>8</td>
<td>PROBLEM SOLVING</td>
</tr>
<tr>
<td>9</td>
<td>IMAGERY</td>
</tr>
<tr>
<td>10</td>
<td>REVIEW</td>
</tr>
</tbody>
</table>
Session 5: introduced additional relaxation techniques for coping with problematic bodily arousal, and constituted shortened versions of the full PMR method. The aim was for participants to progress to using these shortened techniques in everyday situations, and particularly in evaluation contexts (e.g., exams, social encounters).

Sessions 6 & 7: dealt with the role that the cognitive modality, specifically worry, plays in evaluation anxiety. The first worry session described the nature and functions of worry and inappropriate or poor problem solving (e.g., Tallis, Eysenck, & Mathews, 1991), whereas the purpose of session 7 was to modify negative and positive meta-beliefs about worry. The techniques used were adapted from Wells (1997) and include offering counter-evidence, showing that worries offer inaccurate representations of situations, highlighting the disadvantageous nature of worry as a coping strategy, and showing that situations can turn out positively when worry is not used.

Session 8: focused on problem solving training as an effective method for dealing with an unhelpful behavioural response to challenging situations. Participants were reminded that avoidance and procrastination are rarely useful strategies for coping, but that more positive problem solving can serve as a useful alternative to worry (i.e., as a strategy to reduce the impact of maladaptive cognitions). Focusing on solving problems step-by-step can help participants realize that they do not have to use worry to ensure successful outcomes. This process also serves to direct attention away from worry and back to the task at hand, further illustrating the controllable nature of worry.

Session 9: introduced guided imagery techniques, which have been shown to be effective in reducing both the worry and emotionality components of the problem (e.g., Sapp, 1994). For example, relaxation images can aid the relaxation process and thus be effective for reducing bodily arousal; coping images can provide an alternative (less catastrophic) view of a situation, thereby challenging irrational thinking patterns; and finally, mental practice techniques can help students prepare for challenging situations through the mental rehearsal of appropriate behavioural response patterns.

Session 10: provided participants with a review of the stress management techniques. The trainer provides a summary of how the different techniques impact upon the interacting response modalities.

The control group

Those who were allocated to the control group simply continued their education as per usual (i.e., they received no special attention). Our decision not to run a ‘placebo’-type intervention was driven by ethical concerns. The GCSE is extremely important within the UK education system and we were unwilling to withdraw students from important GCSE-related lessons in order to present them with an intervention that we designed to be ineffective. We acknowledge that, whilst this methodology is not ideal, it is nonetheless ethical and pragmatic, given the nature of the sample involved.

Questionnaire measures

General health questionnaire: 28-item version (GHQ; Goldberg & Hillier, 1979)

The GHQ was administered to assess general mental health (McDowell & Newell, 1996). The GHQ is one of the most widely used psychometric measures in health and psychiatry and has
good reliability and validity (Goldberg et al., 1997). Each item is on a four point Likert-type scale, which assesses how a person has been feeling over the past few weeks. Higher scores indicated greater degrees of mental ill-health. Alpha coefficients for this study were .92 and .95 for Times 1 and 2, respectively.

**Need for achievement scale (Motivation; Paspalanov, 1984)**

This scale comprises of 42 yes/no items, including 16 distractor questions, and is worded in such a way that it can be used in educational settings. It includes questions pertaining to typical experiences, evaluations and behavioural patterns with regard to work and learning (e.g., “I like to make long-term plans for my studies and achievements”). This motivation scale was chosen because we felt it assessed effectively achievement goal theory which, as noted, is the model that accounts for the proposed relationship between the SMI and GCSE performance. The wording of some of the items was amended slightly to make them more suitable for the age of our sample. The scale is reported to have good validity, and the reliability for this study was adequate (Time 1 = .65; Time 2 = .75).

**Revised test anxiety scale (TA; Benson & Bandalos, 1992)**

This measure of test anxiety is comprised of 25 items, which combine to form four sub-scales relating to tension, worry, bodily symptoms and test-irrelevant thinking, as well as a single higher-order factor (Benson & El-Zahhar, 1994). Students are asked to respond to their feeling when taking a test, and asked to rate each item based on how they generally feel using a four-point Likert rating scale (1 = almost never; 2 = sometimes; 3 = often; 4 = almost always). Confirmatory factor analysis indicates that the TA has excellent factor reliability across a variety of different sample groups (e.g., Benson & Bandalos, 1992; Benson & El-Zahhar, 1994; Hodapp & Benson, 1997). Alpha coefficients for the global TA scale were .89 and .91 for Times 1 and 2, respectively.

**Dysfunctional attitude survey (DAS; Weissman, 1979)**

Is a 40-item measure of the functionality of those beliefs that form a cognitive vulnerability to depression and general psychopathological conditions (e.g., “People will probably think less of me if I make a mistake”) (Segal & Shaw, 1988). Responses were recorded on a seven-point Likert scale labelled “Disagree totally” to “Agree totally”, scored 1 to 7 respectively. Cronbach alpha coefficients of .89 (Time 1), and .92 (Time 2) were obtained.

**Performance indicators**

**Cognitive abilities test (CAT)**

The Cognitive Abilities Test (CAT) is the most widely used assessment of reasoning ability employed in the United Kingdom (Strand, 2002). The test, which is used as a measure of intelligence for school pupils aged from 8–15, indicates a pupil’s verbal reasoning ability, quantitative reasoning ability, and non-verbal reasoning ability. Educational research has shown that pupils’ scores on the CAT are powerful indicators of subsequent attainment in the GCSE examinations (Fernandes & Strand, 2001). The pupils involved in this study completed the CAT in their previous school year.
Examinations

Our primary performance indicator was pupil’s formal GCSE examination results, which are standard national examinations that are used as indicators of academic competence. They are the culmination of the mandatory UK formal education system, and they determine whether pupils are able to progress onto further and then higher education. Each pupil takes a number of GCSEs in subjects such as English, Maths, and Science. Students are graded alphabetically from A+ to G, and for the purposes of the current study we converted each grade to a numerical score, ranging from 8 (A+) to 1G. A higher score should therefore be taken as indicative of better examination performance.

Procedure

At the start of the spring term (in January 2000), the parents of all pupils at the school were contacted and informed that a study into examination stress management was being conducted and asked for permission to allow their son or daughter to participate. Therefore, pupils and parents were generally aware of the aims of the study, although they were not aware of more specific hypotheses, such as those concerning the proposed mediators. All pupils who had parental permission to participate, and who wanted to do so, were then administered a battery of questionnaires (see above) in groups the week prior to the intervention starting. From the initial sample of 209 (and before any of the measures had been scored), 80 pupils were randomly allocated to the examination stress intervention group and so were invited to attend weekly stress management sessions for ten weeks. During the last week of the spring term (in March, 2000) all pupils were again asked to complete the second set of questionnaire measures. Pupils took their formal GCSE’s during the months of May and June 2000, approximately eight weeks after the SMI ended.

Results

As a result of participant attrition and consequent listwise deletion, the SMI group size was reduced to 40. This 50% attrition rate was due to participant absence at one of the two questionnaire administration points, or failure to attend to at least 60% (6 out of 10) of the training sessions (note that failure to attend occurred mostly towards the end of the intervention). In order to meet the statistical requirements of equal numbered groups, only 50% of the control group was selected (using systematic sampling procedures) for our analyses. This sampling procedure (see Pedhazur & Schmelkin, 1991) ensured that we retained a similar number of males and females for the analyses. Thus, we ensured that we not only had equal cell sizes, but that the number of males and females was similar across both groups (SMI = 18 males, 22 females; control = 20 males, 20 females).

A series of analysis of variance (ANOVA) tests were conducted on the Time 1 outcome and mediator variables with group at time 1 (intervention vs. control) and inclusion at time 2 (included vs. excluded) as the between group factors. There was a significant difference in levels of intelligence between those who were included and those who were not (F(1,193) = 29.69, p < .001), with those excluded having significantly lower levels of intelligence (mean CAT
score = 4.58) than those who were included in the main analyses (mean CAT score = 5.86). Consistent with this finding, a similar pattern was found for number of examinations taken (included = 9.04; excluded = 7.97; F(1, 201) = 19.42, p < .001). This suggests that those included in the final analysis were of slightly higher intelligence. Finally, there was a significant difference between those in the intervention (mean = 135.63) and the control group (mean = 124.53) on the dysfunctional attitudes scale (F(1, 185) = 6.49, p < .05).

Means and standard deviation for the measures taken for the final sample are presented in Table 2. No significant differences between groups were found for either age or intelligence.

### Bivariate relationships

Intercorrelations among all variables at both Time 1 (below the diagonal) and Time 2 (above the diagonal) are presented in Table 3. Consistent with our mediation hypotheses, high levels of dysfunctional attitudes were significantly associated with mental ill-health at both Times 1 and 2. Furthermore, high levels of motivation at both time points were significantly related to high levels of overall GCSE performance. Interestingly, test anxiety was not significantly related to GCSE performance, and so we examined the RTAS subscales (tension, worry, bodily symptoms and test-irrelevant thinking). At time 2, but not time 1, worry (r = -.24, p < .05) and test-irrelevant thinking (r = -.23, p < .05) were negatively correlated with GCSE points. This suggests that the cognitive component of test anxiety becomes important, when examined closer to examinations.

### Outcome variables

To examine the effectiveness of the SMI in improving mental health, 2 x 2 repeated measures analyses of variance (ANOVA) were performed on the mental health and test anxiety variables. Group (SMI vs. control) served as the between-groups variable and time (Time 1 vs. Time 2) as the within-groups variable. Where significant main or interaction effects were found,

<table>
<thead>
<tr>
<th></th>
<th>SMI</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time 1</td>
<td>Time 2</td>
</tr>
<tr>
<td>Age</td>
<td>15.68 (.47)</td>
<td>N/A</td>
</tr>
<tr>
<td>Intelligence</td>
<td>5.91 (1.40)</td>
<td>N/A</td>
</tr>
<tr>
<td>GHQ</td>
<td>24.64 (13.26)</td>
<td>18.69 (10.84)</td>
</tr>
<tr>
<td>Test Anxiety</td>
<td>41.81 (8.61)</td>
<td>40.54 (9.04)</td>
</tr>
<tr>
<td>Motivation</td>
<td>12.66 (4.23)</td>
<td>13.24 (4.46)</td>
</tr>
<tr>
<td>Dysf. Att.</td>
<td>129.98 (24.22)</td>
<td>123.10 (22.83)</td>
</tr>
<tr>
<td>GCSE points</td>
<td>N/A</td>
<td>59.73 (18.22)</td>
</tr>
</tbody>
</table>

SMI = stress management intervention group; GHQ = mental ill health; Test Anxiety = Revised Test Anxiety Scale; Dysf. Att. = dysfunctional attitudes; GCSE = overall GCSE performance points; *p < .05; **p < .01; ***p < .001.
all four possible simple (pairwise) comparisons were conducted. For between-group comparisons at Time 2, we entered the relevant Time 1 scores as a covariate. To control for inflated alpha, a Bonferroni correction was applied to all simple effects tests (Pedhazur & Schmelkin, 1991).

Finally, we conducted independent groups ANCOVA in order to determine the impact of the SMI on GCSE performance. Intelligence was entered as a covariate given the association it has with examination performance. Gender was also included as a covariate because of previous evidence that suggests sex differences in examination performance (Zeidner, 1998).

Mental ill-health (GHQ)
Although no significant main effects were found, the Group × Time interaction on GHQ scores was significant \( F(1, 74) = 6.26, p < .05, \eta_p^2 = .08 \). Simple main effects tests indicated that mental ill-health scores significantly improved (i.e., decreased) from Time 1 to Time 2 in the SMI group \( F(1, 38) = 17.75, \ p < .01; \eta_p^2 = .32 \), to a statistically large extent. Furthermore, at Time 2 (controlling for Time 1 scores) those in the intervention group were significantly healthier than those in the control group \( F(1, 73) = 6.94, p < .05; \eta_p^2 = .09 \), with a difference of a medium magnitude.

Test anxiety
There were no significant main, or interaction, effects for this variable.

Overall GCSE performance
An independent groups ANCOVA revealed a significant difference in performance between the SMI and the control groups for GCSE points \( F(1, 76) = 7.86, p < .01; \eta_p^2 = .10 \). This effect, of a
medium magnitude, revealed that the SMI group obtained significantly higher GSCE points (mean = 59.7) than did those in the control group (mean points = 51.3). In order to place this result in real terms, we calculated the grade point average (GPA) for each participant, which revealed that the magnitude of this between-group difference was equivalent to, on average, one-letter grade for all examinations taken i.e., SMI group GPA = 6.56 (B grade); control group GPA = 5.53 (C grade).

Mediator analysis

An additional aim of this study was to identify the mechanisms (or mediators) by which improvements in the outcome variables occur. Before formally testing for mediation, however, it can be instructive to examine the extent to which the mediator variables (i.e., motivation and dysfunctional attitudes) change over time, as a function of the group variable. To test for this group × time interaction, we conducted 2 × 2 ANCOVA on motivation (controlling for intelligence and gender) and ANOVA on dysfunctional attitudes.

For motivation, both covariates were significant (intelligence, \(F(1, 70) = 4.90, p < .05, \eta^2_p = .07\); gender, \(F(1, 70) = 5.74, p < .05, \eta^2_p = .07\), as was the main effect of group (\(F(1, 70) = 4.92, p < .05, \eta^2_p = .07\)). Furthermore, a significant Group × Time interaction was observed (\(F(1, 70) = 5.00, p < .05, \eta^2_p = .07\)). Simple effects analyses indicated that the SMI group had significantly higher levels of motivation at Time 2 than did the control group (\(F(1, 69) = 8.13, p < .05; \eta^2_p = .11\)), a difference of a medium magnitude. Regarding dysfunctional attitudes, neither of the main effects was significant, but the Group × Time interaction was found to be significant (\(F(1, 78) = 6.04, p < .05, \eta^2_p = .07\)). In particular, simple effects tests revealed a decrease in dysfunctional attitudes in the SMI group between Time 1 and 2, an effect that was just outside the adjusted level of significance (\(F(1, 78) = 4.44, p < .0125; \eta^2_p = .05\)). Furthermore, at Time 2, the SMI group reported lower levels of dysfunctional beliefs than did the control group, a difference that was small to medium in magnitude, but was also not significant at the corrected alpha level (\(F(1, 77) = 5.08, p < .0125; \eta^2_p = .06\)).

The above analyses demonstrated that our hypothesized mediator variables changed over time, as a function of the SMI. We need still, however, to assess the degree to which these changes actually explain the significant interaction effects that occurred for the mental health and performance outcomes. Baron and Kenny (1986) state that a variable functions as a mediator when the magnitude of a significant relationship between an independent variable (IV) and a dependent variable (DV) reduces, after controlling for the proposed mediator variable. If the magnitude of this relationship is reduced to non-significance, then the proposed mediator variable is said to mediate fully this association.

This operational definition of mediation is assessed through “four tests of mediation” that are recommended by Baron and Kenny (1986). These are as follows:

**Test 1**: Establish that the predictor variable (\(X\)) is correlated with the outcome variable (\(Y\)). This step determines if there is, indeed, an effect to be mediated.

**Test 2**: Establish that the predictor variable is correlated with the mediator. This step treats the proposed mediator (\(M\)) as an outcome variable.

**Test 3**: Establish that the mediator is correlated with the outcome variable, while controlling for the predictor variable. Such control is necessary, as the mediator and outcome variables could be
correlated, for the sole reason that they are both caused by the predictor variable. If such a correlation exists, then there is evidence for partial mediation.

**Test 4:** If $M$ completely mediates the $X - Y$ relationship, then this relationship should become non-significant, when controlling for $M$. The regression equation used for Test 3 is used to establish this effect.

As can be seen in Table 4, the significant effects of group (SMI vs. control) on changes in mental ill-health were reduced to non-significance, after controlling for change in dysfunctional attitudes. Similarly, the significant effects of group on GCSE performance were rendered non-significant, after controlling for change in motivation. Together, these findings are consistent with our mediation hypotheses: that the reduction in dysfunctional cognitions that occurred as a result of our SMI, fully mediated the improvements in mental health that was seen in this group; and, that the superior GCSE performance observed in the SMI group was fully mediated by changes in motivation.

Table 4
Hierarchical regression analyses for determining whether change in dysfunctional attitudes (Dysf. Att) mediates change in general mental health (GHQ), and whether change in motivation mediates change in GCSE performance, as a result of the SMI

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Test</th>
<th>Step</th>
<th>Predictor</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHQ Time 2</td>
<td>1</td>
<td>1</td>
<td>GHQ Time 1</td>
<td>.51***</td>
<td>.26***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Group</td>
<td>.25**</td>
<td>.32***</td>
<td>.07**</td>
</tr>
<tr>
<td>Dysf. Att Time 2</td>
<td>2</td>
<td>1</td>
<td>DAS Time 1</td>
<td>.72***</td>
<td>.52***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Group</td>
<td>.17*</td>
<td>.55***</td>
<td>.03*</td>
</tr>
<tr>
<td>GHQ Time 2</td>
<td>3 &amp; 4</td>
<td>1</td>
<td>GHQ Time 1</td>
<td>.51***</td>
<td>.26***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Dysf. Att Time 1</td>
<td>.05</td>
<td>.26***</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>Dysf. Att Time 2</td>
<td>.66***</td>
<td>.54***</td>
<td>.28***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Group</td>
<td>.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCSE points</td>
<td>1</td>
<td>1</td>
<td>Gender</td>
<td>.18*</td>
<td>.64***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intelligence</td>
<td>.80***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Group</td>
<td>-.18**</td>
<td>.68***</td>
<td>.03**</td>
</tr>
<tr>
<td>Motivation Time 2</td>
<td>2</td>
<td>1</td>
<td>Gender</td>
<td>-.02</td>
<td>.55***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intelligence</td>
<td>.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Motivation Time 1</td>
<td>.73***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Group</td>
<td>-.22**</td>
<td>.60***</td>
<td>.05**</td>
</tr>
<tr>
<td>GCSE points</td>
<td>3 &amp; 4</td>
<td>1</td>
<td>Gender</td>
<td>.10</td>
<td>.68***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intelligence</td>
<td>.75***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Motivation Time 1</td>
<td>.23**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>Motivation Time 2</td>
<td>.20*</td>
<td>.73***</td>
<td>.04*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Group</td>
<td>-.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* *$p < .05$; **$p < .01$; ***$p < .001$. 
Discussion

The current study examined the ability of a brief multi-method CBT SMI (Flaxman et al., 2002) to impact GCSE performance and mental health in school children. Furthermore, it sought to establish the mechanisms, or mediators, by which these two benefits occurred. The above results indicate that we met these objectives, and we shall now discuss the findings.

Examination performance

Results indicate that our CBT SMI had a significant effect on the GCSE examination performance of a group of secondary school children. When compared to a control group, pupils who were randomly allocated to receive the SMI scored on average eight points higher, even when controlling for gender and intelligence. The magnitude of this effect means that, on average, each student in the SMI group received exam scores that were a letter grade higher, than did students in the control group. In an age of school league tables and increasing pressure for pupils to perform better, the potential outcome of such a psychological intervention, if adopted by UK schools, is a significant improvement in examination performance. We also believe that these findings are unique in that, as far as we are aware, no study has investigated the effects of an SMI on the performance of school children taking a standardized examination in the UK.

Alongside demonstrating that our SMI had a significant impact on examination performance, we were also able to identify the mechanism by which this effect occurred. As hypothesized, we found evidence that an increase in motivation (need for achievement) in the intervention group accounted for the group difference in examination performance that we found. This result is consistent with research that indicates that motivation plays a crucial role in the academic success of school children (e.g., Ames & Ames, 1984).

Mental health

The second aim of the current study was to examine whether our SMI improves the overall mental health of pupils. As predicted we found that pupils who were allocated to the intervention group exhibited a significant improvement in general mental health compared to controls. This finding is consistent with those from worksite SMIs that occur in adult organizations (e.g., Bond & Bunce, 2000). We believe that the current study is unique, though, in that it examines the effectiveness of an SMI in the ‘worksite’ of children: the school. The encouraging mental health and performance findings from this study suggest that applied psychologists may effectively use their strategies for the benefit of both adult-centred and child-centred organizations.

In addition to finding that our CBT SMI improved mental health, we were also able to identify the mechanism by which the intervention produced this benefit. As predicted, we found that positive changes in GHQ scores in the intervention group occurred because the dysfunctional beliefs and cognitions of children in this group became more functional and less distorted. This change is consistent with our research aims: to reduce pupils’ dysfunctional (worrisome) beliefs about academic success by providing them with strategies that they could use to make them less distorted and more functional. These mediation effects, concerning mental health, are consistent
with core CBT principles; that is, if one challenges dysfunctional cognitions, an improvement in mental health occurs (e.g., Beck, 1976; Bond & Dryden, 2002).

Even though we found evidence that our SMI improved examination performance and mental health, the fact that different mechanisms mediate such effects suggests that our intervention worked on a number of different, possibly independent, levels. If, as was found in the current study, improvements in mental health and examination performance occur via different mechanisms, it would be interesting to determine which aspects of the intervention reduces stress, and which affects academic performance.

Interestingly, whilst our SMI improved examination performance and general mental health, it did not affect test anxiety levels. This finding is in contrast to other studies that report reductions in test anxiety following SMIs (e.g., Vagg & Papsdorf, 1995). Indeed, we predicted reductions in test anxiety because others have reported success in their SMIs. One possible explanation could be due to the measure of test anxiety used here (Revised Test Anxiety Scale; Benson & Bandalos, 1992). It is very possible that had alternative measures of examination stress been used then a different pattern of results could have been found. Indeed, a general point is that the results found here are very much dependent on the measures used.

A second possibility could be that test anxiety is conceptualized as a trait that makes people vulnerable to anxiety in the context of evaluations (Sarason, 1980). As a result, this trait should be relatively stable and difficult to change over a relatively short period of time. Given that our two measurements (i.e., Times 1 and 2) were only 10 weeks apart, it is conceivable that this duration was too short to see any notable change in the trait of test anxiety. One could also argue that, since test anxiety is supposedly stable, little change would be expected anyway. This leads to the interesting possibility that our intervention was able to provide pupils with the necessary skills to reduce the disruptive influence of examination stress without changing their stable fears. If so, then one does not necessarily have to remove the underlying fear trigger, but merely provide pupils with the mechanisms to control it. Such a view would be consistent with a number of test anxiety theories, including Eysenck and Calvo’s (1992) processing efficiency theory, which argues that those high in test anxiety have the means to perform a task, but just less efficiently, due to cognitive interference through worry. They believe that test anxiety disrupts how well individuals perform a task, not actual ability to do the task. However, whilst it is possible that test anxiety is stable, the fact that the cognitive components of test anxiety were only found to be negatively related to examination performance when measured closer to examinations, suggests some state-related effects. Clearly, further research is required in order to understand the degree to which an SMI affects test anxiety, the nature of this relationship, and whether the duration required for such an intervention to have an effect is important.

Limitations

Not unexpectedly with applied research of this nature, there are a number of potential limitations with the current study that should be considered before any definite conclusions can be made. One concern relates to the attrition rate within the intervention group, which resulted in a loss of around 50% of the original sample. The main reason for attrition was due to poor attendance (<60%) at the intervention sessions. Given that there were differences in intelligence between those included in the final analysis and those not, it is possible that the intervention is
only effective for those higher in intelligence. It is also possible that, although there were no group differences in motivation at time 1, those with higher levels of intelligence maintained motivation and so continued to attended sessions, whereas those lower in intelligence did not sustain motivation levels over time. Of course such views are speculative, especially since no mid-intervention assessment of motivation was taken. However, if correct, then the potential impact of such reasons for attrition on the conclusions drawn is that the effectiveness of the intervention may be limited to a selective sample of students who are highly motivated. Future research needs to determine whether this is the case, and if so, consider more appropriate methods for those concerned.

A second potential limitation concerns our decision not to include a placebo condition, which limits the strength of the conclusions that can be drawn regarding the efficacy of the intervention. It is, for example, possible that the group effects found here could have been due to the mere fact that we interacted with the intervention group, rather than something specific to our intervention. It is also possible that the higher motivation in the SMI group at time 2 could be due to an unspecified effect of being aware that they belong to a ‘special’ group. The decision not to interact with the control group, through the use of a placebo intervention, was made because we felt that removing children from the classroom at such a critical period in their academic careers could not be justified. It may have been possible to run the intervention before or after school. However, given the issue of attrition, this method may have reduced the overall number of students completing the intervention. Therefore, until a suitable follow-up study is conducted, which compares the specific SMI used here with an alternative intervention, no definite conclusions can be made.

**Implications**

If the current results are reliable then they have a number of theoretical and practical implications. Our research indicates that SMIs have the potential to improve mental health and academic performance, through separate, and theoretically derived, mechanisms. On a more practical level, if SMIs can help children to perform better and be more mentally healthy, it is likely that students will get more out of education. Research is planned to investigate the effect of our SMI on other outcome variables, such as absenteeism. Although it is well established that poor mental health is associated with increased turnover and absenteeism in adult-centred organizations (e.g., Quick, Quick, Nelson, & Hurrell, 1997), it would be interesting to investigate such indicators of productivity amongst children, in an educational setting. We predict that, as with adult organizations, improved mental health would be associated with increased attendance and academic success.

**Conclusions**

In sum, the current study presents strong evidence that the CBT SMI outlined by Flaxman et al. (2002) has great efficacy and potential use in terms of improving educational performance, and the mental health of school children. This study also provides theoretically, and practically, important information as to why this intervention works. Examination performance seems to
benefit from higher levels of motivation, whereas mental health improvements occur through a reduction in dysfunctional cognitions and attitudes. These results clearly demonstrate that our intervention has potential for future stress management, and further research is warranted.

References


