Putwain, D, Nicholson, L and Pekrun, R

Test Anxiety, Emotion Regulation, and Achievement: Lagged and Contemporaneous Reciprocal Relations

Test Anxiety, Emotion Regulation, and Achievement: Lagged and Contemporaneous Reciprocal Relations

David W. Putwain[1], Laura J. Nicholson[1] and Reinhard Pekrun[2,3]

[1] School of Education, Liverpool John Moores University, UK.

[2] Department of Psychology, University of Essex, UK

[3] Institute for Positive Psychology and Education, Australian Catholic University, Australia.

Author Note

David W. Putwain https://orcid.org/0000-0001-5196-4270

Laura J. Nicholson https://orcid.org/0000-0001-9255-9650

Reinhard Pekrun https://orcid.org/0000-0003-4489-3827

We have no known conflict of interest to disclose. The materials, dataset, and analytic code, on which analyses within this manuscript were conducted can be accessed at https://doi.org/10.17605/OSF.IO/XVFSY.

Correspondence concerning this article should be addressed to: David W. Putwain, School of Education, Liverpool John Moores University, Maryland St, Liverpool, L1 9DE. Phone: +44(0)151 231 5270. Email: D.W.Putwain@ljmu.ac.uk

Abstract

Many studies have shown that test anxiety is negatively related to achievement, but studies testing reciprocal relations with achievement are largely lacking. Furthermore, non-adaptive forms of emotion regulation have been linked to elevated anxiety but there are few studies of emotion regulation and test anxiety specifically, and emotion regulation is rarely considered alongside relations between test anxiety and achievement. Given that strategies to regulate emotions about exams could have a substantive bearing on test anxiety, we aimed to gain a more comprehensive understanding of how test anxiety relates to achievement by also considering emotion regulation strategies. Data were collected from 533 students (mean age 16.4 years; male = 25.8%) in a sample of upper secondary students over five waves with six-month intervals. Achievement was measured at the first, third, and fifth waves, and test anxiety and emotion regulation, at the second and fourth waves. Data were analyzed in a series of structural equation models to model relations between test anxiety and achievement over time and test anxiety and emotion regulation contemporaneously. Test anxiety showed negative reciprocal relations with achievement and positive reciprocal relations with two emotion regulation strategies, catastrophizing and rumination. Of the remaining emotion regulation strategies, test anxiety followed putting into perspective in a negative unidirectional relation; positive reappraisal and refocus on planning followed test anxiety in negative unidirectional relations, and acceptance, self-blame, and other-blame, followed test anxiety in positive unidirectional relations. Catastrophizing, rumination, and putting into perspective may be useful foci for test anxiety intervention.

Keywords: test anxiety, emotion regulation, achievement; high-stakes exams

Educational Impact and Implications Statement

Lower achievement leads to higher subsequent test anxiety and higher test anxiety leads to lower subsequent achievement in a negative debilitating cycle. Intervention is therefore required to limit anxiety-related underachievement. Two forms of non-adaptive emotion regulation, catastrophizing and rumination, were positively related to test anxiety in a cyclic fashion, that is higher rumination and catastrophizing led to higher test anxiety and vice versa. In addition, an adaptive form of emotion regulation, putting into perspective, led to lower test anxiety. Strategies to reduce rumination and catastrophizing and enhance putting into perspective would be useful foci for intervention.

Test Anxiety, Emotion Regulation, and Achievement: Lagged and Contemporaneous Reciprocal Relations

High-stakes exams, with impactful consequences for subsequent life trajectory (Morris, 2011), and low-stakes tests used to measure student progress and provide feedback (Dixson & Worrell, 2016), are ubiquitous in many education systems. The outcomes of high-stakes tests influence access to higher-level education and training (Carnoy, 2005), entry to the job market (Thompson, 2017), and earning potential (Machin et al., 2020). Both low- and high-stakes tests and exams can considerably impact one's self-worth, where judgments become contingent on achieving high or aspired grades (Jones, 2007; Wenzel & Reinhard, 2021).

Given the profound consequences of high-stakes tests and exams, it is not surprising that many students find them anxiety-provoking. In the 2015 assessment of the OECD's Programme for International Student Assessment (PISA), 55.5% of students "agreed" or "strongly agreed" that they felt very anxious about tests even if well prepared (OECD, 2017). Studies in the United Kingdom, the location of the present study, have found that approximately 16.4% of students preparing for secondary school exit exams (Putwain & Daly, 2014), and 18.8% of students preparing for upper secondary school exit exams (Putwain, 2020), reported high levels of test anxiety.

High levels of test anxiety can contribute to poor mental and physical health and jeopardize future life trajectory by interfering with achievement. A robust evidence base has shown that test anxiety (e.g., Hembree, 1988; von der Embse et al., 2018) and domain-specific anxieties that incorporate anxieties about testing, such as mathematics anxiety (e.g., Barroso et al., 2021), are negatively related to achievement. Importantly, longitudinal studies allowing for a directional interpretation of findings, by controlling for prior achievement and other relevant covariates, have shown that higher test anxiety drives subsequent lower

achievement (e.g., Pekrun, 1992; Putwain et al. 2015). Theoretically, lower achievement provides feedback that would also lead to subsequent higher anxiety by lowering competence appraisals (Pekrun, 2006). Studies examining reciprocal relations between test anxiety and achievement, however, are largely lacking (see Pekrun, 1992, and Steinmayr et al., 2016, for exceptions).

Given the impact of test anxiety on achievement, it is important to know how students deal with feelings of anxiety about tests, especially in the face of failure feedback or an exam that was perceived not to go as well as hoped for. Students can become fixated on worry and exaggerate the consequences of the exam, becoming trapped in an ongoing cycle of anxiety. Theoretically, such an emotion regulation strategy would enhance test anxiety, thereby reducing achievement (Wells, 2009). Alternatively, students could try to keep things in perspective and devise plans to try to reduce the likelihood of failure in the future. Such strategies would reduce test anxiety, thereby enhancing performance (Harley, Pekrun, et al., 2019).

Accounting for emotion regulation provides a more comprehensive account of the processes that lead some students to greater, and other students to lower, test anxiety while studying for and taking high-stakes exams. Studies have shown how different emotion regulation strategies positively or negatively relate to test anxiety (e.g., Decuir-Gunby et al., 2009), but reciprocal relations with test anxiety have yet to be considered. Since emotion regulation strategies positively correlated with test anxiety could either be causes of high anxiety (e.g., rumination) or responses to anxiety (e.g., deep breathing), it is critical to untangle directional relations between emotion regulation and test anxiety.

In the present study we draw on control-value theory (CVT; Pekrun, 2006, 2024; Pekrun, Marsh, Elliot, et al., 2023), the integrated model of emotion regulation in achievement situations (ERAS; Harley, Pekrun, et al., 2019), and the Self-Regulatory

Executive Function Model of anxiety (S-REF; Wells, 2009) to examine reciprocal relations between test anxiety, emotion regulation, and achievement. We used longitudinal assessments of test anxiety, emotional regulation, and achievement in a sample of students preparing for and taking high-stakes upper secondary school exit exams in England. The results of these exams are typically used for university entrance. Other studies have examined reciprocal relations between composite academic anxiety (comprising class, homework, and testing) and achievement (Pekrun et al., 2017). Our literature review identified only one study (Steinmayr et al., 2016) to examine reciprocal relations between students' test anxiety and their achievement. This study, however, only used two measurement waves. Our study moves beyond two waves to allow for a more robust test of reciprocal relations between test anxiety and achievement and considers these relations alongside strategies to regulate test anxiety.

The Key Constructs: Test Anxiety and Emotion Regulation

Test Anxiety

Test anxiety is a situation-specific form of anxiety arising from the appraisal of a performance-evaluative situation as a psychological threat (Spielberger & Vagg, 1995). It may comprise affective (e.g., feeling panicky), physiological (e.g., pounding heart), cognitive (e.g., freezing' during an exam), motivational (e.g., avoiding the exam), and expressive (e.g., auto manipulation) components (Pekrun, 2006; Scherer & Moors, 2019). Moreover, test anxiety can be conceptualized as a state (i.e., a specific episode of anxiety before or during an exam) and as a trait. In the present research, we conceptualize test anxiety as a trait, defined as the relatively stable tendency to generally appraise exams as a threat and respond with elevated state anxiety (Lotz & Sparfeldt, 2017). We chose to focus on measuring two affective-physiological and two cognitive components as providing the most salient indicators of test anxiety (see Putwain, von der Embse, et al., 2020). The two affective-physiological components were "tension" (feelings of anxiety, such as panic) and

"physiological indicators" (e.g., heart pounding). The two cognitive components were "worry" (thoughts concerning failure) and "cognitive interference" (perception of reduced efficiency in memory and attention).

Emotion Regulation

Emotion regulation (ER) refers to how persons deliberately or automatically influence the type, intensity, timing, and expression of emotions in relation to personal goals and situational demands (Gross, 2015). Goals direct emotion regulation to achieve a specific purpose or objective and are important for two reasons: They provide self-regulatory direction and provide a mechanism to judge progress. Any specific single goals are located within a dynamic self-system comprising multiple, hierarchically organized (and possibly contradictory) goals (Shutz & Davis, 2000). For instance, one might have a hedonic goal to avoid unpleasant states such as anxiety. At the same time, one may need high grades to gain a place on a competitive university course (an instrumental goal), which raises extrinsic value and hence potential anxiety.

The Process Model of ER (PMER; Gross, 1998, 2015) proposes that different ER strategies are used at five stages in the emotion-generative process (also see Harley, Pekrun, et al., 2019). These are situation selection (e.g., choosing a conducive study setting), situation modification (e.g., using effective test-taking strategies during an exam), attentional deployment (e.g., fixating on catastrophic consequences of failure), cognitive change (e.g., blaming one's teachers if an exam did not go well), and response modulation (e.g., using diaphragmatic breathing). CVT proposes a similar classification of strategies (Pekrun, 2006, 2024; Pekrun & Stephens, 2009) but adds one group of strategies that is especially important to regulate achievement emotions: increasing one's competence, which can serve to attain success and all the positive emotions coming with success (competence-oriented regulation). The effectiveness of a selected strategy is judged in progressing towards or reaching a goal.

Thus, ER strategies could be more effective in one situation and less in another, or useful in the short-term but not in the long-term (e.g., Bonanno et al., 2004). Nonetheless, some habitually used strategies are typically adaptive (e.g., reappraisal, putting into perspective, competence development) and others are typically maladaptive because they lead to greater psychological distress and behavioral problems (e.g., self-blame, rumination; Garnefski et al., 2007; Garnefski & Kraaij, 2018).

We used the Cognitive Emotion Regulation Questionnaire (CERQ: Garnefski et al., 2002) to measure ER. The CERQ is based on Lazarus and Folkman's (1984) transactional model of stress and coping, which also informed the development of CVT (Pekrun, 2006, 2025) and the ERAS (Harley, Pekrun, et al., 2019). Nine strategies that can be used after a negative event are assessed in the CERQ (in our case, strategies following an exam that did not go well). Specifically, the CERQ focuses on the appraisal processes that precede taking action (Garnefski et al., 2001), including attentional deployment and a broad range of strategies for cognitive change.

To explain the strategies measured by the CERQ, attentional deployment can focus on or distract from the unpleasant or distressing emotions associated with the negative situation (Schutz & Davis, 2000; Schutz et al., 2014). As such, attentional deployment comprises strategies that can up- or down-regulate test anxiety (Davis et al., 2008; Decuir-Gunby et al., 2009). We measured attentional deployment using CERQ subscales for rumination and positive refocusing (thinking about a pleasant alternative). Cognitive change involves the use of strategies providing critical or reassuring self-talk (Davis et al., 2008; Schutz et al., 2008) and that aim to gain, keep, or regain task focus (Schutz & Davis, 2000; Schutz et al., 2014). We measured the former aspect of cognitive change using CERQ subscales for reappraisal and putting into perspective, and the latter using the CERQ subscale for refocus-on-planning.

Refocus-on-planning also aligns with competence development in the CVT. CERQ strategies are classified in Table 1 according to their stage in the PMER/ERAS and CVT.

Test Anxiety and Emotion Regulation

The strategies used to regulate emotions about exams (before, during, and after) are likely to have a substantial impact on the intensity of test anxiety. The S-REF model of anxiety (Wells, 2009; Zeidner & Matthews, 2005) proposes that situational threats, like a forthcoming exam, trigger an episode of executive processing which includes ongoing appraisals of the situation, goal congruence, and progress toward the goal, as well as plans for coping with the threat. Executive processes draw on self-beliefs including competence appraisals and generic plans for dealing with the situation. With positive competence beliefs and adaptive plans, the episode of executive processing may be brief. Ongoing monitoring of the situation would occur below the threshold of conscious awareness. A subsequent episode of controlled executive processing would only be triggered if there was a discrepancy between the situation and one's goals (e.g., falling behind in one's study timetable).

In the S-REF model, another facet of self-knowledge, namely metacognitive beliefs about anxiety, can result in the use of ER strategies that prolong the period of executive processing to increase and maintain test anxiety. Metacognitive beliefs that worry and anxiety must be controlled (e.g., *"*If I don't control my worry, it will control me"), are dangerous (e.g., "worrying too much can harm me"), or positive (e.g., "analyzing my problems will help me find answers"), result in increased monitoring of one's internal state, failing to disrupt worry, persistent overthinking, and in some cases it intensifies negative outcomes. Rumination and catastrophizing especially would result in increased executive threat processing. Rumination involves recalling past failures and focusing on negative self-perceptions, resulting in an enhanced expectation of failure (Clark & Wells, 1995). Catastrophizing involves contemplating negative scenarios involving failure and its

consequences in an attempt to plan coping options. Reflecting on negative scenarios also intensifies threat processing. The resulting intensification of anxiety, however, fails to provide an "internal signal" that one can stop worrying, hence the period of executive processing is prolonged (Wells & Carter, 2001).

Meta-analyses have shown that anxiety symptoms are negatively related to habitual use of problem solving, reappraisal, and acceptance (*r*s = -.07 to -.42), and positively related to avoidance, rumination, and emotional suppression (*r*s = .21 to .46), in child and adolescent populations (Kraft et al., 2023; Schäfer et al., 2016). Relatively few studies have examined how ER is linked specifically to test anxiety. Nonetheless, cross-sectional studies of ER in university students show that ER strategies that maintain or retain task focus (e.g., problem solving, competence development) are negatively related to test anxiety, while those that focus on emotions (e.g., self-blame, emotional suppression), show positive relations (Burić et al., 2016; Decuir-Gunby et al., 2009; Harley, Jarrell, et al., 2019; see also Stockinger et al., 2024).

Theoretically, test anxiety would be linked to ER in a reciprocal fashion. In the ERAS model (Harley, Pekrun, et al., 2019), achievement emotions including test anxiety prompt the use of different regulation strategies, namely situation selection, situation modification and competence development, attentional deployment, cognitive change, and response modulation, at different stages of the emotion-generative process. These strategies, in turn, can impact features of the achievement situation, attention, and control-value appraisals, to influence the type and intensity of achievement emotions generated, as well as emotional responses directly. In the S-REF model of anxiety (Wells, 2009; Zeidner & Matthews, 2005), we have already described how some ER strategies would terminate, while others would prolong, executive threat processing, to diminish or enhance anxiety. In turn, awareness of state anxiety prompts the use of further ER strategies.

While reciprocal relations are theoretically posited and have substantive implications for intervention, studies have yet to provide empirical tests. ER strategies that impact test anxiety would be important tools in interventions to help manage or reduce test anxiety. If ER strategies are outcomes of test anxiety, however, they would be less useful to test anxiety intervention but may nevertheless assist in promoting mental health.

Test Anxiety and Achievement

Meta-analyses have shown that test anxiety negatively correlates with achievement in the region of r = -.20 to -.40 (Hembree, 1988; Seipp, 1991; von der Embse et al., 2018). Indeed, test anxiety shows one of the strongest relations with achievement of all psychological attributes and processes (Richardson et al., 2012; Schneider & Peckel, 2017). Unpacking the directionality in this correlation (a classic chicken and egg conundrum, e.g., Marsh et al. 2024) is critical to understanding its substantive educational implications.

If the test anxiety-achievement correlation is driven by anxiety-induced cognitive interference (interference model; Zeidner, 1998), intervention should focus on reducing anxiety to raise achievement or using exam procedures and alternative assessments that are less prone to generate anxiety. If, on the other hand, higher test anxiety is merely an epiphenomenon of lower achievement and competence perceptions (deficit model), then effective interventions should use better instruction to build competence and mastery.

Pekrun (1991, 1992) has proposed a reciprocal effects model (REM) of anxiety and achievement that integrates both perspectives. This model became subsequently part of CVT (Pekrun, 2006). In this model, anxiety is proposed to often (although not always) lower achievement through cognitive-motivational processes including impaired working memory functions (e.g., less flexible switching), reduced working memory capacity, decreased intrinsic motivation, and increased avoidance motivation. These anxiety-induced effects reduce achievement. Poor achievement (i.e., exam scores and grades), in turn, provides

negative feedback on competence which lowers control appraisals, including the belief that one can achieve a successful exam outcome. Alongside the value attached to exam outcomes, control appraisals are a proximal antecedent of test anxiety. Over successive iterations, test anxiety and achievement would relate in reciprocal vicious cycles.

Few studies have tested directional relations between test anxiety and achievement in academic settings. Exceptions have shown test anxiety to negatively predict subsequent test scores in Year 6 primary school children after controlling for prior cognitive ability (Putwain et al., 2013), and grades on Year 11 secondary school exit exams after controlling for prior achievement, coping, and academic buoyancy (Putwain et al., 2015). These and other studies (e.g., Pekrun, 1991, 1992) establish how test anxiety can negatively impact achievement and is not merely the effect of low achievement.

Studies examining the reverse direction, how achievement impacts subsequent test anxiety, or reciprocal relations, are largely lacking. An exception is the study by Steinmayr et al. (2016). The study used a two-wave design with test anxiety and achievement measured at a twelve-month interval in secondary school, controlling for autoregressive paths and effects of mood and life satisfaction. The cognitive component of test anxiety (worry) predicted grade point average but not vice versa. The study was limited by the small sample size and collinearity of the affective (emotionality) and cognitive (worry) test anxiety variables that were included in the model. However, despite limited direct evidence that achievement could influence subsequent test anxiety, the principle remains sound: Feedback about achievement impacts competence beliefs (e.g., Arens et al., 2017; Forsblom et al., 2022), and these beliefs impact test anxiety (e.g., Preiss et al., 2006).

Furthermore, the REM of anxiety and achievement is supported by longitudinal evidence on related forms of anxiety, such as students' mathematics anxiety. For example, using data from the Project for the Analysis of Learning and Achievement in Mathematics

(PALMA), Pekrun et al. (2017) examined reciprocal relations between mathematics anxiety (defined as aggregated class-, learning-, and test-related anxiety), end-of-year math grades, and math test scores over five annual waves in a sample of secondary school students. Anxiety was a negative predictor of grades and test scores which, in turn, were negative predictors of anxiety, controlling for autoregressive relations, intelligence, and socio-demographic variables (for a replication of these findings using random-intercept cross-lagged modeling, see Pekrun, Marsh, Suessenbach, et al., 2023).

In a similar study of elementary school students, Lichtenfeld et al. (2023) examined reciprocal relations between learning- and test-related math anxiety and math grades. Relations with learning-related and test-related anxiety were established over three and two annual waves, respectively. Grades significantly predicted test anxiety but not vice versa. Reciprocal relations were shown between grades and learning-related anxiety. On balance, the weight of available evidence is supportive of the REM of test anxiety and achievement, but further studies are needed to test the model.

Emotion Regulation and Achievement

Studies have shown effective emotion regulation is associated with better school adjustment and developmental outcomes in pre-schoolers and children (e.g., Djambazova-Popordanoska, 2016; Garner, 2010). For example, children with better teacher-reported emotion regulation in preschool showed better academic competence six months later having transitioned to school, resulting from greater attention to academic tasks (Trentacosta & Izard, 2007). Theoretically, ER is expected to impact achievement indirectly as a function of directing attention towards or away from task engagement and disruptive emotions such as test anxiety (Tyson et al., 2009). Studies have supported this proposition, showing indirect relations between adaptive ER and achievement in children through lower internalizing symptoms, such as anxiety (Wong et al., 2023), and greater classroom engagement (Kwon et

al., 2018). Studies of ER and achievement in populations of adolescent students, however, are largely lacking. In addition, there is a lack of studies assessing specific ER strategies like those included in the CERQ (e.g., Cécillon et al., 2024; Xu et al., 2019) and, to the best of our knowledge, no studies that consider ER specifically in relation to test anxiety, using longitudinal designs to control for prior achievement.

Optimal Intervals Between ER and Test Anxiety: Lagged and Contemporaneous Models

Establishing the optimal time interval between predictor and outcome for longitudinal designs is not straightforward and depends on the nature of the variables, theoretical expectations regarding the timescale of the processes, and the population under study (Dormann & Griffin, 2015). Lagged effects analyses of relatively stable constructs, linking trait-like achievement emotions or academic self-concept with achievement, typically use intervals lasting months or years (e.g., Marsh, 1990; Omer & Chen, 2023; Pekrun et al., 2017). The ER processes specified in the ERAS model (Harley, Pekrun, et al., 2019) and S-REF (Wells, 2009; Zeidner & Matthews, 2005) can unfold over shorter time periods.

The dataset used in the present study used a twelve-month interval between assessments of test anxiety and ER. It is not clear, however, whether the typical use of ER strategies and trait-like test anxiety would show reciprocal relations over a twelve-month interval (also see Hollenstein, 2015). Failure to find reciprocal effects from twelve-month lags would not necessarily indicate no relations as they may have faded over time and would have been captured with shorter time lags.

When optimal intervals are possibly short but uncertain, a plausible alternative is to model contemporaneous relations (Marsh et al., 2024; Muthén & Asparouhov, 2024). Contemporaneous does not imply simultaneous; rather, processes linking predictor and outcome variables develop more quickly than the consecutive intervals used in classic lagged

analyses (Rohrer & Murayama, 2023). For instance, test anxiety may trigger the use of ER strategies that last several weeks or months but may not extend to twelve months. In this case, estimates of contemporaneous relations capture hypothetical unobserved temporal relations that occur within twelve months.

As ER processes occur quickly and may not be captured effectively in longer time intervals (Hollenstein & Lanteigne, 2018), we examined relations with test anxiety using a contemporaneous effects modeling approach (Marsh et al., 2024). That is, reciprocal relations between test anxiety and ER were modeled using data collected at the same wave. The estimates of directional relations derived from these data represent the effects of test anxiety on subsequent ER use, and vice versa, that are theorized to have temporal lags that occur within a shorter period than the twelve-month measurement interval. For comparative purposes, models with lagged test anxiety and ER relations are reported in the Supplementary Materials.

Aims of the Present Study

The present study aimed to test reciprocal lagged relations between test anxiety and achievement and reciprocal contemporaneous relations between test anxiety and ER in a sample of upper-secondary students. Data were collected over five waves (T_1 to T_5). Achievement was collected at the first, third, and fifth, waves, and test anxiety and ER at the second and fourth waves. Thus, we were able to test lagged test anxiety and achievement, and contemporaneous test anxiety and ER, relations, twice (see Figure 1). Although not a specific focus of the current study, for completeness, we included relations between ER and achievement. We leave the nature of relations as an open research question, but would largely expect these to be positive for adaptive, and negative for nonadaptive strategies as a function of lower test anxiety. Girls, and those from economically-deprived backgrounds, typically report higher test anxiety (King et al., 2024; Robson et al., 2023). Furthermore, in England,

girls show higher school achievement, but students from economically deprived backgrounds show lower achievement (Strand, 2014). We therefore included gender and free school meal eligibility (a proxy for low family income) as demographic covariates to control their possible confounding influence.

We tested the following hypotheses:

Hypothesis 1: Test anxiety, emotion regulation, and achievement show significant stability (positive autoregressive paths), including effects of T_2 test anxiety/emotion regulation on T_4 test anxiety/emotion regulation; and T_1 and T_3 achievement on T_3 and T_5 achievement.

Hypothesis 2. Test anxiety and achievement are linked by negative reciprocal relations (negative effects of test anxiety at T_2 and T_4 on achievement at T_3 and T_5, respectively; negative effects of achievement at T_1 and T_3 on test anxiety at T_2 and T_4, respectively).

Hypothesis 3. Test anxiety and emotion regulation are linked by reciprocal contemporaneous relations at T_2 and T_4; positive for non-adaptive strategies, and negative for adaptive strategies.

Method

Procedure, Participants, and Missing Data

Data were collected over five waves (T_1 to T_5) from students in upper secondary education (Years 12 and 13). T_1 data were grades in two key subjects (English and mathematics) from secondary school exit exams taken towards the end of Year 11 (May and June). T_1 data were reported by students retrospectively at T_2. T_2 data were self-reported test anxiety, emotion regulation, and socio-demographic variables, collected in January and February of Year 12 (mean age 16.4 years, *SD* = .52). T_3 data were aggregated grades from internal college exams taken at the end of Year 12 (June and July). These grades were accessed from official records. T_4 data were self-reported test anxiety, emotion regulation,

and socio-demographic variables, collected in January and February of Year 13 (mean age 17.4 years, *SD* = .51). T_5 data were aggregated grades from exit exams taken at the end of Year 13 (June) and accessed from official college records.

There were 533 and 535 participants at T_1 and T_2 respectively (i.e., two students at T_2 had not reported their achievement; 99.7% consented to T_1 achievement data being used); 439 participants at T_3 (all returners from T_1 and T_2; 82.1% of the T_2 sample consented to T_3 achievement data being used); 450 participants at T_4 (318 returners, 132 new); and 361 participants at T_5 (all returners from T_4 and 258 returners from T_1, T_2, and T_3; 80.2% of the T_4 sample consented to T_5 achievement data being used). In total 237 persons participated in all five waves (approximately 44% of the T_1/T_2 sample). Additional participant information and demographic information are provided in the Supplementary Materials.

The project received approval from an institutional research ethics committee (17/EHC/001). Letters were sent to the principals of schools who work in partnership with the first and second author's employing institution. Three schools agreed to participate resulting in an opportunistic sample. Written consent was provided by principals and participants at T_2 and T_4. Self-reported data at T_2 and T_4 were linked using an anonymous code unique to each participant. In addition, participants could choose to only participate in the self-report phase of the study and withhold their achievement data. If permission to access T_3 and T_5 achievement data from college was provided, participants provided their unique college identification number. Self-reported data were collected electronically using an online survey platform which prompted participants to respond if they had missed an item to minimize within-person missing data. The T_2 and T_4 questionnaire took approximately 10-15 minutes to complete.

In total, 25.6% of values were missing which is not uncommon in longitudinal survey studies (e.g., Jeličić et al., 2009; Nicholson et al., 2017). Data and model-based methods to

handle missing data have been shown to generate unbiased parameter estimates under assumptions of Missing Completely at Random (MCAR) or Missing at Random (MAR), even at rates of 40% attrition (Enders, 2010; Kristman et al., 2005). Little's omnibus test of Missing Completely at Random (MCAR; Little, 1988) was not statistically significant, $\chi^2(104) = 119.82$, $p = .14$, indicating that MCAR could be assumed. Accordingly, we used a model-based approach, namely full-information maximum likelihood (FIML), to handle missing data in subsequent analyses.

Transparency and Openness

The dataset, analytic code, and materials, used for the present study are deposited with the Open Science Framework (Author and Author, 2025). The study, and hypotheses, were not preregistered.

Measures

Test Anxiety

Test anxiety was measured using the 16-item MTAS (Putwain et al., 2021). The MTAS comprises four subscales, each with four items. Two subscales measure cognitive aspects of test anxiety, including worry (e.g., "Before a test/ exam, I am worried I will fail") and cognitive interference (e.g., "During tests/ exams, I find it hard to concentrate"). The remaining subscales measure the affective-physiological aspects of test anxiety, including tension (e.g., "Even when I have prepared for a test/ exam I feel nervous about it") and physiological indicators (e.g., "Before I take a test/ exam my hand trembles"). Participants responded to items on a five-point scale (1 = "strongly disagree" to 5 = "strongly agree"). Studies have shown strong factorial validity for both lower-order and hierarchical factor models, test-retest reliability, and internal consistency, for the MTAS along with measurement invariance for gender, age, and family income background (Fenouillet et al., 2023; Putwain, von der Embse, et al., 2021; von der Embse et al., 2021). In the present study,

our research questions focused on the relations between global test anxiety, achievement, and ER. As such, we used higher-order models including test anxiety as a secondary factor, rather than models separating between test anxiety components. Internal consistency was good (Hierarchical McDonald's ωs = .75; see Table 2).

Emotion Regulation

Emotion regulation was measured using the 36-item CERQ (Garnefski et al., 2002). The CERQ comprised nine subscales, each with four items, that capture cognitive emotion regulation strategies used *after* a stressful or adverse event. In the present study, we asked participants to specifically think about strategies used after an exam that did not go as well as was hoped for. At T_2, participants were instructed to think about their most recent high-stakes exams, the General Certificate of Secondary Education (GCSE) exams taken at the end of Year 11, "Please think about a recent exam (e.g., one of your GCSEs) that didn't go as well as you wanted it to. When this happened, what did you think?". At T_4, we changed the wording to "e.g., one of your end-of-year 12 or mock exams"[1].

Four of the nine subscales assessed non-adaptive strategies. These were self-blame (e.g., "I think about the mistakes I have made in this matter"), other-blame (e.g., "I feel that others are responsible for what has happened"), rumination (e.g., "I often think about how I feel about what I have experienced"), and catastrophizing (e.g., "I often think that what I have experienced is the worst that can happen to a person"). The remaining five subscales assessed adaptive strategies. These were positive refocusing (e.g., "I think of something nice instead of what has happened"), refocus on planning (e.g., "I think about how I can best cope with the situation"), positive reappraisal (e.g., "I look for the positive sides to the matter"), putting into perspective (e.g., "I think that it hasn't been too bad compared to other things"),

[1] Mock (i.e., practice) exams taken under standardized conditions are commonly used in Year 12 to predict actual grades. Offers for university entrance in the UK are typically made on the basis of predicted grades estimated from mock exams.

and acceptance (e.g., “I think that I must learn to live with it”). Participants responded to the items on a 5-point scale (1 = “almost never” to 5 = “almost always”). Studies have shown that the measure has good construct validity, test-retest reliability, and internal consistency (e.g., Garnefski & Kraaij, 2007; Ireland et al., 2017). In the present study, internal consistency was, in the main, good (McDonald’s ωs = .63 to .86; see Table 2).

Achievement

T_1 achievement was measured as the mean grade in two compulsory GCSE subjects, namely English and mathematics. GCSE exams are graded on a 9-point numerical scale whereby 9 is the highest grade and 1 is the lowest. A grade 4 is considered a minimum pass grade. We chose English and mathematics for three reasons. First, as compulsory subjects, they were taken by all students. Second, minimum pass grades in English and mathematics are required for entry to academic upper secondary irrespective of which subjects are studied. Third, as numerous combinations of arts, mathematics and science, social science, and humanities, subjects may be chosen in upper secondary education, English and mathematics have broad applicability to different subject combinations and hence are more appropriate as a measure of prior achievement than other subjects.

T_3 achievement was measured as the mean grade obtained in the end-of-Year 12 exams. These are exams taken under standardized conditions to provide a practice (‘mock’) exam used to estimate a predicted grade on which university entry will be based. Mock exams may be perceived as high-stakes by some students as acceptance onto competitive university courses is contingent on achieving predicted grades. T_5 achievement was measured based on students’ grades in exit exams taken at the end of Year 13. T_3 and T_5 achievement

were graded on a 7-point letter scale that was converted to a numerical scale[2]. A mean grade was used due to the numerous possible combinations of subject options.

Socio-demographic Covariates

Gender (0 = male, 1 = female, 2 = non-binary, and 3 = prefer not to say) and FSM (0 = not eligible, 1 = eligible) were self-reported at T_2 and T_4. For analytic purposes, gender was treated as a binary variable, as insufficient numbers of participants chose non-binary or prefer not to say responses to warrant the inclusion of four dummy-coded variables. For completeness, we included a binary ethnicity variable (0 = British/European white, 1 = South Asian/British South Asian, Black African, Caribbean/British Black African or Caribbean). There were insufficient numbers of participants from Chinese/British Chinese or "other" backgrounds to justify their inclusion as additional dummy variables.

Analytic Approach

Our analytic strategy consisted of three broad steps. First, we estimated measurement models for test anxiety and the emotion regulation variables. Based on these models, we also tested the measurement invariance of the latent variables over time. Second, we estimated latent bivariate correlations from a comprehensive measurement model containing all study variables, including test anxiety, emotion regulation strategies, and achievement, along with gender and FSM as socio-demographic covariates. Third, we estimated a series of longitudinal structural equation models (SEMs) to test the hypothesized relations between test anxiety, emotional regulation, and achievement (see Figure 1). All analyses were conducted in Mplus 8.11 (Muthen & Muthen, 2018) using the maximum likelihood estimator.

Step 1: Separate measurement models and invariance over time. Test anxiety was estimated hierarchically as a single second-order factor (test anxiety) with the four MTAS

[2] Upper secondary school exit exams are General Certificate of Secondary Education Advanced Level (A Level) or Business and Technology Education Council (BTEC) qualification. A-Levels are graded on a seven-point scale and BTECs are graded on a four-point scale. BTEC grades were converted to A-Level equivalents using the standardized system operated by the Universities and Colleges Admissions Service in England.

subscales, comprising four items each, as lower-order factors. To account for items showing low-level cross-loading to non-target factors (e.g., worry items loading on tension) we first estimated a lower-order exploratory structural equation model (ESEM) based on the four MTAS subscales. Higher-order models cannot be estimated in ESEM, so we then estimated a hierarchical test anxiety factor based on the ESEM solution. This is referred to as the ESEM within confirmatory factor analysis (EwC) approach and will show identical degrees of freedom and model fit to the lower-order ESEM (Marsh et al., 2020; Morin et al., 2013). ER strategies were modeled as lower-order models with each ER strategy comprising four items. A minimum prerequisite for longitudinal modeling of relations between variables is invariance of factor loadings and intercepts on each measurement occasion (metric invariance; Widaman et al., 2010). Accordingly, we performed a series of invariance tests for hierarchical test anxiety and the nine CERQ ER strategies across T_2 and T_4 (see Table S1 in Supplementary Materials).

Step 2: Integrative measurement model and bivariate correlations. Like test anxiety, ER strategies may also show cross-loading to cognate non-target factors (e.g., positive reappraisal items loading on putting into perspective). Therefore, the measurement model including all variables was constructed as a set-exploratory structural equation model using target rotation (set-ESEM; Marsh et al., 2020; Morin et al., 2013). When cross-loadings are not accounted for (e.g., using traditional confirmatory factor analysis), model fit can decline and correlations between factors can be inflated (Marsh et al., 2014). There were four sets: test anxiety (using the hierarchical EwC model) and the nine ER strategies were modeled as distinct sets at T_2 and T_4. Achievement scores were treated as single-item latent variables and gender and FSM were added as manifest variables.

Step 3: Longitudinal structural equation models (SEMs). To test hypothesized relations between test anxiety, emotional regulation, and achievement, we estimated nine

separate SEMs, one for each emotion regulation strategy. We used this strategy to reduce complexity given that the study included nine different emotion regulation strategies, and to avoid problems caused by multicollinearity between the strategies. Based on the measurement model, test anxiety was modeled as a secondary factor, and the emotion regulation strategies were modeled as primary factors. Achievement scores were modeled as single-item latent variables (error variance was set to zero; Yu & Martin, 2014) and gender was included as a manifest variable. We chose not to include FSM or ethnicity as these variables did not show substantive correlations with other variables (see Table 3). Moreover, ethnicity, which typically intersects with gender and economic disadvantage in the UK (Strand, 2014), is difficult to interpret given the broad range of ethnicities combined in this variable. Generally, inclusion of ethnicity as a covariate may not be warranted if not theoretically justified (Singh et al., 2024).

In preliminary SEM analyses, we modeled reciprocal relations between test anxiety and ER as contemporaneous or lagged effects. Relations with achievement were modeled as lagged effects in both types of models. Using the terminology proposed by Marsh et al. (2024), the first type of model is a contemporaneous and cross-lagged panel model (CCLPM), and the second type is a pure cross-lagged panel model (CLPM). Confirming expectations, there were few lagged relations in the CLPMs (see Table S4 in Supplementary Materials), suggesting that effects linking test anxiety and ER may not be preserved over a twelve-month lag. In contrast, the CCLPMs showed numerous relations between test anxiety and ER, suggesting that estimates of contemporaneous effects are an effective method of capturing ER dynamics, and confirming that contemporaneous effects models are a plausible alternative for modeling processes that occur between the measurement waves in longitudinal panel designs (Marsh et al., 2024; Muthén & Asparouhov, 2024).

As such, using the CCLPM approach, the final models estimated relations linking test anxiety and ER as contemporaneous effects, and relations of both constructs with achievement as lagged effects. Figure 1 shows the paths that were estimated in each model. These include: (a) autoregressive effects for all variables (T_1 achievement on T_3 and T_5 achievement; T_3 achievement on T_5 achievement; T_2 test anxiety on T_4 test anxiety; T_2 emotion regulation on T_4 emotion regulation); (b) reciprocal effects linking test anxiety and achievement (effects of test anxiety at T_2 and T_4 on achievement at T_3 and T_5 respectively, and of achievement at T_1 and T_3 on test anxiety at T_2 and T_4 respectively); (c) reciprocal contemporaneous effects linking T_2 test anxiety and emotion regulation, as well as T_4 test anxiety and emotion regulation; and (d) reciprocal effects linking emotion regulation and achievement (effects of regulation at T_2 and T_4 on achievement at T_3 and T_5 respectively, and of regulation at T_2 and T_4 on achievement at T_3 and T_5 respectively). Residual variances for the corresponding test anxiety and emotion regulation items at T_2 and T_4 were allowed to correlate.

We also estimated models with contemporaneous and lagged effects either freely estimated or constrained to be equivalent over time (i.e., developmental equilibrium; see Marsh et al., 2019). Model fit was assessed using the root mean error of approximation (RMSEA), the standardized root mean residual (SRMR), the confirmatory fit index (CFI), and the Tucker-Lewis index (TLI). Based on simulation studies, Hu and Bentler (1999) recommended RMSEA values ≈ .06, SRMR values ≈.08, and CFI and TLI values ≈ .95 as indicating a good fit of a model with the data. When testing complex models with naturalistic data, however, these values need to be interpreted with a degree of flexibility (Heene et al., 2011; Lance et al., 2006).

Results

Descriptive Statistics, Measurement Models, and Latent Bivariate Correlations

Descriptive data are shown in Table 3. Skewness and kurtosis statistics were all within ±1, and McDonald's omegas were >.7 for all variables except for positive reappraisal (ωs = .68 and .69. at T_2 and T_4 respectively) and T_4 acceptance (ω = .63). The measurement models showed acceptable (acceptance) to excellent (rumination) model fit and were, in the main, good.

Factor loadings from the separate measurement models showed one T_2 and T_4 acceptance item with lower than ideal factor loadings (λ <. 3). Treatment of low-loading items is a matter of conjecture. On the one hand, low-loading items could be dropped to purify construct measurement (e.g., Stevens, 1992). On the other hand, removing items from a theoretically specified construct risks altering its nature and moving to a more exploratory framework (Knekta et al., 2019). If a model is not mis-specified (i.e., no mismatch between the factor structure of the true and estimated models; structural error) and low factor loadings are the result of either sampling or model error, the retention of low loading items may be justified (e.g., Ximénez, 2009). As previous studies have shown CERQ subscales to demonstrate good factorial validity, with acceptable factor loadings (e.g., Garnefski & Kraaij, 2007), we assumed no structural error and chose to retain the low-loading items.

With one exception (refocus on planning), all measures showed strict measurement invariance, namely equivalence of item-factor loadings, item intercepts, and item residual invariance, at T_2 and T_4 with no substantial loss of fit (see Table S2). Refocus on planning showed partial metric invariance where equivalence constraints were dropped for one item. Partial metric invariance is sufficient for longitudinal modeling (Widaman et al., 2010). The comprehensive set-ESEM/EwC model showed a reasonably good fit to the data, $\chi^2(5002) = 7156.63$, RMSEA = .025, SRMR = .049, CFI = .925, and TLI = .910, and so we proceeded to inspect bivariate correlations (see Table 4). In the main, positive refocusing, refocus on planning, positive reappraisal, and putting into perspective, were negatively correlated with

test anxiety; acceptance, rumination, catastrophizing, and self-blame, were positively correlated with test anxiety. Achievement correlated negatively with acceptance, other-blame, and test anxiety; achievement correlated positively with refocus on planning and rumination. Female participants reported higher rumination, catastrophizing, and test anxiety, and lower positive reappraisal, positive refocusing (T_2 only), and putting into perspective (T_4 only). Moreover, female participants showed higher achievement. Economic deprivation was largely unrelated to study variables; those eligible for FSM showed lower T_4 putting into perspective and T_4 other-blame. Participants from a South Asian/British South Asian, Black African, Caribbean/British Black African or Caribbean heritage showed higher T_2 refocus on planning and T_4 positive reappraisal, lower T_2/T_4 putting into perspective, and lower T_1/T_3 achievement, compared to those from a British/European white background. Given the heterogeneous categories used in this dummy ethnicity variable, we caution against overinterpreting these correlations.

Longitudinal Structural Equation Modeling

As we described above, separate CCLPMs were estimated, one for each emotion regulation strategy. Table 5 shows the standardized path coefficients for substantive variables with paths freely estimated and with developmental equilibrium constraints. There was no substantial loss of fit for the models with developmental constraints; all differences in fit indices were $\Delta < .003$. Since the constrained models are more parsimonious and provide more precise parameter estimates (Marsh et al., 2019), we proceeded to interpret coefficients from these (note that as constraints are applied to unstandardized coefficients, standardized coefficients can still differ over time intervals).

Our results showed autoregressive paths for achievement (T_1 to T_3 and T_5; T_3 to T_5), test anxiety, and emotion regulation (both T_2 to T_4). Test anxiety was a *negative* predictor of achievement (T_2 and T_4 test anxiety to T_3 and T_5 achievement, respectively). Achievement

was, in turn, a *negative* predictor of test anxiety (T_1 and T_3 achievement to T_2 and T_4 test anxiety, respectively). Test anxiety showed *positive* contemporaneous effects on acceptance and self-blame, at T_2 and T_4, but not vice versa. Test anxiety showed *negative* contemporaneous effects on refocus on planning and positive reappraisal, at T_2 and T_4, but not vice versa. Putting into perspective showed *negative* contemporaneous effects on test anxiety, T_2 and T_4, but not vice versa. Rumination and catastrophizing showed *positive* contemporaneous reciprocal relations with test anxiety at T_2 and T_4. Finally, positive refocusing and test anxiety were unrelated.

Rumination was a *positive* predictor of achievement (T_2 and T_4 rumination to T_3 and T_5 achievement, respectively). Achievement was, in turn, a *positive* predictor of rumination (T_1 and T_3 achievement to T_2 and T_4 rumination, respectively). T_2 acceptance and T_4 positive refocusing were *negative* predictors of T_5 achievement. Achievement was a *negative* predictor of other-blame (T_1 and T_3 achievement to T_2 and T_4 other-blame, respectively). Finally, T_1 achievement was a *positive* predictor of T_4 refocus on planning.

Coefficients for gender are shown in the Supplementary Materials (Table S2). Female students consistently showed higher T_1 and T_3 achievement. Coefficients for the effects of gender on T_5 achievement were not significant in all models but were in the same direction. Female students consistently reported higher T_2 test anxiety. T_4 coefficients for test anxiety and gender were not significant in all models but were also in the same direction. Female students showed lower rumination and self-blame at T_2 and T_4 and lower putting into perspective at T_4 only.

Discussion

This study aimed to test lagged reciprocal relations between test anxiety and achievement, and contemporaneous reciprocal relations between test anxiety and ER. The current study is the first to comprehensively examine these possible reciprocal relations. The

sample comprised students preparing for and taking high-stakes upper-secondary exit examinations. All CCLPMs showed statistically significant autoregressive paths for achievement, test anxiety, and emotion regulation, supporting Hypothesis 1. Test anxiety showed negative lagged reciprocal relations with achievement, supporting Hypothesis 2, and positive contemporaneous reciprocal relations with rumination and catastrophizing, supporting Hypothesis 3. In addition, positive unidirectional relations from test anxiety to ER were shown for acceptance and self-blame, and negative unidirectional relations from test anxiety to ER were shown for refocus on planning and positive reappraisal. A negative unidirectional relation was shown from putting into perspective to test anxiety. These unidirectional relations provide partial support for Hypothesis 3.

Reciprocal Relations Between Test Anxiety and Achievement

Numerous studies have shown negative relations between test anxiety and achievement (Hembree, 1988; von der Embse et al., 2018). Theoretically, this relationship would be expected to be reciprocal (Pekrun, 1992, 2006, 2024). While studies have shown how test anxiety negatively predicts achievement while controlling for prior achievement or ability (e.g., Putwain et al., 2013, 2015), studies examining how test anxiety impacts subsequent achievement, or reciprocal studies are largely lacking. Findings of the present study confirm that, like related forms of achievement anxiety (e.g., Lichtenfeld et al., 2023; Pekrun, Marsh, Suessenbach, et al., 2023), test anxiety and achievement were related in a negative reciprocal fashion; higher test anxiety predicted lower subsequent achievement, which, in turn, predicted higher subsequent test anxiety.

Students with lower achievement, or higher test anxiety, may therefore become trapped in a vicious cycle of detriment to their well-being and educational flourishing. The size of the coefficients for relations between test anxiety and achievement, after controlling for autoregressive paths, were substantial (achievement to test anxiety, βs = -.07 to -.14; test

anxiety to achievement, βs = -.13 to -.28; see Orth et al., 2024, for an interpretation of effect sizes in longitudinal analyses). From a CVT perspective (Pekrun, 2006, 2017, 2024), anxiety can motivate persons to invest short-term effort to reduce failure (also see Eysenck et al., 2007), but also disrupt cognitive processes when learning, studying for, and taking exams (Moran, 2016). Anxious students rely on rigid forms of learning (Spada et al., 2006; Spada & Moneta, 2014), have reduced working memory capacity (Angelidis et al., 2019), and can avoid exam study (Cassady et al., 2024). In the long term, anxiety would also undermine motivation (e.g., Hancock, 2001; Sutter-Brandenberger et al., 2018). The net result is a negative relation between test anxiety and long-term achievement.

Lower achievement on tests, homework, and class assignments provides competence feedback that would lead to lower academic self-concept. This is an important self-belief that underpins control appraisals in CVT (Pekrun, 2006, 2017, 2018) and inputs to executive processes in the S-REF model (Wells, 2009; Zeidner & Matthews, 2005). Reduced competence beliefs and perceptions of control would increase the likelihood of threat appraisals and subsequent test anxiety. Evidence for related forms of anxiety (e.g., mathematics anxiety) has shown lower achievement is related to high anxiety through lower academic self-concept (Kyttälä & Björn, 2010; Van der Beek et al., 2017).

Notwithstanding that T_1 achievement was measured on a 9-point scale, and T_3 and T_5 achievement on a 7-point scale, it is also salient to briefly explain why the achievement scores declined from T_1 to T_5. This is likely a result of T_3 and T_5 achievement being assessed at a higher level than T_1. Moreover, T_3 achievement was based on curriculum content for Year 12 only whereas T_5 achievement covered curriculum content from both Years 12 and 13. While the level of assessment was the same, the exam demands for T_5 were greater than for T_3.

Reciprocal Relations Between Test Anxiety and ER

Studies have shown how nonadaptive ER strategies, such as catastrophizing and self-blame, positively relate to anxiety problems, while adaptive strategies, like reappraisal and planning, are negatively related (Garnefski et al., 2007; Garnefski & Kraaij, 2018). Furthermore, test anxiety is positively related to ER strategies such as avoidance and suppression and negatively related to reappraisal (Burić et al., 2016; Harley, Jarrell, et al., 2019). While there are theoretical reasons to anticipate reciprocal relations between test anxiety and ER (e.g., Harley, Pekrun, et al., 2019; Zeider & Matthews, 2005), evidence has been lacking. To the best of our knowledge, the current study is the first to explore these possible reciprocal relations. We showed unidirectional and reciprocal links between test anxiety and ER that were replicated at T_2 and T_4, adding to the robustness of the findings.

Positive reciprocal relations were shown between test anxiety and rumination and catastrophizing. In the S-REF model (Wells, 2009; Zeider & Matthews, 2005), such strategies intensify executive monitoring of threat, resulting in elevated anxiety. Elevated anxiety, in turn, results in continued use of rumination and threat in further ineffective attempts to control threat. Similarly, in the ERAS (Harley, Pekrun, et al., 2019), rumination and catastrophizing focus attention on the self (emotions and negative outcomes) at the expense of task activity (e.g., exam preparation), leading to greater threat appraisal and anxiety. Similarly to the S-REF model (Wells, 2009; Zeidner & Matthews, 2005), anxiety then prompts further attention focus on the self in an ineffective method to control threat.

Relations between test anxiety and putting into perspective were unidirectional. Specifically, test anxiety was reduced following putting things in perspective rather than vice versa. Putting into perspective was an adaptive strategy in that it was associated with lower anxiety. In the S-REF model (Wells, 2009; Zeidner & Matthews, 2005), keeping a perspective would terminate executive threat processing, and it is conceptualized as an adaptive strategy

in the ERAS (Harley, Pekrun, et al., 2019). Conceptually, in the S-REF model and the ERAS, situational threat appraisals are reduced, hence anxiety is lower.

Positive reappraisal and refocus on planning showed negative unidirectional relations suggesting that test anxiety reduced positive reappraisal and refocus on planning rather than vice versa. These are adaptive strategies that maintain task focus and would potentially help to reduce the likelihood of future failure and lower resultant test anxiety. It would be beneficial if highly test anxious students responded with greater, rather than less, use of these strategies. However, as noted above, in the S-REF model (Wells, 2009; Zeidner & Matthews, 2005), test anxiety prompts the continued use of ineffective ER (e.g., rumination) that maintains anxiety rather than positive reappraisal. Furthermore, anxious persons are rigid in their use of ER strategies (e.g., Hollenstein, 2015) which will further contribute to an ongoing cycle of executive threat processing.

Acceptance and self-blame showed positive unidirectional relations suggesting that test anxiety promoted the use of these strategies. It could be argued that these strategies were not non-adaptive since they followed rather than preceded higher test anxiety. They may be non-adaptive, however, in other ways. For instance, higher T_2 acceptance was associated with lower T_5 achievement. While active non-judgemental acceptance of one's emotional state can provide a degree of cognitive flexibility to overcome rigid use of ineffective ER (e.g. Swain et al., 2013), this may not be the case with resigned situational acceptance as measured in the present study. Passive resigned situational acceptance that a past situation cannot be changed (e.g., exam failure) may result in 'giving up' and a failure to strategize avoiding that outcome in the future (Nakamura & Orth, 2005). While resigned situational acceptance may not directly lead to enhanced test anxiety, without a change of strategy there is a chance that test anxiety reoccurs in the future.

Self-blame involves an internal attribution characterized by a pejorative judgmental pattern of thought common across various anxiety problems and emotion disorders (Legerstee et al., 2010; Rodríguez-Menchón et al., 2021). In the S-REF model (Wells, 2009; Zeidner & Matthews, 2005), failure feedback prompts executive processes to access negative core beliefs about incompetence, inability to cope, and worthlessness that subsequently promote a pejorative attributional style. Like passive resigned situational acceptance, self-blame may not directly lead to further anxiety but is ineffective in that there is a focus on failure rather than adaptive practical steps to avoid that outcome in the future.

Achievement and ER

Although we did not specifically hypothesize relations between achievement and ER, expecting them to be largely indirect through test anxiety, there were notable relations beyond the variance in achievement accounted for by test anxiety. Many of these relations were in plausible directions. For example, higher achievement was associated with lower other-blame. Presumably with higher grades, there is no need to blame others, such as one's teachers. Some relations, however, were counterintuitive. Rumination showed positive reciprocal relations with achievement. This would not be expected given that rumination is considered non-adaptive and associated with higher test anxiety. When rumination occurs after a stressful event and is focused on failure to make satisfactory progress toward avoiding failure in subsequent situations, individuals may simultaneously appraise the situation as a threat and not give up on their goal (Schultheiss et al., 2008; Smith & Alloy, 2009). Thus, after an exam that did not go as well as hoped for, a student could plausibly experience greater test anxiety and still strive towards achieving academic goals. This interpretation is consistent with the S-REF model (Wells, 2009; Zeidner & Matthews, 2005) whereby goal discrepancy (i.e., between academic goals and prior achievement) would involve continued executive threat monitoring that maintains anxiety. It is also consistent with Pekrun's (2006,

2024) CVT which posits that worries and anxiety can motivate people to invest effort to avoid failure. Indeed, Krys et al. (2020) showed that undergraduate students' rumination about a difficult statistics exam was positively related not only to stress (defined as unpredictable and excessive academic demands) but also to achievement on the statistics exam.

Optimal Time Intervals and Implications for Study Design

The six-month interval between test anxiety and achievement in successive waves was sufficient to show substantive, statistically significant negative relations. In contrast, coefficients were non-significant for the twelve-month interval between T_2 test anxiety and T_5 achievement and small for the twelve-month interval between T_1 achievement and T_4 test anxiety. This suggests that relations may gradually fade over time (see Dorman & Griffin, 2015). Theoretically, changes in test anxiety would occur in response to success/failure feedback and the resulting change in self-beliefs and self-regulatory processes. It is plausible that the effects of ongoing feedback from classwork, homework, and tests weaken the impact of the earlier exam grades. For instance, students may perform well in homework and informal class tests over twelve months, strengthen their control beliefs, and gradually lower their test anxiety. Hence, the link from achievement to subsequent test anxiety fades over the longer interval as feedback gradually shapes self-beliefs.

In terms of the reverse causal direction, some of the effects of test anxiety on achievement are more immediate (e.g., interference during a test) and others longer-term (e.g., shallow study strategies). For example, shorter intervals may capture the effects of test anxiety on achievement mediated by cognitive interference, whereas longer intervals could capture effects mediated by motivation and study strategy. It is plausible that the six-month interval captured the negative effects of test anxiety on motivation and study strategy. Over a twelve-month period, however, there is greater scope for students to be able to change their

motivations and study strategies in positive ways to weaken the relation with subsequent achievement.

Despite measuring test anxiety as a trait and ER as habitual strategies, the twelve-month interval between successive measurements of test anxiety and ER strategies revealed few statistically significant relations (see Supplementary Materials, Table S5). This may be due in part to anxiety and self-regulatory processes changing over time in ways to weaken possible relations, and partly due to ER strategies being reported in relation to difficulties in high-stakes exams at T_1 and T_3. While the outcomes of these tests are likely to be particularly salient for students' personal, academic, and occupational goals, subsequent academic events may prompt other more or less effective ER strategies that impact test anxiety. For theoretical reasons, we modeled relations between test anxiety and ER as contemporaneous. However, they represent sequential processes within short time lags rather than being truly simultaneous. We could expect significant relations between test anxiety and ER if measured over intervals lasting weeks, or perhaps months, even if they do not persist over a twelve-month period. Future studies should include shorter intervals that could be assistive in strengthening claims for directionality in test anxiety and ER relations.

Study Limitations and Suggestions for Future Research

Our study showed novel, statistically significant negative lagged reciprocal relations between test anxiety and achievement, and contemporaneous unidirectional and reciprocal relations between test anxiety and ER. Nonetheless, there are a few limitations to highlight. First, as noted, lagged assessments of test anxiety and ER using shorter time intervals would allow for a more robust test of directionality. Future studies could use, for example, intervals of several weeks or a month between the measurement of test anxiety and ER, rather than the twelve-month interval used in the present study. Alternatively, studies could use more intensive within-person designs measuring test anxiety and ER once or more times a day on

successive days over a defined period. The latter approach, however, may not necessarily generate sufficient variance in ER strategies to effectively test relations with test anxiety given the rigidity in ER strategies used by anxious persons (Hollenstein, 2015). Studies adopting the former approach could instruct participants to respond about ER strategies used more generally in response to academic, or exam-related, difficulties which may assist in capturing typical strategies used.

Second, we did not measure all possible types of ER strategies corresponding to all stages of the emotion-generative process as considered in the ERAS and PMER. Response-oriented strategies like suppression or diaphragmatic breathing were not included in the present study. Establishing relations between test anxiety and ER strategies used at different stages in the ERAS and PMER emotion generative process would facilitate the theorization of directional relations. That is, whether ER strategies precede, or are a response to, test anxiety. In addition, there are behavioral strategies (e.g., Kraaij & Garnefski, 2019), such as help-seeking and avoidance, that students may use in relation to exams that theoretically would show unidirectional or reciprocal relations with test anxiety.

Furthermore, our findings regarding the relations between achievement and rumination were surprising. Future studies are required to establish if this was a unique spurious finding. In doing so, studies may also wish to include perfectionism as a possible common variable that drives ER strategies to meet goals that result in test anxiety *and* higher achievement. Furthermore, studies could differentiate between different facets of rumination that may be more (e.g., self-reflection) or less (e.g., perseverative focus on anxiety and the problematic situation) adaptive, and if these relate to test anxiety and achievement in different ways.

Relatedly, our assessment of ER strategies used an instruction that prompted students to reflect on a recent high-stakes exam that had not gone as well as hoped for. We reasoned

that this would be a situation that would elicit future test anxiety and prompt the use of ER strategies to manage this anxiety. While the specificity is necessary to provide a meaningful context for the use of ER, it does somewhat limit generalizability to other situations where exams went poorly, and grades were lower than expected. While this type of incident may not be uncommon, it is one of several that might prompt ER strategies to manage test anxiety. Other situations that could elicit test anxiety (e.g., pressure from self, teachers, or family) in conjunction with personal variables, could result in alternate ER strategies. Future studies could explore if different triggers for self-referent processes about exams elicit different uses of ER.

Third, we relied on self-reported grades for achievement at T_1 and socio-demographic data at T_2 and T_4. While this expedited data collection, there is a risk of inaccurate data being reported. The secondary school exit exam grades used for T_1 achievement data are a truly once-in-a-lifetime experience (Brown & Woods, 2022) that would help mitigate retrospective misremembering. In addition, student-reported grades have been shown to correlate strongly with grades drawn from official records (Sticca et al., 2017). Moreover, there is no guarantee that official records will be entirely error-free. Indeed, there may be some benefits from using self-reported data. In England, student sex is recorded as a binary characteristic. For gender, where students may not identify with their birth sex, there is no option other than to ask students to self-report. Using self-reported achievement and socio-demographics also reduces the burden on participating institutions. For future studies, we recommend that researchers carefully consider the advantages and disadvantages of using and combining self-reports with official records.

Fourth, we adopted a specific approach to measuring test anxiety, which considered its cognitive, affective, and physiological components but not expressive or motivational elements. This may have somewhat constrained our assessment. Future studies should

consider the advantages of using a broader approach to measuring test anxiety. The motivational component (e.g., avoidance motivation) could provide a more balanced assessment of test anxiety and demonstrate specific links to behavioural variables, such as emotion regulation, behavioral problems (e.g., procrastination), and achievement. The expressive component may be relevant for understanding how other people react to test-anxious individuals. Such advantages should be balanced against the potential disadvantages (see Putwain et al., 2021). The distinction between indicators and effects of test anxiety can become blurred with motivation; some expressive components (e.g., staring into space) may not adequately differentiate between on-task behavior (e.g., thinking about the answer to an exam question) and test anxiety-induced off-task behavior (e.g., distraction by worries).

Fifth, a nuanced assessment of the impact of ethnic heritage was not possible with our sample. Such a study would require a substantially larger sample with sufficient numbers of participants from the major ethnic backgrounds represented in the English population (i.e., those used by the Office for National Statistics of the UK). Future studies could include such samples in the UK and beyond, ideally nationally representative. Alternatively, variables related to educational pressures (such as test anxiety and ER) could be added to existing large-scale surveys. These data would allow differences in ethnic heritage to be investigated for test anxiety and ER. In addition, the possible moderating effect of ethnicity on relations between ER, test anxiety, and achievement, could be tested to establish whether effects are generalizable across populations.

Sixth and finally, the reciprocal lagged relations between achievement and test anxiety were in line with theory. Nonetheless, there are mediating mechanisms to explain these links that we could not account for with the present dataset. Achievement would impact subsequent test anxiety through changes to underlying competence beliefs and control appraisals. Test anxiety would impact subsequent achievement through cognitive interference

when studying for and taking exams, the use of narrow information processing strategies when learning and studying for exams, and avoidance motivation. Subsequent longitudinal studies should incorporate measures of competence/control beliefs and cognitive-motivational component processes of learning to investigate these mediational links.

Implications for Intervention

Cognitive-behavioral interventions (CBIs) are well-established as effective methods of reducing test anxiety (e.g., Putwain et al., 2022). They typically combine specific elements of cognitive, emotional, physiological, and behavioral intervention synergistically, with psychoeducation (Williams & Chellingsworth, 2010). Modern CBIs can flexibly incorporate a variety of strategies into this framework including emotion regulation skills. Indeed some 'classic' emotional intervention techniques such as progressive muscle relaxation, diaphragmatic breathing, and visualization are, in PMER terms, response modulation ER strategies.

The findings of the present study indicate that rumination, catastrophizing, and putting into perspective, in particular, would be beneficial cognitive intervention foci for test anxiety CBI. Specifically, CBI strategies could be used to disrupt rumination and catastrophizing, notwithstanding the possible benefit for achievement resulting from rumination, and putting into perspective can be developed. Intervention strategies could assist students to identify and challenge ruminative thinking and catastrophizing in numerous ways. These could include assessing depersonalized evidence for and against catastrophic outcomes (the "what would stand up in a court of law" question) and using scaling questions to reduce the perception of catastrophic outcomes. Evidence-based thinking can be used to generate plausible non-catastrophic alternate outcomes. In addition, identifying and challenging metacognitive beliefs that drive rumination (e.g., that worrying is helpful) could also be assistive. Putting in perspective can be developed with exercises to identify the costs of becoming test anxious,

writing about test anxious experiences from the perspective of a detached observer, listing positive things about one's life, and thinking about how the event will be viewed in the future. These types of strategies could be delivered within the context of a CBI-based psychological intervention, but could also be included within school- or health-based curricula or well-being support.

Conclusion

Using novel CCLPMs to test lagged relations between test anxiety and achievement, and contemporaneous relations between test anxiety and ER, our findings confirm theoretical predictions that test anxiety is reciprocally and negatively related to achievement over time. Intervention to break this harmful cycle could focus on reducing test anxiety, building achievement through mastery and instructional intervention, or both. Relations between test anxiety and ER were contemporaneous. Positive reciprocal relations were shown between test anxiety and catastrophizing/rumination. Furthermore, putting into perspective had a negative contemporaneous effect on test anxiety. These forms of ER are thereby emphasized as important foci for intervention. Rumination and catastrophizing can be identified and challenged, and putting into perspective developed. Test anxiety had positive contemporaneous effects on acceptance and self-blame. Although they may not lead to higher test anxiety, they may be non-adaptive in other, indirect ways. Our study demonstrated the benefit of considering both lagged and contemporaneous effects in longitudinal designs, and the importance of judging intervals between measurement waves in relation to theory.

References

Angelidis, A., Solis, E., Lautenbach, F., van der Does, W., & Putman, P. (2019). I'm going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. *PloSOne, 14*(2), e0210824. https://doi.org/10.1371/journal.pone.0210824

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621– 634. http://dx.doi.org/10.1037/edu0000163

Author and Author (2025). Test Anxiety, emotion regulation, and academic buoyancy. https://osf.io/xvfsy/?view_only=ac56440662b14e17854c11ed05f17413

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin, 147*(2), 134–168. https://doi.org/https://doi. org/10. 1037/ bul0000307

Bonanno, G. A., Papa, A., Lalande, K., Westphal, M., & Coifman, K. (2004). The importance of being flexible - The ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science, 15*(7), 482-487. https://doi.org/10.1111/j.0956-7976.2004.00705.x

Brown, K., & Woods, K. (2022). Thirty years of GCSE: A review of student views and experiences. *Assessment in Education: Principles, Policy & Practice, 29*(1), 51–76. https://doi.org/10.1080/0969594X.2022.2053946

Burić, I., Sorić, I., & Penezić, Z. (2016). Emotion regulation in academic domain: Development and validation of the academic emotion regulation questionnaire

(AERQ). *Personality and Individual Differences, 96*, 138-147. https://doi.org/10.1016/j.paid.2016.02.074

Carnoy, M. (2005). Have state accountability and high-stakes tests influenced student progression rates in high school? Educational Measurement: Issues and Practice, 24(4), 19-31. https://doi.org/10.1111/j.1745-3992.2005.00020.x

Cassady, J.C., Helsper, A., & Quagliano, Q. (2024) The Collective Influence of Intolerance of Uncertainty, Cognitive Test Anxiety, and Academic Self-Handicapping on Learner Outcomes: Evidence for a Process Model. *Behavioral Sciences,* 14(2):96. https://doi.org/10.3390/bs14020096

Cécillon, F.-X. Mermillod, M. Leys, C. Lachaux, J.-P. Le Vigouroux, S. & Shankland, R. (2024). Trait Anxiety, Emotion Regulation, and Metacognitive Beliefs: An Observational Study Incorporating Separate Network and Correlation Analyses to Examine Associations with Executive Functions and Academic Achievement. *Children, 11*, 123. https://doi.org/10.3390/children1101012

Clark, D. M. and Wells, A. (1995). A cognitive model of social phobia. In Heimberg, R. G., Liebowitz, M. R., Hope, D. A. and Schneier, F. R. (Eds.), *Social Phobia: diagnosis, assessment and treatment* (pp. 69–93). Guilford Press.

Davis, H. A., DiStefano, C., & Schutz, P. A. (2008). Identifying patterns of appraising tests in first-year college students: Implications for anxiety and emotion regulation during test taking. *Journal of Educational Psychology, 100*(4), 942–960. https://doi.org/10.1037/a0013096

Decuir-Gunby, J. T., Aultman, L. P., & Schutz, P. A. (2009). Investigating transactions among motives, emotional regulation related to testing, and test emotions. *The Journal of Experimental Education, 77*(4), 409-438. https://doi.org/10.3200/jexe.77.4.409-43

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory into Practice, 55*(2), 153-159. https://doi.org/10.1080/00405841.2016.1148989

Djambazova-Popordanoska, S. (2016). Implications of emotion regulation on young children's emotional wellbeing and educational achievement. *Educational Review*, *68*(4), 497–515. https://doi.org/10.1080/00131911.2016.1144559

Dorman, C., & Griffin, M. A. (2015). Optimal time lags in panel studies. *Psychological Methods, 20*(4), 489–505. https:// doi. org/ 10. 1037/ met00 00041

Dormann, C., & van de Ven, B. (2014). Timing in methods for studying psychosocial factors at work. In M. Dollard, A. Shimazu, R. B. Nordin, P. Brough, & M. Tuckey (Eds.), *Psychosocial factors at work in the asia pacific* (pp. 89–116). Springer. https://doi.org/10.1007/978-94-017-8975-2_4.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336. https://doi.org/10.1037/1528-3542.7.2.336

Fenouillet, F., Nelson, V., Lorant, S., Masson, J., & Putwain, D.W. (2023). French study of Multidimensional Test Anxiety Scale in relation to performance, age and gender. *Journal of Psychoeducational Assessment, 41*(7), 828–834. https://doi.org/10.1177/073428292311868

Forsblom, L., Pekrun, R., Loderer, K., & Peixoto, F. (2021). Cognitive appraisals, achievement emotions, and students' math achievement: A longitudinal analysis. *Journal of Educational Psychology, 114*(2), 346–367. https://doi.org/10.1037/edu00 00671

Garnefski, N., & Kraaij, V. (2007). The Cognitive Emotion Regulation Questionnaire. Psychometric features and prospective relationships with depression and anxiety in adults. *European Journal of Psychological Assessment, 23*(3), 141–149. https://doi.org/10.127/1015-5759.23.3.141

Garnefski, N., & Kraaij, V. (2018). Specificity of relations between adolescents' cognitive emotion regulation strategies and symptoms of depression and anxiety. *Cognition and Emotion, 32*(7), 1401–1408. https://doi.org/10.1080/02699931.2016.1232698

Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual Differences, 30*(8), 1311-1327. https://doi:10.1016/S0191-8869(00)00113-6

Garnefski, N., Kraaij, V. & Spinhoven, P. (2002). *Manual for the use of the cognitive emotion regulation questionnaire*. The University of Leiden

Garnefski, N., Rieffe, C., Jellesma, F., Meerum Terwogt, M., & Kraaij, V. (2007). Cognitive emotion regulation strategies and emotional problems in 9–11-year-old children: The development of an instrument. *European Child & Adolescent Psychiatry, 16*(1), 1–9. https://doi.org/10.1007/s00787-006-0562-3

Garner, P. W. (2010) Emotional competence and its influences on teaching and learning. *Educational Psychology Review, 22*, 297–321. https://doi.org/10.1007/s10648-010-9129-4

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26(*1), 1-26. doi: /10.1080/1047840X.2014.940781

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271–299. https://doi.org/10.1037/1089-2680.2.3.271

Hancock, D. R. (2001). Effects of test anxiety and evaluative threat on students' achievement and motivation. *The Journal of Educational Research, 94*(5), 284–290 https:// doi. org/ 10. 1080/ 00220 67010 95987 64.

Harley, J. M., Jarrell, A., & Lajoie, S. P. (2019). Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. *Instructional Science, 47*, 151-180. https://10.1007/s11251-018-09480-z

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist, 54*(2), 1–21. https://doi.org/10.1080/00461520.2019.1587297

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods, 16*(3), 319–336. https://doi.org/10.1037/a0024917

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58(*1), 47–77. https://doi.org/10.2307/1170348.

Hollenstein, T. (2015). This time, it's real: affective flexibility, time scales, feedback loops, and the regulation of emotion. *Emotion Review, 7*(4), 308–315. https://doi.org/10.1177/1754073915590621

Hollenstein, T., & Lanteigne, D. M. (2018). Emotion regulation dynamics in adolescence. In P. M. Cole & T. Hollenstein (Eds), Emotion regulation: A matter of time (pp. 158–176). Routledge. https://doi.org/10.4324/9781351001328

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ireland, M. J., Clough, B. A., & Day, J. J. (2017). The cognitive emotion regulation questionnaire: Factorial, convergent, and criterion validity analyses of the full and short versions. *Personality and Individual Differences, 110*, 90–95. https://doi.org/10.1016/j.paid.2017.01.035

Jeličić, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology, 45*(4), 1195–1199. https://doi.org/10.1037/a0015665

Jones, B. D. (2007). The unintended outcomes of high-stakes testing. *Journal of Applied School Psychology, 23*(2), 65-86. https://doi.org/10.1300/J370v23n02_05

King, R.B., Cai, Y. & Elliot, A.J. (2024). Income inequality is associated with heightened test anxiety and lower academic achievement: A cross-national study in 51 countries. *Learning and Instruction, 89*, 101825. https://doi.org/10.1016/j.learninstruc.2023.101825

Krys, S., Otte, K. P., & Knipfer, K. (2020). Academic performance: A longitudinal study on the role of goal-directed rumination and psychological distress. *Anxiety, Stress, & Coping, 33*(5), 545-559. https://doi.org/10.1080/10615806.2020.1763141

Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE Life Sciences Education, 18*(1)1-17. https://doi.org/10.1187/cbe.18-04-0064

Kyttälä, M., & Björn, P.M. (2010). Prior mathematics achievement, cognitive appraisals and anxiety as predictors of Finnish students' later mathematics performance and career orientation, *Educational Psychology, 30*(4), 431-448.doi: 10.1080/01443411003724491

Kraft, L., Ebner, C., Leo, K., & Lindenberg, K. (2023). Emotion regulation strategies and symptoms of depression, anxiety, aggression, and addiction in children and adolescents:

A meta-analysis and systematic review. *Clinical Psychology: Science and Practice, 30(*4), 485–502. https://doi.org/10.1037/cps0000156

Kraaij, V., & Garnefski, N. (2019). The behavioral emotion regulation questionnaire: development, psychometric properties and relationships with emotional problems and the cognitive emotion regulation questionnaire. *Personality and Individual Differences, 137*, 56-61. https://doi.org/10.1016/j.paid.2018.07.036

Kristman, V, L., Manno, M., & Cote, P. (2005). Methods to account for attrition in longitudinal data: Do they work? A simulation study. *European Journal of Epidemiology, 20*, 657–662. https://doi.org/10.1007/s10654-005-7919-7

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria what did they really say? *Organizational Research Methods*, *9*(2), 202–220. https://doi.org/10.1177/1094428105284919

Lazarus, R., & Folkman, S. (1984). *Stress, appraisal and coping*. Springer.

Legerstee, J.S., Garnefski, N., Jellesma, F.C., Verhulst, F.C., & Utens, E.M.W.J (2010) . Cognitive coping and childhood anxiety disorders. *European Journal of Adolescent Psychiatry, 19*, 143–150. https://doi.org/10.1007/s00787-009-0051-6

Lichtenfeld, S., Pekrun, R., Marsh, H. W., Nett, U. E., & Reiss, K. (2023). Achievement emotions and elementary school children's academic performance: Longitudinal models of developmental ordering. *Journal of Educational Psychology, 115*(4), 552–570. https://doi.org/10.1037/edu0000748

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83*(404), 1198–1202. https://doi.org/10.2307/2290157

Lotz, C., & Sparfeldt, J.R. (2017). Does test anxiety increase as the exam draws near? – Students' state test anxiety recorded over the course of one semester. *Personality and Individual Differences 104*, 397–400. http://dx.doi.org/10.1016/j.paid.2016.08.032

Machin, S., McNally, S., & Ruiz-Valenzuela, J. (2020). Entry through the narrow door: The costs of just failing high stakes exams. *Journal of Public Economics, 190*, 104224. https://doi.org/10.1016/j.jpubeco.2020.104224

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646-656. https://doi.org/10.1037/0022-0663.82.4.646

Marsh, H.W. Guo, J., Dicke, T., Parker, P.D., & Craven. R.G. (2020). Confirmatory factor analysis (CFA), exploratory structural equation modelling (ESEM), and Set-ESEM: Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research, 55*(1), 102-119. http://dx.doi.org/10.1080/00273171.2019.1602503

Marsh, H.W., Guo, J., Pekrun, R. Lüdtke, O., & Núñez-Regueiro, F. (2024). Cracking chicken-egg conundrums: Juxtaposing contemporaneous and lagged reciprocal effects models of academic self-concept and achievement's directional ordering. *Educational Psychology Review, 36,* 53. https://doi.org/10.1007/s10648-024-09887-w

Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology, 10*, 85–110. http://dx.doi.org/10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331–353. https://doi.org/10.1037/edu0000281

Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin, 142*(8), 831–864. https://doi.org/10.1037/bul0000051

Morin, A.J., Marsh, H.W., & Nagengast, B. (2013). Exploratory structural equation modelling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A second course (2nd ed.)*. Information Age Publishing, Inc.

Morris, A. (2011). *Student standardised testing: Current practices in OECD countries and a literature review (OECD education working papers No. 65)*. OECD Publishing. https://doi.org/10.1787/19939019

Muthén, B., & Asparouhov, T. (2024). Can cross-lagged panel modeling be relied on to establish cross-lagged effects? The case of contemporaneous and reciprocal effects. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000661

Muthén, L. K., & Muthén, B. O. (2018). M*plus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nakamura, Y. M., & Orth, U. (2005). Acceptance as a coping reaction: Adaptive or not? *Swiss Journal of Psychology, 64*(4), 281–292. https://doi.org/10.1024/1421-0185.64.4.281

Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development, 41*(1), 143-153. https://doi.org/10.1177/0165025415618275

OECD. (2017). *PISA 2015 Results (Volume III): Students' Well-Being*. OECD Publishing. https://doi.org/10.1787/9789264273856-en

Omer, S.M., & Chen, S. (2023). Examining the dynamics of mathematics anxiety, perceived cost, and achievement: A control-value theory approach. *Contemporary Educational Psychology, 73*, 102169. https://doi.org/10.1016/j.cedpsych.2023.102169

Orth, U., Meier, L. L., Bühler, J. L., Dapp, L. C., Krauss, S., Messerli, D., & Robins, R. W. (2024). Effect size guidelines for cross-lagged effects. *Psychological Methods, 29*(2), 421–433. https://doi.org/10.1037/met0000499

Pekrun, R. (1991). Prüfungsangst und Schulleistung: Eine Längsschnittanalyse [Test anxiety and academic achievement: A longitudinal analysis]. *Zeitschrift für Pädagogische Psychologie, 5*, 99-109.

Pekrun, R. (1992). The expectancy-value theory of anxiety: Overview and implications. In D. G. Forgays, T. Sosnowski, & K. Wrzesniewski (Eds.), *Anxiety: Recent developments in self-appraisal, psychophysiological and health research* (pp. 23-41). Hemisphere.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2024). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review, 36*, 83. https://doi.org/10.1007/s10648-024-09909-7

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653-1670. https://doi.org/10.1111/cdev.12704

Pekrun, R., Marsh, H. W., Elliot, A. J., Stockinger, K., Perry, R. P., Vogl, E., Goetz, T., van Tilburg, W. A. P., Lüdtke, O., & Vispoel, W. P. (2023). A three-dimensional taxonomy of achievement emotions. *Journal of Personality and Social Psychology, 124*(1), 145–178. https://doi.org/10.1037/pspp0000448

Pekrun, R., Marsh, H. W., Suessenbach, F., Frenzel, A. C., & Goetz, T. (2023). School grades and students' emotions: Longitudinal models of within-person reciprocal effects.

Learning and Instruction, 83, 101626. https://doi.org/10.1016/j.learninstruc.2022.101626

Pekrun, R., and Stephens, E. J. (2009). Goals, emotions, and emotion regulation: perspectives of the control-value theory. *Human Development, 52*(6), 357–365. https://doi.org/10.1159/000242349

Preiss, R.W., Gayle, B.M., & Allen, M. (2006). Test anxiety, academic self-efficacy, and study skills: A meta-analytic review. In B.M. Gayle, R.W. Preiss & M. Allen (Eds.), *Classroom interaction and instructional processes: A meta-analytic review* (pp. 99-111). Lawrence Erlbaum Associates.

Putwain, D.W., Chamberlain, S., Daly, T., & Saddredini, S. (2015). Academically buoyant students are less anxious about and perform better in high-stakes examinations. *British Journal of Educational Psychology, 85*(3), 247–263. https://doi.org/10.1111/bjep.12068

Putwain, D. W., Wood, P., & Pekrun, R. (2022). Achievement emotions and academic achievement: Reciprocal relations and the moderating influence of academic buoyancy. *Journal of Educational Psychology, 114*(1), 108–126. https://doi.org/10.1037/edu0000637

Putwain, D.W. (2020). *Examination Pressures on Children and Young People: Are They Taken Seriously Enough?* London: The British Academy.

Putwain, D.W., & Daly, A.L. (2014). Test anxiety prevalence and gender differences in a sample of English secondary school students. *Educational Studies, 40*(5), 554–570. https://doi.org/10.1080/03055698.2014.953914

Putwain, D. W., Pekrun, R., Rainbird, E., & Roberts, C. (2022). Cognitive-behavioural intervention for test anxiety: Could teachers deliver the STEPS Program and what training would they require? In L. R. V. Gonzaga, L. L. Dellazzana-Zanon, & A. M.

Becker da Silva (Eds.), *Handbook of stress and academic anxiety* (pp. 381-399). https://doi.org/10.1007/978-3-031-12737-3_25

Putwain, D. W., von der Embse, N. P., Rainbird, E. C., & West, G. (2021). The development and validation of a new Multidimensional Test Anxiety Scale (MTAS). *European Journal of Psychological Assessment, 37*(3), 236-246. https://doi.org/10.1027/1015-5759/a000604

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. https://doi.org/10.1037/a0026838 Rottweiler, A-L., Taxer, J.L., Nett. U.E. (2018). Context Matters in the Effectiveness of Emotion Regulation Strategies. *AERA Open 4*(2), 1-13.doi: 10.1177/2332858418778849

Rogosa, D. (1980). A critique of cross-lagged correlation. Psychological Bulletin, 88(2), 245-258. https://doi.org/10.1037/0033-2909.88.2.245

Rohrer, J. M., & Murayama, K. (2023). These are not the effects you are looking for: Causality and the within-/between-persons distinction in longitudinal data analysis. *Advances in Methods and Practices in Psychological Science*, *6*(1), 25152459221140842. https://doi.org/10.1177/25152459221140842

Rodríguez-Menchón, M., Orgilés, M., Fernández-Martínez, I., Espada, J. P., & Morales, A. (2021). Rumination, catastrophizing, and other-blame: The cognitive-emotional regulation strategies involved in anxiety-related life interference in anxious children. *Child Psychiatry & Human Development, 52*(1), 63–76. https://doi.org/10.1007/s10578-020-00988-5

Schäfer, J. Ö., Naumann, E., Holmes, E. A., Tuschen-Caffier, B., & Samson, A. C. (2017). Emotion regulation strategies in depressive and anxiety symptoms in youth: A meta-

analytic review. *Journal of Youth and Adolescence, 46*(2), 261–276. https://doi.org/10.1007/s10964-016-0585-0

Scherer, K. R., & Moors, A. (2019). The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology, 70*(1), 719–745. https://doi.org/10.1146/annurev-psych-122216-011854

Schneider, M., & Peckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin, 143*(6),565-600. https://doi.org/10.1037/bul0000098

Schutz, P. A., & Davis, H. A. (2000). Emotions and self-regulation during test taking. *Educational Psychologist, 35*, 243–255. https://doi.org/10.1207/S15326985EP3504_03

Schutz, P. A., Davis, H. A., Decuir-Gunby, J. T., & Tillman, D. (2014). Regulating emotions related to testing. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), International handbook of emotions in education (p. 348–367). Routledge.

Schutz, P. A., Benson, J., & Decuir-Gunby, J. T. (2008). Approach/avoidance motives, test emotions, and emotional regulation related to testing. *Anxiety, Stress, and Coping, 21*(3), 263–281. https://doi.org/10.1080/10615800701787672

Seipp, B. (1991) Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27-41, https://doi.org/10.1080/08917779108248762

Schultheiss, O. C., Jones, N. M., Davis, A. Q., & Kley, C. (2008). The role of implicit motivation in hot and cold goal pursuit: Effects on goal progress, goal rumination, and emotional well-being. *Journal of Research in Personality*, *42*(4), 971–987. https://doi.org/10.1016/j.jrp.2007.12.009

Singh, L., Barokova, M., Bazhydai, M., Baumgartner, H. A., Franchin, L., Kosie, J. E., Lew-Williams, C., Omane, P. O., Reinelt, T., Schuwerk, T., Sheskin, M., Soderstrom, M., Wu, Y., & Frank, M. C. (2024). Tools of the trade: A guide to sociodemographic

reporting for researchers, reviewers, and editors. *Journal of Cognition and Development*. Advance online publication. https://doi.org/10.1080/15248372.2024.2431106

Smith, J. M., & Alloy, L. B. (2009). A roadmap to rumination: A review of the definition, assessment, and conceptualization of this multifaceted construct. *Clinical Psychology Review*, *29*(2), 116–128. https://doi.org/10.1016/j.cpr.2008.10.003

Spada, M. M., Nikcevic, A., Moneta, G. B., & Ireson, J. (2006) Metacognition as a mediator of the effect of test anxiety on a surface approach to studying. *Educational Psychology, 26(*5), 615–624. https://doi.org/10.1080/01443410500390673

Spada, M. M., & Moneta, G. B. (2014). Metacognitive and motivational predictors of surface approach to studying and academic examination performance. *Educational Psychology, 34(*4), 512-523. https://doi.org/10.1080/01443410.2013.814196

Spielberger, C. D., & Vagg. R.P. (1995). Test Anxiety: A Transactional Process Model. In by C. D. Speilberger & P. R. Vagg (Eds.) *Test Anxiety: Theory, Assessment and Treatment*, (pp. 3–14). Taylor & Francis.

Steinmayr, R., Crede, J., McElvany, N., & Wirthwein, L. (2015). Subjective well-being, test anxiety, academic achievement: testing for reciprocal effects. *Frontiers in psychology*, *6*:1994. https://doi.org/10.3389/fpsyg.2015.01994

Stevens, J.P. (1992). *Applied multivariate statistics for the social sciences (2nd edition)*. Erlbaum.

Sticca, F., Goetz, T., Bieg, M., Hall, N. C., Eberle, F., & Haag, L. (2017). Examining the accuracy of students' self-reported academic grades from a correlational and a discrepancy perspective: Evidence from a longitudinal study. *PLoS ONE, 12*(11), Article e0187367. https://doi.org/10.1371/journal.pone.0187367

Stockinger, K., Dresel, M., Marsh, H. W., & Pekrun, R. (2024). Strategies for regulating achievement emotions: Conceptualization and relations with university students' emotions, well-being, and health. *Learning and Instruction, 98*, 102089.https://doi.org/10.1016/j.learninstruc.2025.102089

Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and 'getting it' for the white working class. *Research Papers in Education, 29*(2), 131-171. https://doi.org/10.1080/02671522.2013.767370

Sutter-Brandenberger, C. C., Hagenauer, G., & Hascher, T. (2018). Students' self-determined motivation and negative emotions in mathematics in lower secondary education – Investigating reciprocal relations. *Contemporary Educational Psychology, 55*, 166–175. https://doi.org/10.1016/j.cedpsych.2018.10.002

Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and commitment therapy in the treatment of anxiety: A systematic review. *Clinical Psychology Review, 33*(8), 965–978. http://dx.doi.org/10.1016/j.cpr.2013.07.002.

Thompson, R. (2017). Opportunity structures and educational marginality: The post-16 transitions of young people outside education and employment. *Oxford Review of Education, 43*(6), 749–766. https://doi.org/10.1080/03054985.2017.1352502

Trentacosta, C. J., & Izard, C. E. (2007). Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion, 7*(1), 77–88. https://doi.org/10.1037/1528-3542.7.1.77

Tyson, D. F., Linnenbrink-Garcia, L., & Hill, N. E. (2009). Regulating debilitating emotions in the context of performance: achievement goal orientations, achievement-elicited emotions, and socialization contexts. *Human Development*, *52*(6), 329–356. https://doi.org/10.1159/000242348.

Van der Beek, J.P.J., Van der Ven, S.H.G., Kroesbergen, E.H., & Leseman, P.P.M. (2017). Self-concept mediates the relation between achievement and emotions in mathematics. *British Journal of Educational Psychology, 87*(3), 90-107.doi: 10.1111/bjep.12160

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

von der Embse, N. P., Putwain, D. W. & Francis, G. (2021). Interpretation and use of the Multidimensional Test Anxiety Scale (MTAS). *School Psychology, 36*(2), 86–96. https://doi.org/10.1037/spq0000427

Widaman, K.F., Ferrer, E., & Conger, R.D. (2010). Factorial Invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10-18. https://doi.org/10.1111/j.1750-8606.2009.00110.x

Wells, A. (2009). *Metacognitive therapy for anxiety and depression*. Guilford Press.

Wells, A., & Carter, K. (2001). Further tests of a cognitive model of generalized anxiety disorder: metacognitions and worry in GAD, panic disorder, social phobia, depression, and nonpatients. *Behavior Therapy, 32*(1), 85–102. https://doi.org/10.1016/S0005-7894(01)80045-9

Wenzel, K., & Reinhard, M. A. (2021). Does the end justify the means? Learning tests lead to more negative evaluations and to more stress experiences. *Learning and Motivation, 73*, 101706. https://doi.org/10.1016/j.lmot.2020.101706

Williams, C. & Chellingsworth, M. (2010). *CBT: A clinicians guide to using the five areas approach*. Hodder Arnold.

Wong, T. K., Colasante, T., & Malti, T. (2023). A longitudinal examination of school-related and mental health mediators linking emotion regulation to academic achievement. *Journal of School Psychology*, *101*, 101253. https://doi.org/10.1016/j.jsp.2023.101253

Ximénez, C. (2009). Recovery of weak factor loadings in confirmatory factor analysis under conditions of model misspecification. *Behavior Research Methods, 41*(4), 1038–1052. https://doi.org/10.3758/BRM.41.4.1038

Xu, J., Du, J., Liu, F., & Huang, B. (2019). Emotion regulation, homework completion, and math achievement: Testing models of reciprocal effects. *Contemporary Educational Psychology, 59*, 101810. https://doi.org/10.1016/j.cedpsych.2019.101810

Yu, K., & Martin, A. J. (2014). Personal best (PB) and 'classic' achievement goals in the Chinese context: their role in predicting academic motivation, engagement and buoyancy. *Educational Psychology, 34*(5), 635-658. https://doi.org/10.1080/01443410.2014.895297

Zeidner, M. (1998). *Test anxiety: State of the art*. Plenum.

Zeidner, M., & Matthews, G. (2005). Evaluation anxiety. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 141–163). Guilford Press.

Table 1

CERQ Emotion Regulation Strategies and Stage in the PMER/ERAS, and CVT

Strategy	Definition	Stage in PMER/ERAS and CVT
Refocus on planning	Thinking about what steps can be taken to deal with the negative event	Situation modification, competence development
Positive reappraisal	Attaching a positive meaning to the negative event as an opportunity for personal growth	Cognitive change
Putting into perspective	Reducing the severity of the negative event by comparing it to other situations	Cognitive change
Rumination	Persistent thinking about the thoughts and feelings associated with the negative event	Attention deployment
Catastrophizing	Explicitly magnifying the negative outcomes of the negative situation	Cognitive change
Positive refocusing	Thinking about pleasant alternatives to the negative event	Attention deployment
Self-blame	A pejorative judgement attributing responsibility for the negative situation or outcome to oneself	Cognitive change
Other-blame	A pejorative judgement attributing responsibility for the negative situation or outcome to others	Cognitive change
Acceptance	Resigning oneself that the negative situation cannot be changed	Cognitive change

Table 2
Descriptive Data for Test Anxiety and Emotion Regulation Strategies

	Mean	SD	ω	Skewness	Kurtosis	Factor Loadings
T_2 Acceptance	2.91	0.90	.72	0.25	-0.51	.20 - .84
T_2 Positive Refocusing	2.15	0.93	.86	0.67	-0.30	.72 - .84
T_2 Refocus on Planning	2.89	0.65	.76	0.07	-0.65	.60 - .71
T_2 Positive Reappraisal	2.52	0.77	.68	0.40	-0.28	.38 - .83
T_2 Putting Into Perspective	3.03	1.00	.80	0.05	-0.74	.53 - .80
T_2 Rumination	2.77	1.04	.82	0.27	-0.73	.57 - .83
T_2 Catastrophizing	2.26	0.97	.77	0.74	-0.13	.37 - .80
T_2 Self-Blame	2.26	1.03	.86	-0.16	-0.98	.63 - .84
T_2 Other-Blame	1.81	0.53	.75	0.80	0.85	.46 - .80
T_2 Test Anxiety	3.53	0.71	.75	-0.40	-0.24	.50 - .86
T_4 Acceptance	2.78	0.73	.63	0.30	-0.39	.19 - .78
T_4 Positive Refocusing	2.73	0.79	.89	0.23	-0.42	.74 - .85
T_4 Refocus on Planning	3.05	0.88	.79	0.10	-0.74	.63 - .76
T_4 Positive Reappraisal	2.95	0.86	.69	0.18	-0.44	.37 - .83
T_4 Putting Into Perspective	2.89	0.78	.81	-0.07	-0.46	.54 - .81
T_4 Rumination	2.58	0.76	.81	0.27	-0.32	.55 - .82
T_4 Catastrophizing	2.01	0.78	.78	0.72	-0.15	.38 - .86
T_4 Self-Blame	3.05	0.79	.85	0.14	-0.44	.74 - .88
T_4 Other-Blame	2.34	0.55	.80	-0.01	-0.25	.50 - .84
T_4 Test Anxiety	3.61	2.81	.75	-0.36	-0.08	.50 - .87
T_1 Achievement	6.56	1.08	—	0.24	-0.60	—
T_3 Achievement	4.03	1.47	—	-0.29	-0.75	—
T_5 Achievement	3.40	0.99	—	-0.06	-0.11	—

Note. Hierarchical omega (ω_H) was estimated for test anxiety.

Table 3

Latent Bivariate Correlations for Test Anxiety, Emotion Regulation, Achievement and Socio-demographic Variables

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1. T_2 Acceptance	**.11**	**-.24**	**-.18**	**.16**	**.17**	**.28**	**.30**	**.23**	**.24**	**.48**	.11	**-.29**	-.05	**.17**	.08	**.23**	**.20**	**.27**	**.16**	**-.11**	**-.20**	**-.22**	.03	-.04	-.02
2. T_2 Positive Refocusing		**.32**	**.31**	**.50**	-.02	-.05	-.08	**.17**	**-.15**	-.03	**.53**	**.15**	.13	**.30**	-.10	-.01	**-.13**	.08	**-.12**	-.01	-.08	-.02	**-.10**	-.03	.01
3. T_2 Refocus on Planning			**.35**	**.25**	**.17**	.04	**.09**	**-.10**	-.01	-.09	**.18**	**.56**	**.34**	**.18**	**.21**	.02	.04	**-.17**	-.06	.04	**.12**	.09	.05	-.03	**.12**
4. T_2 Positive Reappraisal				**.30**	-.07	**-.11**	**-.09**	-.04	**-.43**	**-.24**	**.22**	**.34**	**.31**	.11	-.14	.07	-.12	-.12	**-.20**	-.01	.02	.01	**-.16**	-.03	.09
5. T_2 Putting Into Perspective					-.04	**-.21**	.03	.07	**-.18**	-.02	**.36**	**.22**	**.27**	**.51**	.04	-.09	-.04	.03	**-.14**	.06	.06	-.04	-.01	-.04	**-.10**
6. T_2 Rumination						**.60**	**.50**	.06	**.65**	**.18**	-.10	.03	-.14	-.10	**.55**	**.45**	**.30**	.01	**.48**	**.12**	**.13**	.08	**.23**	.05	.01
7. T_2 Catastrophizing							**.40**	**.25**	**.66**	**.24**	-.10	**-.15**	-.13	**-.19**	**.41**	**.69**	**.30**	.11	**.53**	-.08	-.02	-.02	**.23**	.05	.09
8. T_2 Self-Blame								-.03	**.44**	**.28**	-.08	.01	-.11	-.07	**.33**	**.28**	**.57**	**.13**	**.35**	-.06	-.08	-.06	.02	.01	-.06
9. T_2 Other-Blame									.09	-.01	-.01	**-.15**	.05	.09	-.03	**.09**	-.05	**.50**	-.02	**-.17**	**-.14**	**-.13**	-.03	-.07	.01
10. T_2 Test Anxiety										**.36**	**-.13**	**-.17**	**-.44**	**-.12**	**.39**	**.49**	**.36**	.11	**.94**	-.05	-.08	-.09	**.48**	.03	-.08
11. T_4 Acceptance											**.11**	**-.18**	-.09	.06	**.24**	**.27**	.26	**.26**	**.36**	**-.09**	-.08	-.10	.10	-.01	-.01
12. T_4 Positive Refocusing												**.26**	**.42**	**.52**	-.05	**-.10**	-.07	**.14**	**-.13**	.02	.10	.01	-.09	-.02	.01
113. T_4 Refocus on Planning													**.41**	**.23**	**.14**	-.05	**.17**	**-.13**	**-.19**	**.24**	**.22**	**.17**	.03	-.08	-.05
14. T_4 Positive Reappraisal														**.42**	-.08	**-.18**	-.09	.02	**-.54**	.01	.03	.01	**-.21**	.04	**.21**
15. T_4 Putting Into Perspective															**-.14**	**-.24**	.04	.09	**-.20**	.04	-.04	-.06	**-.16**	**-.14**	**-.16**
16. T_4 Rumination																**.55**	**.45**	**.13**	**.55**	**.14**	**.24**	**.16**	**.20**	.01	.08
17. T_4 Catastrophizing																	**.37**	**.29**	**.64**	.05	-.09	-.10	**.11**	.01	.10
18. T_4 Self-Blame																		-.08	**.44**	-.02	-.09	**-.13**	-.05	-.04	-.05
19. T_4 Other-Blame																			.10	-.10	**-.12**	**-.11**	-.07	**-.13**	-.05
20. T_4 Test Anxiety																				-.07	**-.14**	**-.16**	**.44**	.01	-.05
21. T_1 Achievement																					**.50**	**.45**	**.11**	-.06	**-.13**
322. T_3 Achievement																						**.56**	**.15**	-.05	.05
23. T_5 Achievement																							**.11**	-.08	**-.20**
24. Gender																								.05	-.01
25. Free School Meals																									**.27**
26. Ethnicity																									

Note. **Bold** coefficients: $p < .05$.

Table 4

Model Fit for the Structural Equation Models

	χ^2	RMSEA	SRMR	CFI	TLI
Models with developmental equilibrium constraints:					
Acceptance	1333.54	.033	.071	.948	.937
Positive Refocusing	1196.89	.029	.053	.964	.956
Refocus on Planning	1258.86	.031	.061	.955	.946
Positive Reappraisal	1247.24	.031	.060	.950	.939
Putting Into Perspective	1186.84	.028	.053	.963	.955
Rumination	1257.56	.031	.057	.948	.949
Catastrophizing	1272.02	.031	.056	.956	.947
Self-Blame	1295.24	.032	.059	.956	.946
Other-Blame	1171.74	.028	.056	.964	.956
Models with all paths freely estimated:					
Acceptance	1300.08	.033	.070	.951	.940
Positive Refocusing	1170.33	.028	.051	.966	.958
Refocus on Planning	1232.62	.030	.060	.957	.948
Positive Reappraisal	1294.07	.032	.060	.950	.939
Putting Into Perspective	1157.61	.028	.052	.965	.957
Rumination	1227.97	.030	.056	.960	.951
Catastrophizing	1246.78	.031	.055	.959	.950
Self-Blame	1266.18	.032	.057	.958	.948
Other-Blame	1195.85	.029	.057	.961	.953

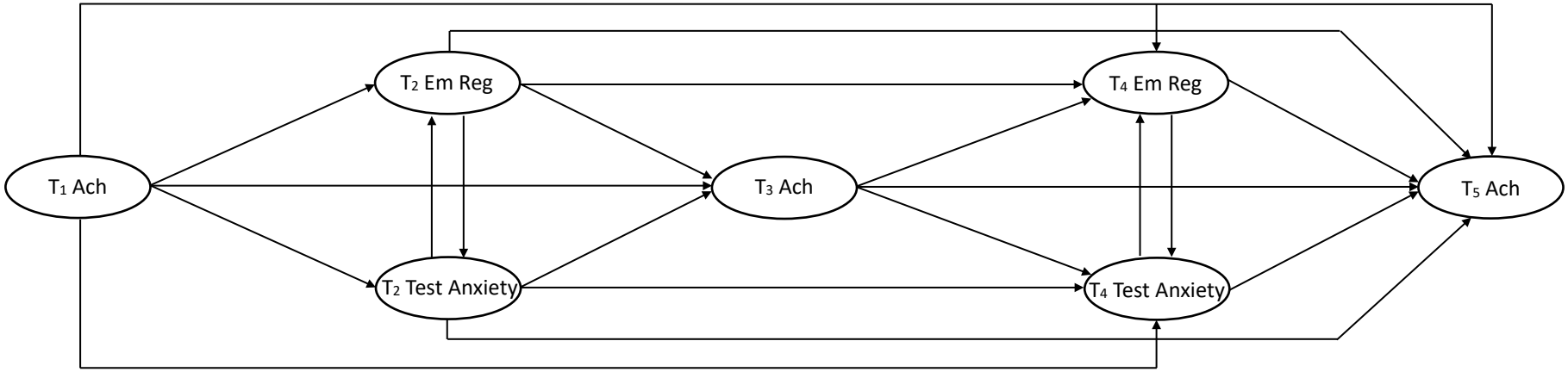
Note. χ^2 tests for all models were $p < .001$ (781 degrees of freedom for constrained, and 772 degrees of freedom for freely estimated models)

Table 5
Structural Equation Models: Standardized Path Coefficients

	Acceptance model		Positive refocusing model		Refocus on planning model		Positive reappraisal model		Putting into perspective model		Rumination model		Catstroph-izing model		Self-blame model		Other-blame model	
	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*
Autoregressive effects																		
T_1 Ach → T_3 Ach	**.32**	**.03**	**.33**	**.03**	**.32**	**.03**	**.31**	**.03**	**.32**	**.03**	**.31**	**.03**	**.33**	**.03**	**.32**	**.03**	**.32**	**.03**
T_1 Ach → T_5 Ach	**.14**	**.05**	**.14**	**.05**	**.14**	**.05**	**.15**	**.05**	**.15**	**.05**	**.15**	**.05**	**.13**	**.05**	**.15**	**.05**	**.14**	**.05**
T_3 Ach → T_5 Ach	**.54**	**.04**	**.55**	**.04**	**.53**	**.04**	**.52**	**.04**	**.54**	**.05**	**.51**	**.04**	**.55**	**.04**	**.54**	**.05**	**.54**	**.05**
T_2 TA → T_4 TA	**.77**	**.07**	**.74**	**.05**	**.75**	**.05**	**.72**	**.09**	**.73**	**.05**	**.54**	**.07**	**.51**	**.06**	**.71**	**.07**	**.75**	**.05**
T_2 ER → T_4 ER	**.37**	**.06**	**.54**	**.05**	**.60**	**.05**	**.43**	**.07**	**.53**	**.05**	**.37**	**.06**	**.49**	**.08**	**.44**	**.05**	**.51**	**.06**
Achievement & anxiety																		
T_1 Ach → T_2 TA	**-.09**	**.03**	**-.08**	**.03**	**-.10**	**.03**	**-.10**	**.04**	**-.08**	**.03**	**-.09**	**.03**	**-.07**	**.03**	**-.07**	**.03**	**-.09**	**.03**
T_1 Ach → T_4 TA	.02	.05	.01	.05	.02	.05	.02	.05	.02	.05	.01	.04	-.02	.04	.01	.05	.01	.05
T_3 Ach → T_4 TA	**-.12**	**.04**	**-.10**	**.04**	**-.12**	**.04**	**-.13**	**.05**	**-.10**	**.04**	**-.14**	**.04**	**-.07**	**.03**	**-.11**	**.04**	**-.11**	**.05**
T_2 TA → T_3 Ach	**-.16**	**.06**	**-.15**	**.05**	**-.17**	**.06**	**-.21**	**.08**	**-.13**	**.05**	**-.26**	**.08**	**-.16**	**.07**	**-.13**	**.06**	**-.15**	**.06**
T_2 TA → T_5 Ach	.17	.09	.10	.09	.13	.09	.13	.11	.07	.09	.12	.11	.10	.07	.08	.09	.11	.09
T_4 TA → T_5 Ach	**-.19**	**.07**	**-.18**	**.07**	**-.21**	**.07**	**-.28**	**.11**	**-.17**	**.07**	**-.27**	**.09**	**-.26**	**.11**	**-.15**	**.07**	**-.19**	**.07**
Anxiety & emotion regulation																		
T_2 TA → T_2 ER	**.43**	**.12**	-.04	.08	**-.19**	**.08**	**-.37**	**.10**	.01	.09	**.53**	**.12**	**.22**	**.07**	**.47**	**.11**	.15	.09
T_4 TA → T_4 ER	**.47**	**.10**	-.04	.07	**-.18**	**.08**	**-.40**	**.14**	.01	.09	**.46**	**.09**	**.35**	**.12**	**.41**	**.09**	.11	.07
T_2 ER → T_2 TA	-.04	.09	-.05	.05	-.02	.06	-.07	.10	**-.13**	**.07**	**.23**	**.09**	**.53**	**.06**	.05	.08	-.02	.05
T_4 ER → T_4 TA	-.04	.08	-.06	.06	-.02	.06	-.07	.09	**-.13**	**.06**	**.27**	**.10**	**.33**	**.06**	.08	.08	-.03	.07
Achievement & emotion regulation																		
T_1 Ach → T_2 ER	.03	.03	.05	.04	.01	.04	-.04	.04	.01	.04	**.13**	**.03**	-.02	.03	.01	.03	**-.10**	**.04**
T_1 Ach → T_4 ER	-.02	.06	-.01	.06	**.15**	**.06**	.05	.06	.05	.06	-.01	.05	.09	.05	.06	.05	.05	.06
T_3 Ach → T_4 ER	.05	.06	.05	.04	.02	.05	-.05	.06	.01	.05	**.18**	**.04**	-.03	.05	.01	.04	**-.09**	**.04**
T_2 ER → T_3 Ach	-.04	.04	-.07	.03	.01	.04	-.07	.05	-.01	.04	**.16**	**.06**	.10	.07	-.04	.02	-.01	.02
T_2 ER → T_5 Ach	**-.12**	**.06**	.06	.06	-.01	.07	-.02	.08	-.10	.06	-.10	.09	-.02	.10	.03	.06	-.04	.06
T_4 ER → T_5 Ach	-.04	.04	**-.10**	**.04**	.03	.05	-.09	.06	-.01	.04	**.20**	**.07**	.10	.07	-.05	.05	-.02	.05

Note. Ach = achievement; TA = test anxiety; ER = emotion regulation. **Bold** coefficients: $p < .05$.

Figure 1
Hypothesized Model



Note. Gender was included in the model as a covariate but omitted from Figure 1 to avoid over-cluttering.

- Supplementary Materials -

This document contains materials designed to supplement the main text. The materials include the following:

Additional Sample Information

Participants were drawn from three English 6th form colleges specialising in providing two year-courses (Years 12 and 13) in upper secondary education that are used to access higher education. Students were studying courses in General Certificate of Education: Advanced Level (A Level) and/or Business and Technology Education Council (BTEC) qualifications. Students typically study three A Level subjects. BTEC qualifications can be studied flexibly to be the equivalent of one, two, or three A Level subjects. Colleges can offer up to forty different A Level and BTEC subject options depending on staff expertise, resulting in multiple possible subject combinations.

The self-reported socio-demographic characteristics of participants at T_2 and T_4 are shown in Table S1. Over the two academic years when data were collected (2021-22 and 2022-23), 51% of students in English schools and 6th form colleges were male, 34.5% (2021-22) and 35.7% (2022-23) were from minority ethnic (i.e., non-white) backgrounds, and 22.5% (2021-22) and 23.8% (2022-23), were eligible for FSM (Department for Education, 2022, 2023). Our sample, therefore, was broadly characteristic of English schools and 6th form colleges for ethnic heritage but contained a smaller proportion of males and persons from low-income backgrounds than is typical.

References

Department for Education (2022). *Schools, pupils and their characteristics: January 2022*. His Majesties Stationary Office.

Department for Education (2023). *Schools, pupils and their characteristics: January 2023*. His Majesties Stationary Office.

Table S1
Socio-Demographics of Participants at Waves 2 and 4

	Wave 2	Wave 4
Gender:		
Male	138 (25.8%)	102 (22.6%)
Female	376 (70.2%)	329 (73.1%)
Prefer not to say	3 (0.6%)	3 (0.7%)
Non-binary	18 (3.4 %)	16 (3.6%)
Ethnicity		
Black	80 (15.0%)	72 (16.0%)
Chinese	3 (0.6%)	3 (0.7%)
South Asian	62 (11.6%)	47 (11.1%)
White	339 (63.4%)	282 (62.7%)
Other	27 (5.0%)	19 (4.2%)
Dual Heritage	24 (4.4%)	24 (5.3%)
Free School Meals	77 (14.4%)	60 (13.3%)
Total	535	450

Table S2
Longitudinal Invariance Tests for Test Anxiety and Emotional Regulation

	χ^2(df)	RMSEA	SRMR	CFI	TLI	ΔRMSEA	ΔCFI	ΔTLI
Hierarchical Test Anxiety								
Configural	708.84 (420)***	.032	.047	.964	.957			
Metric	717.40 (432)***	.032	.049	.964	.958	.000	.000	+.001
Scalar	742.34 (448)***	.031	.050	.962	.958	-.001	-.002	.000
Error	749.36 (464)***	.030	.050	.954	.961	-.001	-.008	+.003
Acceptance								
Configural	67.96 (15)***	.073	.059	.935	.880			
Metric	67.78 (18)***	.065	.059	.939	.906	-.008	+.004	+.026
Scalar	86.48 (22)***	.066	.042	.921	.900	+.001	-.018	-.006
Partial Scalar[a]	75.51 (21)***	.063	.044	.932	.910	-.002	-.007	+.004
Positive Refocusing								
Configural	34.08 (15)**	.044	.034	.986	.974			
Metric	37.79 (18)**	.040	.035	.986	.987	-.004	+.001	+.013
Scalar	43.39(22)**	.038	.042	.984	.980	-.002	-.002	-.007
Error	45.00 (26)*	.033	.044	.986	.985	-.005	+.002	+.005
Refocus on Planning								
Configural	27.06 (15)*	.035	.027	.986	.974			
Metric	20.62 (18)*	.030	.028	.988	.981	-.050	+.002	+.007
Scalar	44.87(22)**	.039	.045	.973	.966	+.009	-.015	-.015
Partial Scalar[a]	38.63 (21)**	.035	.037	.979	.972	-.005	-.009	-.009
Positive Reappraisal								
Configural	20.94 (15)	.024	.027	.990	.981			
Metric	21.43 (18)	.017	.029	.994	.991	-.007	+.004	+.010
Scalar	26.38 (22)	.017	.032	.993	.991	.000	-.001	.000
Error	34.00 (26)	.021	.044	.987	.986	+.004	-.006	-.005

Putting into Perspective								
Configural	19.76 (15)	.022	.025	.996	.992			
Metric	23.63 (18)	.022	.032	.995	.992	.000	-.001	.000
Scalar	27.13 (22)	.019	.035	.995	.994	-.003	.000	+.002
Error	30.10 (26)	.015	.039	.996	.996	-.004	+.001	+.002
Rumination								
Configural	16.84 (15)	.014	.019	.998	.997			
Metric	19.64 (18)	.012	.024	.999	.998	-.002	+.001	+.001
Scalar	24.66 (22)	.013	.024	.999	.997	+.001	.000	-.001
Error	25.87 (26)	.000	.025	1.000	1.000	-.013	+.001	+.003
Catastrophising								
Configural	37.48 (15)**	.047	.032	.978	.960			
Metric	42.77 (18)**	.045	.040	.976	.963	-.002	-.002	+.003
Scalar	47.90 (22)**	.042	.041	.975	.968	-.003	-.001	+.005
Error	50.16 (26)**	.037	.042	.977	.975	-.005	+.002	+.007
Self-Blame								
Configural	28.34 (15)*	.037	.032	.991	.983			
Metric	32.24 (18)*	.034	.036	.990	.985	-.003	-.001	+.002
Scalar	34.92 (22)*	.030	.038	.991	.989	-.004	+.001	+.004
Error	35.74 (26)	.024	.041	.993	.993	-.006	+.002	+.004
Other-Blame								
Configural	19.46 (15)	.021	.038	.990	.981			
Metric	25.08 (18)	.024	.048	.984	.975	+.003	-.006	-.006
Scalar	30.06 (22)	.023	.049	.982	.977	-.001	-.002	+.002
Error	29.95 (26)	.007	.055	.999	.998	-.016	+.017	+.011

Note. * p <.05. ** p <.01. *** p <.001.
[a] Constraint on item 2 relaxed.

Table S3
Structural Equation Models: Standardized Path Coefficients for Gender

	Acceptance model		Positive refocusing model		Refocus on planning model		Positive reappraisal model		Putting into perspective model		Rumination model		Catstroph-izing model		Self-blame model		Other-blame model	
	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*
Effects of gender																		
T_1 Ach	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**
T_3 Ach	**.14**	**.05**	**.14**	**.05**	**.14**	**.05**	**.16**	**.05**	**.14**	**.05**	**.17**	**.05**	**.15**	**.05**	**.14**	**.05**	**.14**	**.05**
T_5 Ach	.09	.05	.09	.05	.09	.05	**.11**	**.05**	.08	.05	**.12**	**.05**	**.12**	**.06**	.08	.05	.09	.05
T_2 TA	**.50**	**.04**	**.50**	**.04**	**.50**	**.04**	**.49**	**.04**	**.50**	**.04**	**.50**	**.04**	**.38**	**.04**	**.50**	**.04**	**.50**	**.04**
T_4 TA	.09	.06	.09	.06	**.10**	**.05**	.08	.05	.07	.06	**.13**	**.06**	**.18**	**.05**	.12	.06	.09	.05
T_2 ER	-.13	.08	-.09	.06	.12	.06	.07	.07	-.02	.07	**-.15**	**.05**	.12	.08	**-.21**	**.07**	-.10	.05
T_4 ER	-.11	.07	-.03	.06	.01	.06	.03	.07	**-.16**	**.07**	**-.22**	**.05**	-.12	.06	**-.25**	**.07**	-.09	.05

Note. Ach = achievement; TA = test anxiety; ER = emotion regulation. **Bold** coefficients: $p < .05$.

Table S4
Model Fit for the Structural Equation Models With Lagged Emotion Test Anxiety and Emotion Regulation Paths

Model	χ^2	RMSEA	SRMR	CFI	TLI
Acceptance	1310.13	.033	.070	.950	.939
Positive Refocusing	1173.00	.028	.051	.966	.958
Refocus on Planning	1233.39	.030	.060	.957	.948
Positive Reappraisal	1294.54	.032	.059	.950	.940
Putting Into Perspective	1157.92	.028	.052	.965	.957
Rumination	1233.32	.030	.055	.961	.952
Catastrophizing	1240.86	.030	.054	.959	.950
Self-Blame	1268.69	.031	.057	.958	.949
Other-Blame	1167.60	.028	.055	.964	.956

Note. χ^2 tests for all models were $p < .001$ (778 degrees of freedom)

Table S5

Structural Equation Models with Lagged Test Anxiety and Emotion Regulation Paths: Standardized Coefficients

	Acceptance model		Positive refocusing model		Refocus on planning model		Positive reappraisal model		Putting into perspective model		Rumination model		Catstroph-izing model		Self-blame model		Other-blame model	
	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*
Autoregressive effects																		
T_1 Ach → T_3 Ach	**.47**	**.04**	**.48**	**.04**	**.47**	**.04**	**.47**	**.04**	**.48**	**.04**	**.46**	**.04**	**.48**	**.04**	**.48**	**.04**	**.48**	**.04**
T_1 Ach → T_5 Ach	**.20**	**.05**	**.20**	**.05**	**.20**	**.05**	**.21**	**.05**	**.21**	**.05**	**.21**	**.05**	**.20**	**.05**	**.21**	**.05**	**.20**	**.05**
T_3 Ach → T_5 Ach	**.44**	**.05**	**.45**	**.05**	**.43**	**.05**	**.41**	**.06**	**.44**	**.05**	**.40**	**.06**	**.44**	**.05**	**.44**	**.05**	**.43**	**.05**
T_2 TA → T_4 Ach	**.77**	**.06**	**.74**	**.05**	**.76**	**.05**	**.76**	**.06**	**.74**	**.05**	**.75**	**.10**	**.70**	**.10**	**.74**	**.07**	**.75**	**.05**
T_2 ER → T_4 Ach	**.36**	**.07**	**.54**	**.05**	**.61**	**.05**	**.45**	**.07**	**.55**	**.05**	**.55**	**.09**	**.61**	**.08**	**.48**	**.04**	**.50**	**.05**
Achievement & Anxiety																		
T_1 Ach → T_2 TA	**-.09**	**.03**	**-.08**	**.03**	**-.10**	**.03**	**-.11**	**.03**	**-.08**	**.03**	**-.06**	**.03**	**-.07**	**.03**	**-.08**	**.03**	**-.09**	**.03**
T_1 Ach → T_4 TA	.02	.05	.01	.05	.02	.05	.02	.05	.02	.05	.02	.05	.02	.05	.02	.05	.01	.05
T_3 Ach → T_4 TA	**-.13**	**.05**	**-.11**	**.05**	**-.14**	**.05**	**-.14**	**.05**	**-.11**	**.05**	**-.10**	**.05**	**-.11**	**.05**	**-.12**	**.05**	**-.12**	**.05**
T_2 TA → T_3 Ach	**-.14**	**.05**	**-.13**	**.05**	**-.15**	**.05**	**-.20**	**.06**	**-.12**	**.04**	**-.24**	**.07**	**-.22**	**.08**	**-.11**	**.05**	**-.13**	**.04**
T_2 TA → T_5 Ach	.14	.09	.07	.09	.09	.09	.09	.10	.05	.09	.13	.13	.11	.14	.05	.10	.08	.08
T_4 TA → T_5 Ach	**-.19**	**.07**	**-.18**	**.07**	**-.21**	**.07**	**-.28**	**.09**	**-.17**	**.07**	**-.30**	**.10**	**-.28**	**.11**	**-.15**	**.08**	**-.18**	**.07**
Anxiety & Emotion regulation																		
T_2 TA → T_4 ER	**.35**	**.08**	.01	.07	-.10	.07	**-.25**	**.08**	.09	.07	.11	.11	.10	.12	**.26**	**.09**	.06	.08
T_2 ER → T_4 TA	-.03	.06	-.02	.05	.03	.05	.02	.06	-.01	.05	-.03	.09	.04	.10	.01	.06	-.09	.05
T_2 TA & T_2 ER	**.34**	**.06**	**-.13**	**.06**	**-.18**	**.06**	**-.36**	**.06**	**-.22**	**.06**	**.68**	**.04**	**.71**	**.04**	**.51**	**.05**	.09	.06
T_4 TA & T_4 ER	**.27**	**.09**	-.11	.08	**-.26**	**.10**	**-.52**	**.09**	**-.21**	**.08**	**.64**	**.06**	**.67**	**.06**	**.32**	**.08**	.16	.08
Achievement & Emotion regulation																		
T_1 Ach → T_2 ER	-.01	.03	.05	.04	.04	.04	.01	.04	.01	.04	**.08**	**.03**	-.05	.03	.04	.03	**-.11**	**.04**
T_1 Ach → T_4 ER	-.01	.06	-.01	.06	**.15**	**.06**	.05	.06	.05	.06	-.03	.06	.09	.05	.06	.05	.05	.06
T_3 Ach → T_4 ER	-.01	.04	.06	.05	.05	.05	.01	.05	.01	.05	**.12**	**.05**	-.08	.05	-.05	.05	**-.12**	**.04**
T_2 ER → T_3 Ach	-.03	.04	-.06	.02	.02	.04	-.07	.04	-.01	.03	**.17**	**.06**	.13	.07	-.04	.04	-.02	.02
T_2 ER → T_5 Ach	**-.13**	**.06**	.04	.06	-.01	.07	-.02	.07	-.02	.05	-.10	.10	-.06	.11	.04	.07	-.06	.06
T_4 ER → T_5 Ach	-.03	.04	-.09	.04	.02	.05	-.10	.06	-.09	.04	**.22**	**.08**	.15	.08	-.05	.05	-.01	.05

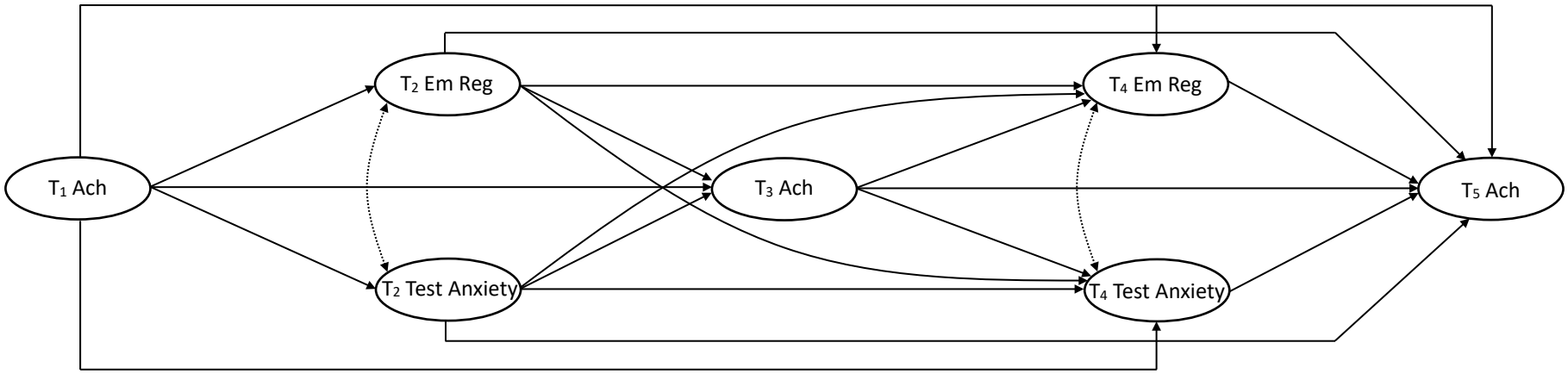
Note. Ach = achievement; TA = test anxiety; ER = emotion regulation. **Bold** coefficients: $p < .05$.

Table S6
Structural Equation Models with Lagged Test Anxiety and Emotion Regulation Paths: Standardized Path Coefficients for Gender

	Acceptance model		Positive refocusing model		Refocus on planning model		Positive reappraisal model		Putting into perspective model		Rumination model		Catstroph-izing model		Self-blame model		Other-blame model	
	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*	ß	*SE*
Effects of gender																		
T_1 Ach	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**	**.11**	**.04**
T_3 Ach	**.14**	**.05**	**.14**	**.05**	**.15**	**.05**	**.16**	**.05**	**.14**	**.05**	**.16**	**.05**	**.15**	**.05**	**.14**	**.05**	**.14**	**.05**
T_5 Ach	.09	.05	.08	.05	**.10**	**.05**	**.11**	**.05**	.08	.05	**.11**	**.05**	**.11**	**.05**	.08	.08	.09	.05
T_2 TA	**.50**	**.04**	**.50**	**.04**	**.50**	**.04**	**.49**	**.04**	**.50**	**.04**	**.49**	**.04**	**.49**	**.04**	**.50**	**.04**	**.50**	**.04**
T_4 TA	.08	.06	.10	.06	.08	.06	.08	.06	.09	.06	.11	.06	.11	.06	.09	.06	.09	.06
T_2 ER	.06	.05	**-.11**	**.05**	.03	.05	**.11**	**.05**	-.02	.05	**.24**	**.04**	**.25**	**.04**	.03	.05	-.03	.05
T_4 ER	-.07	.07	-.05	.06	-.03	.06	-.02	.07	**-.21**	**.06**	**-.02**	**.06**	-.04	.06	**-.18**	**.07**	-.07	.07

Note. Ach = achievement; TA = test anxiety; ER = emotion regulation. **Bold** coefficients: $p < .05$.

Figure S1
Hypothesized Model with Lagged Relations Between Test Anxiety and Emotion Regulation



Note. Solid lines represent structural paths and dotted lines correlations. Gender was included in the model as a covariate but omitted from Figure 1 to avoid over-cluttering.