Nicholson, L and Putwain, DW

A Cross-Lagged Panel Analysis of Fear Appeal Appraisal and Student Engagement

A Cross-Lagged Panel Analysis of Fear Appeal Appraisal and Student Engagement

Short title: *Fear Appeal Appraisal and Engagement*

Laura J. Nicholson*[1] and David W. Putwain[2]

[1] Faculty of Education, Edge Hill University, Ormskirk, Lancashire, UK

[2] School of Education, Liverpool John Moores University, Liverpool, UK

*Corresponding author information: Dr Laura Nicholson, Faculty of Education, Edge Hill University, Ormskirk, L39 4QP, UK. Email: Laura.nicholson@edgehill.ac.uk. Tel: 01695 657277

A Cross-Lagged Panel Analysis of Fear Appeal Appraisal and Student Engagement

Abstract

Background: Fear appeals are persuasive messages given by teachers to students about the importance of avoiding failure in an upcoming high-stakes test. The relationship between the way that students appraise fear appeals and engagement in lessons has not previously been tested using a robust methodological and analytical design that fully controls for concurrent relationships between the variables and stability over time.

Aim: The present study addressed these limitations using a cross-lagged panel design to probe reciprocal relationships between students' appraisal of fear appeals as a threat and as a challenge, with their engagement in class.

Sample: A total of 2,025 Year 10 and 11 students took part.

Method: Fear appeal appraisal and student engagement were measured at two time-points, four months apart.

Results: After controlling for unlagged and auto-lagged correlations, and gender and year group, the model fit the data well and six cross-lagged paths emerged as statistically significant. Threat appraisal and emotional engagement were reciprocally, negatively related. Threat appraisal also positively predicted emotional disaffection and challenge appraisal negatively predicted behavioural disaffection. In addition, prospective relations were revealed between specific components of student engagement.

Conclusions: Classroom teachers need to be aware of the possible consequences of making fear appeals and moderate this aspect of their practice accordingly. It would also be beneficial for educational interventions to focus on promoting challenge appraisals of fear appeals.

Keywords: Fear appeals, threat appraisal, challenge appraisal, student engagement, student disaffection

This study investigated possible reciprocal relationships between students' appraisal of fear appeals – which are persuasive messages given by teachers to students about the importance of avoiding failure in upcoming high-stakes tests – and their engagement in class. Previous research has demonstrated that appraising fear appeals as threatening is related to lower engagement, whereas challenge appraisals relate to higher engagement (e.g., Putwain *et al.*, 2016), however, these studies have either been cross-sectional in design or lacked the necessary controls in terms of accounting for concurrent relationships between variables and variable stability across time. It is important to fully examine the direction of these relations to ensure that educational efforts and interventions designed to facilitate success and reduce failure, are focusing on the antecedent rather than the outcome. Further, relationships between fear appeal appraisal and disaffection, and between the emotional, and behavioural, dimensions of engagement and disaffection have not previously been fully examined. The aim of the present study was to address these limitations using a cross-lagged panel design, in which we temporally separated variables treated as predictors and outcomes and controlled for auto-lagged and unlagged correlations, as well as influential demographic variables. This allowed a rigorous investigation into the direction of effects.

Appraisal of Fear Appeals

In the field of education, 'fear appeals' denote messages that teachers use to elicit fear of failure in students in an attempt to encourage and motivate them to work hard and do well in an upcoming high-stakes examination (Putwain & Best, 2011). In particular, the messages emphasise the importance and value of the examination and its outcome by prompting fear about the consequences of failing and highlighting actions which may increase and/or decrease the threat of failure (Putwain & Symes, 2014). For example, a teacher may inform students that if they do not revise for their examination they will fail and if they fail, they may not meet the entry criteria for further education or employment.

Students respond differently to the messages within fear appeals depending on the extent to which they appraise them as threatening or challenging. Students who appraise them as threatening are focused on concerns about harm or loss of self-worth and are worried about not being able to fulfil aspirations and demonstrating low ability to themselves and significant others, e.g., their parents and peers, which results in fear and anxiety. Students who find the messages challenging are focused on growth and mastery, and experience hope, excitement and eagerness (Blascovich, 2008; Blascovich & Mendes, 2000; Folkman & Lazarus, 1985; Lazarus, 2006). It follows that threat and challenge appraisals display negative and positive empirical relationships with educationally-relevant cognitions, emotions and behaviours, respectively (McGregor & Elliot, 2002; Putwain & Remedios, 2014; Putwain, Remedios, & Symes, 2015; Putwain & Symes, 2011a, 2011b; Putwain, Symes, & Wilkinson, 2017; Putwain *et al.*, 2016, 2017; Skinner & Brewer, 2002). Further, the same pattern of associations extends to examination performance (Putwain & Remedios, 2014; Putwain & Symes, 2011a; Putwain, Symes, & Wilkinson, 2017).

Student Engagement

Student engagement is multidimensional consisting of emotional (e.g., enjoyment of learning), cognitive (e.g., psychological investment in learning) and behavioural (e.g., involvement in academic tasks) elements (Fredricks, Blumenfeld, & Paris, 2004). In their conceptualization, Skinner, Kindermann and Furrer (2009) embedded cognitive engagement within the behavioural and emotional components. Broadly speaking, they proposed that student engagement comprises mental effort, on-task behaviour, class participation and energized emotion, as well as indicators of disaffection including not only a lack of engagement but also mental withdrawal, ritualistic/pressured participation and alienated/enervated emotion. Specifically, their model contained four components which they argued represented the core indicators of engagement in the classroom, namely behavioural

engagement (e.g., effort, participation), emotional engagement (e.g., interest, enjoyment), behavioural disaffection (e.g., passivity, lack of attention) and emotional disaffection (e.g., boredom, frustration). Behavioural and emotional engagement positively inter-relate but negatively correlate with the positively inter-related components of behavioural and emotional disaffection (e.g., Skinner, Furrer, Marchand, & Kindermann, 2008). Empirical work has confirmed the existence of these components and the importance of measuring each one for a full assessment of student engagement (e.g., Nicholson *et al.*, 2019; Skinner *et al.*, 2008, 2009).

Relationships between Fear Appeal Appraisals and Student Engagement

Student engagement is a malleable construct that is responsive to contextual factors, such as the classroom environment and teacher behaviour (Fredricks *et al.*, 2004; Skinner, 2016). Importantly, it is the key mediator linking contextual factors to educational outcomes of interest, most notably, academic achievement (Fredricks *et al.*, 2004; Reschly & Christenson, 2012; Skinner, 2016). One such contextual factor is the occurrence of fear appeals, and how students appraise them. Indeed, the impact of fear appeals on engagement is dependent on how they are appraised by the student (Putwain, Symes, & McCaldin, 2019; Putwain, Symes, & Wilkinson, 2017; Putwain *et al.*, 2016, 2017), supporting the general proposition that subjective experiences of the classroom mediate the effects of objective features on engagement (Skinner, 2016). Specifically, threat and challenge appraisals are related to lower and higher (behavioural and emotional) engagement, respectively. Moreover, behavioural engagement mediates the effects of appraisals on examination grade, whereby a threat appraisal predicts a lower grade through lower engagement, and a challenge appraisal predicts a higher grade through higher engagement (Putwain, Symes, & Wilkinson, 2017). It is apparent, however, that the aforementioned studies measured appraisals and engagement cross-sectionally, which precludes inferences about the directionality of effects. Furthermore,

a lack of longitudinal research has also prevented robust analyses that can control for concurrent relationships between variables at each time-point and the stability of the same variables across time.

Only one study to date has investigated the relationship between fear appeal appraisal and disaffection. Using a person-centred approach, Nicholson *et al.* (2019) found that students reporting low threat simultaneously with high challenge displayed low levels of both behavioural and emotional disaffection. Notably, the relationship between high threat appraisal and high emotional disaffection emerged even when there was also a high challenge appraisal. Appraisals and disaffection were again measured concurrently, however, preventing directionality inferences. It is conceivable that student engagement may also impact on fear appeal appraisals. More engaged students may be more likely to feel challenged by the messages contained in fear appeals and disaffected students may be more inclined to appraise them as threatening. Indeed, behavioural engagement has been found to positively predict a prospective measure of challenge appraisal and emotional engagement negatively predicted threat appraisal (Putwain, Symes, & Wilkinson, 2017; Putwain *et al.*, 2016). Although the variables were temporally separated in these studies, all variables were not measured at each time-point.

The Present Study

The extant literature has failed to provide a rigorous investigation of either of the directional hypotheses (i.e., that fear appeal appraisal predicts student engagement and vice versa). To address these limitations, we measured threat and challenge appraisals of fear appeals and the four components of student engagement at two time-points in a sample of secondary school students. This cross-lagged panel design fully controlled for (i) concurrent relationships between all variables at each time-point (unlagged correlations) and (ii) stability of the same variables across time (auto-lagged paths), and tests a reciprocal model in which

fear appeal appraisals predict student engagement and engagement predicts appraisals (cross-lagged paths). The model is illustrated in Figure 1.

---- Figure 1 about here ----

Fear appeal appraisal and student engagement were measured in relation to a forthcoming high-stakes mathematics examination. The subject of mathematics was chosen because students are required to pass this examination in order to gain entry into post-compulsory education or employment, and are therefore likely to perceive it as high-stakes. Mathematics has been found to be perceived by students as more important, useful and interesting than other school subjects (Wolters & Pintrich, 1998), although engagement in mathematics classes is relatively low (Pöysä *et al.*, 2018). Due to demonstrated gender and age differences in engagement in mathematics classes and appraisal of fear appeals made in relation to mathematics examinations (Putwain, Symes & Wilkinson, 2017; Putwain *et al.*, 2017), gender and year group were included as covariates in the model.

Reciprocal relationships between fear appeal appraisals and student engagement were expected. Specifically, the following hypotheses were tested:

Hypothesis 1 (H_1): Threat appraisal will be negatively related, and challenge appraisal positively related, in a reciprocal fashion, to behavioural and emotional engagement.

Hypothesis 2 (H_2): Threat appraisal will be positively related, and challenge appraisal negatively related, in a reciprocal fashion, to behavioural and emotional disaffection.

The design also allowed a full investigation into the internal dynamics of student engagement over time. This is useful to clarify the role of engagement in the development of student academic resilience and success (Skinner *et al.*, 2008). Previous research has demonstrated that the emotional components of engagement and disaffection have a stronger and more consistent effect on the behavioural components than vice versa (Skinner *et al.*, 2008). This provides a useful starting point; however, Skinner *et al.* did not utilise a cross-

lagged panel design and concurrent relations were not controlled. They also did not test for relations between the two behavioural (i.e., engagement and disaffection) or two emotional components over time. The current research will therefore significantly add to the existing body of knowledge. As the relations between the engagement components were not the focus of the study, we do not offer specific hypotheses.

Method

Participants

A total of 2,025 students were recruited from 108 classes within six secondary schools in England. All students were in the final two years of compulsory education and working towards their secondary school exit examinations (General Certificate of Secondary Education; GCSE) (1,077 Year 10 students, 947 Year 11 students, 1 unknown). There was a fairly even split of males and females in the sample (1,018 males, 50.3%; 973 females, 48.0%; 34 unknown, 1.7%) and participants reported a mean age of 14.61 years (SD = 0.62). Most of the students described their ethnicity as White (80.9%). The remainder stated that they were Asian (10.8%), Black (1.9%), Other (2.7%), and mixed heritage (3.1%; 0.6% unknown). Two hundred and thirteen students qualified for free school meals due to low parental income (10.5% of the sample; 1.1% undisclosed). A comparison between the characteristics of our sample and the average secondary school student population in England during the school year when data were collected indicates that our sample was fairly representative. Compared to the national average, we had a lower proportion of minority ethnic students (approximately 19.1% versus 26.6%) and a lower number of students eligible for free school meals (approximately 10.5% versus 14.6%; Department for Education, 2015).

Measures

Fear appeal appraisals. Fear appeal appraisals were measured using items from the ‘Revised Teachers Use of Fear Appeals Questionnaire’ (Putwain & Symes, 2014).

Participants responded on a five-point scale (1 = never, 3 = sometimes, 5 = most of the time) with a higher score representing a higher threat/challenge appraisal. Three items assessed the appraisal of fear appeals as a threat (e.g., 'Do you feel worried when your teacher tells you that unless you work hard you will fail your maths GCSE?') and three items assessed a challenge appraisal (e.g., 'Does it make you want to pass GCSE maths when your teacher tells you that unless you work hard you will fail?'). Previous research has established the reliability and validity of data collected using these scales (Putwain & Symes, 2014). In the present study, internal consistency was good at both time-points (McDonald's $\omega \geq .76$; Table 1).

Student engagement. The four components of student engagement were measured using items from the 'Engagement versus Disaffection with Learning Questionnaire' (Skinner *et al.*, 2009), adapted to be specific to GCSE mathematics. Three items assessed each subcomponent; behavioural engagement (e.g., 'I participate in the activities and tasks in my GCSE maths class'), emotional engagement (e.g., 'I enjoy learning things in GCSE maths'), behavioural disaffection (e.g., 'I don't try very hard in GCSE maths') and emotional disaffection (e.g., 'When I'm doing GCSE maths work in class, I feel bored'). Participants responded on a five-point scale (1 = strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree) and a higher score represented higher engagement/disaffection. The reliability and validity of data collected using these scales have been previously demonstrated (Skinner *et al.*, 2008, 2009). Internal consistencies in the present study were acceptable (McDonald's $\omega \geq .67$; Table 1).

Procedure

Prior to data collection, the project was approved by a Faculty Research Ethics Committee and consent obtained from the school principals, parents and individual students. Fear appeal appraisals and student engagement were measured at two time-points; at the

beginning of the academic year (T_1) and approximately four months later (T_2). Tutors (not mathematics teachers) read from a standardised sheet which included information about the study and ethical issues. The students were told that their participation was entirely voluntary, that the questions were not part of a test and that their answers would remain confidential to the researchers. Students then received a questionnaire pack which began with an information sheet re-iterating the information received from the tutor. If willing to participate, they completed a consent form, demographics page and the main questionnaires.

Analytical Strategy

Except where stated, analyses were conducted using M*plus* 8.0 (Muthén & Muthén, 2017). Following assessment of missing data and descriptive statistics, we used confirmatory factor analysis (CFA) to examine a measurement model consisting of six latent variables, with three indicators each: threat appraisal and challenge appraisal, and the four student engagement components, at the two time-points (12 latent variables in total). We tested the temporal measurement invariance of the six variables (in the same model) to ensure that the underlying factors reflected the same construct and functioned in the same way over time. It is crucial to establish this prior to testing relationships between latent factor means across time (Marsh, Nagengast, & Morin, 2013). Bivariate correlations were then estimated between the latent variables and with the covariates, gender (0 = male, 1 = female) and year group (0 = Year 10, 1 = Year 11). Finally, a structural equation model (SEM) was built to test the cross-lagged panel model (Figure 1) in which gender and year group were added as covariates.

To evaluate model fit for the measurement models and SEM, we used the comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR). Values of >.95 for CFI and TLI, and <.08 and <.06, for SRMR and RMSEA, respectively, indicate a good fit between the

hypothesised model and observed data (Hu & Bentler, 1999). These values may be overly stringent, however, and should not be treated as strict cut-off scores (e.g., Heene, Hilbert, Draxler, Ziegler, & Bühner, 2011; Marsh, Hau, & Wen, 2004), although it is important to establish general guidelines a priori (Pendergast, von der Embse, Kilgus, & Eklund, 2017). For interpretation of the path coefficients resulting from the SEM, βs > .05 were considered small, βs > .10 moderate, and βs > .25 large (Keith, 2013).

Results

Descriptive Statistics and Resulting Analytical Decisions

There was a total of 23.75% of missing data in the dataset. Due to whole classes not completing the measures at one of the time-points, using SPSS Version 24, we tested whether data were missing completely at random at the two waves separately and then overall for each class (otherwise, the results would have been simply a test of class effects). Little's (1988) test suggested that missing data were observed completely at random at T_1, $\chi^2(77) = 93.32$, $p = .10$, T_2, $\chi^2(60) = 69.36$, $p = .19$, and for the separate classes (all ps > .05). Missing data were subsequently handled in M*plus* using full information maximum likelihood.

Descriptive statistics are presented in Table 1. Challenge appraisal was higher than threat appraisal and among the engagement components, behavioural engagement was highest and behavioural disaffection lowest. In general, data were normally distributed, however, behavioural engagement showed slightly negatively skewed leptokurtic distributions at both waves. Maximum likelihood with robustness to non-normality and non-independence of observations (MLR) was used to estimate the subsequent models. The intraclass correlation coefficient (ICC_1) revealed that a moderate proportion of the variance in the four engagement components at both time-points was attributable to the class level (5% - 10%). Proportions of variance occurring at the class level were greater for fear appeal appraisal at both time-points (14% - 22%). To account for this clustering of participants

within classes, the M*plus* 'cluster' and 'type = complex' commands were used. Standardised factor loadings reported from the measurement model were all statistically significant and most ranged from moderate to high (λ = .59 - .86), except for one lower but satisfactory loading (λ = .48; Kline, 2011; Table 1).

----Table 1 about here----

Measurement Models

In order to obtain accurate estimates of relations among constructs, residuals for parallel items at each wave were correlated (Marsh & Hau, 1996). The measurement model showed an excellent fit to the data, $\chi^2(502) = 1216.22$, $p < .001$, RMSEA = .027, SRMR = .038, CFI = .967, and TLI = .959. The model fit and the finding that all factor loadings were statistically significant at both waves also establishes configural invariance of the six constructs over time and indicates that the model provides a sound measurement base upon which to conduct the cross-lagged analyses (Collie, Martin, Malmberg, Hall, & Ginns, 2015; Martin, Ginns, Brackett, Malmberg, & Hall, 2013). We subsequently tested a series of nested models with increasingly restrictive invariance tests, in which a ΔCFI/ ΔTLI < .01 and ΔRMSEA < .015 indicated strong invariance over time (Chen 2007; Cheung & Rensvold, 2002). Results of these tests are shown in Table 2. Metric (factor loadings constrained to be equal), scalar (item intercepts constrained to be equal) and residual (residuals constrained to be equal) invariance were established.

----Table 2 about here ----

Latent Bivariate Correlations

Latent bivariate correlations are reported in Table 3. Intercorrelations between the same variables over time were all strong and statistically significant, indicating that the variables were highly stable. The cross-lagged correlations between appraisals and engagement components followed the expected directions at both time-points. Threat

appraisal was negatively correlated with emotional engagement and positively correlated with emotional disaffection. Challenge appraisal was positively correlated with behavioural and emotional engagement and negatively correlated with behavioural disaffection.

----Table 3 about here ----

Structural Equation Modelling

The sample size for the full model was N = 1,991. The model fit the data well, $\chi^2(550) = 1336.47$, p <.001, RMSEA = .027, SRMR = .037, CFI = .964, and TLI = .954. Table 4 displays the unlagged correlation coefficients (after model estimation, as opposed to the latent bivariate correlations reported previously), which displayed a similar pattern at the two time-points. Threat and challenge appraisal were strongly and positively related. All but one of the correlations between the engagement components were statistically significant, in the expected direction and ranged from moderate to very strong in magnitude. Threat appraisal was not related to behavioural or emotional engagement. Threat appraisal correlated weakly with behavioural disaffection (positively) and slightly more strongly with emotional disaffection (positively). Challenge appraisal was moderately positively correlated with behavioural and emotional engagement, weakly negatively related to emotional disaffection, and at T_1 only, moderately negatively correlated with behavioural disaffection.

----Table 4 about here ----

Full results for the auto- and cross-lagged analyses are reported in Table 5, and statistically significant paths illustrated in Figure 2. All auto-lagged paths were statistically significant and all but one revealed very strong relationships between variables over time (ranging from β = .38 for emotional disaffection to β = .75 for threat appraisal). The auto-lagged path coefficient for behavioural engagement was moderate to large (β = .21).

----Table 5 about here ----

---- Figure 2 about here ----

Six cross-lagged paths were statistically significant. This included three paths from T_1 fear appeal appraisals to T_2 student engagement and one from T_1 engagement to T_2 appraisals. Threat appraisal predicted emotional engagement ($\beta = -.09$, $p = .04$) and disaffection ($\beta = .13$, $p = .02$), challenge appraisal predicted behavioural disaffection ($\beta = -.12$, $p = .02$) and emotional engagement predicted threat appraisal ($\beta = -.20$, $p = .03$). Standardised beta coefficients can be interpreted in terms of the number of standard deviations that the outcome will change as a result of one standard deviation change in the predictor. All variables were measured on a five-point scale. For every increase of 1.15 scale points on threat appraisal, responses for emotional engagement decreased by 0.09 points and responses for emotional disaffection increased by 0.12 points. When challenge appraisal increased by 1.10 scale points, behavioural disaffection decreased by 0.10 points, and when emotional engagement increased by 1.01 scale points, threat appraisal decreased by 0.23 points.

These findings suggest that the relationship between threat appraisal and emotional engagement is reciprocal. To confirm this, we constrained the two cross-lagged paths to be equal and compared the model fit (Martin *et al.*, 2013). There was no reduction in the goodness of fit indices, $\chi^2(551) = 1338.58$, $p < .001$, RMSEA = .027, SRMR = .037, CFI = .964, and TLI = .954 (ΔRMSEA = <.001, ΔCFL < .001, ΔTLI = <.001) and no statistically significant difference between the two models, $\Delta\chi^2$ (2) = 2.11, $p > .05$, which verifies the existence of a reciprocal relationship. Although the magnitude of the standardised path coefficient from T_1 emotional engagement to T_2 threat appraisal ($\beta = -.20$) was larger than from T_1 threat appraisal to T_2 emotional engagement ($\beta = -.09$), the difference was not statistically significant, $t(3978) = 0.98$, $p = .33$ (Cohen, Cohen, West & Aiken, 2003; Soper, 2019).

Two of the cross-lagged paths between the engagement components were also statistically significant. Behavioural disaffection predicted behavioural engagement ($\beta = -.41$,

p = .006) and emotional engagement predicted emotional disaffection (β = -.27, p = .02). Thus, for every increase of 0.86 scale points on behavioural disaffection, behavioural engagement decreased by 0.27 points, and for every increase of 1.01 scale points on emotional engagement, emotional disaffection decreased by 0.26 points. In general, the six statistically significant cross-lagged paths ranged from moderate (between fear appeal appraisals and engagement) to large (between engagement components) effect sizes.

Regarding the covariates, female students reported higher threat appraisal (β = .19, p < .001) and emotional disaffection (β = .08, p < .01) and lower emotional engagement (β = -.11, p < .001) at T_1, and lower behavioural disaffection (β = -.12, p < .001) at T_2. Year 11 students reported higher behavioural (β = .15, p < .001) and emotional (β = .09, p < .05) engagement and lower behavioural (β = -.09, p < .01) and emotional (β = -.09, p < .01) disaffection at T_2. All other relations with gender and year group were not statistically significant (ps >.05).

Discussion

A cross-lagged panel design was used to investigate relationships between students' fear appeal appraisals and engagement. The rigorous design was superior to previous research by measuring appraisals and engagement longitudinally and allowing a clear temporal separation between variables treated as predictors and outcomes, to establish the direction of effects. Both autoregressive and concurrent relations were controlled, which meant that statistically significant cross-lagged paths represented effects over and above prior measures of the outcome variables and concurrent relations between the predictor and outcome variable. Controlling for auto-lagged paths (stability effects) is considered the gold standard in longitudinal designs (Adachi & Willoughby, 2015). When stability is high, there is less variance left over to be accounted for by additional predictors and therefore effect sizes tend to be smaller than in cross-sectional research. Even small effect sizes can be meaningful

(Adachi & Willoughby, 2015). The finding that all variables were highly stable over time therefore highlights the importance of the six established cross-lagged paths, all of which represented moderate-large effect sizes. The effects of gender, year group and the nesting of students within classes were also controlled for in the analyses.

The reciprocity aspect of the hypotheses was not widely supported. Only one reciprocal relationship was revealed out of a possible eight. Students who appraised fear appeals at a high level of threat at T_1 reported low emotional engagement at T_2, which supports and strengthens previous cross-sectional findings (Putwain, Symes, & Wilkinson, 2017; Putwain *et al.*, 2016, 2017, 2019). High emotional engagement at T_1 also predicted low threat appraisal at T_2, which again confirms past research that relied on a weaker design for probing this relationship (Putwain *et al.*, 2016). This reciprocal finding extends the research on feed-forward and feedback effects within motivational dynamics, exemplifying a 'vicious cycle' whereby motivationally 'poor' students become 'poorer' (Skinner & Pitzer, 2012, p. 31). Novel findings were obtained for the relationships between appraisals and disaffection. A high threat appraisal at T_1 predicted high emotional disaffection at T_2, and a high challenge appraisal at T_1 predicted low behavioural disaffection at T_2. All of the findings were in the expected direction and partially support the two hypotheses. They provide robust evidence that the way that students appraise fear appeals at the beginning of the academic year relate to their self-reported engagement in class four months later. Also, initially emotionally engaged students were less likely to appraise fear appeals in a threatening way later in the term. There was more support for the proposition that appraisals predict student engagement than vice versa. Further, a pattern seems to be emerging suggesting that threat and challenge appraisals are more closely related to the emotional and behavioural aspects of engagement/disaffection, respectively. The former relationship appears more consistent, reciprocal and is even present,

cross-sectionally at least, when there is a simultaneously high challenge appraisal (Nicholson *et al.*, 2019).

Although the bivariate concurrent correlations between threat appraisal and emotional engagement were statistically significant, the unlagged correlations, which controlled for the auto- and cross-lagged effects, were not. This suggests that the relationship is not concurrent, that is, high threat appraisals do not coincide with low emotional engagement at the same point in time. This extends previous cross-sectional findings which have obscured the true longitudinal nature of the relationship. Reciprocal effects therefore accumulate over time; it may be that the predictor is associated with levels of change in the outcome variable (Adachi & Willoughby, 2015). Indeed, it is conceivable that substantial changes in how students appraise teacher messages and engage in class take place from when students start the academic year to four months later when they are fully focused on their GCSE course (Year 10) and/or when the reality of the upcoming examinations is realised (Year 11). Fear appeals, and the resulting threat appraisal for some students, may act like a self-reflective prompt over time, becoming imprinted in students' heads as the teacher messages are repeated across the school year. Therefore, it may take time for threat appraisals, and also emotional engagement in class, to embed and become established before having the observed effects on each other during this period. Further, the exact timing of when the teachers gave the fear appeal messages, and the extent to which the frequency of fear appeals fluctuates across the school year, is unknown; this could be a contributing factor to these novel findings. Experimental research is necessary to further delineate the purely longitudinal, reciprocal relationship.

The additional investigation into the internal dynamics of engagement revealed that high behavioural disaffection predicted low behavioural engagement, and high emotional engagement predicted low emotional disaffection. These novel findings contribute to our knowledge by demonstrating that for behaviour, indicators of disaffection precede changes in

engagement, but for the emotional aspect, markers of engagement lead to changes in disaffection. Our measures captured relatively mild manifestations of disaffection; we did not measure more active and overt displays of behavioural disaffection such as disruption and truancy, nor did we measure the full array of disaffected emotions including frustration or anxiety in class (our measure assessed boredom and feeling discouraged in class). It would be fruitful for further research to build on our findings by tapping these alternative aspects.

Limitations and Future Directions

All data were self-reported which is defensible given the focus on psychological factors (Martin *et al.*, 2013). Information about private experiences and self-perceptions cannot be accessed via any other means; the student perspective is critical (Reschly & Christenson, 2012). However, future researchers may consider collecting alternative measures using repetitive real-time digital recordings in class. The time lag of four months between waves of data collection may be usefully increased to observe if the demonstrated effects are more or less pronounced or if different effects are detected over a longer time span. Also, the focus on GCSE mathematics limits the generalisability of findings to other school subjects.

It is important to note that our statistical model cannot determine causation (Selig & Little, 2012). Variables were not manipulated nor were participants randomly assigned to groups. However, a cross-lagged panel study does imply directionality and findings can be used to build a causal argument when used in conjunction with theory and other empirical results (Selig & Little, 2012). Experimental research to test the established prospective relationships is now required to confirm the findings, for instance, fear appeal appraisals can be manipulated with the use of vignettes (see Putwain & Symes, 2014, 2016) and effects on student engagement examined. An alternative avenue for further work would be to identify moderators of the examined relationships, such as socio-economic status which is known to

affect academic self-perceptions and outcomes. Finally, a natural extension to this work would be to obtain eventual GCSE mathematics grades to observe the complete process from fear appeal appraisals to examination performance.

Educational Implications

Findings suggest that it is highly favourable for students to appraise fear appeals as challenging, and not threatening. To promote high challenge and low threat appraisals, it is necessary to target their antecedents. When students value their performance on their mathematics test, they will appraise fear appeals as threatening if they do not expect to succeed and challenging if they expect success (Putwain & Symes, 2014). An intervention based on increasing students' self-efficacy or competency beliefs could prove effective, for instance, through increasing mastery or vicarious experiences (Bartsch, Case, & Meerman, 2012; Luzzo, Hasper, Albert, Bibby, & Martinelli, 1999). Efforts directed to increasing emotional engagement at the start of the year should also prove effective for reducing high threat appraisals, as well as for reducing subsequent emotional disaffection. Many facilitators of engagement have been identified, including self-perceptions (e.g., sense of belonging) and objective classroom factors (e.g., the teacher's warmth), all of which could be targeted (Skinner, 2016). Although subjective responses to the classroom (e.g., fear appeal appraisals) are stronger predictors of student engagement than actual, objective conditions (e.g., occurrence of fear appeals), the former are based on the latter and so interventions should focus on these (Skinner, 2016). It follows that teachers need to be educated about the possible negative consequences of threat appraisals and encouraged to either limit their delivery of such messages to whole classes which would undoubtedly include students who would appraise them as threatening, or to use them with individual students with whom they are sure would appraise them in a challenging way. This assumes, however, that teachers are able to accurately judge students' motivational states (Putwain *et al.*, 2017).

Conclusion

Relationships between fear appeal appraisals and student engagement were investigated using a cross-lagged panel design in which all variables were measured at two time-points, four months apart. A reciprocal negative relationship between threat appraisals and emotional engagement was revealed. High threat appraisals also predicted high emotional disaffection, whereas high challenge appraisals predicted low behavioural disaffection. In addition, emotional engagement negatively predicted emotional disaffection and behavioural disaffection negatively predicted behavioural engagement. These findings extend those of previous research employing a less rigorous design and contribute significantly to our understanding of how psychological responses to messages received in the classroom relate to subsequent emotions and behaviour in lessons. The research also adds to the existing body of knowledge concerning the internal dynamics of engagement.

Acknowledgements

The authors would like to thank Ben Porter, Ghada Nakhla, Anthony Liversidge and Monika Reece for their help with data collection, and Hannah Wilkinson for her support with data entry and management.

References

Adachi, P., & Willoughby, T. (2015). Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: Implications for psychological science. *European Journal of Developmental Psychology, 12*(1), 116-128. doi:10.1080/17405629.2014.963549

Bartsch, R. A., Case, K. A., & Meerman, H. (2012). Increasing academic self-efficacy in statistics with a live vicarious experience presentation. *Teaching of Psychology, 39*(2), 133-136. doi:10.1177/0098628312437699

Blascovich, J. (2008). Challenge and threat. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 431-446). New York: Erlbaum.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 59-82). Cambridge: Cambridge University Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. doi:10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233-255. doi:10.1207/S15328007SEM0902_5

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3. ed. ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Collie, R. J., Martin, A. J., Malmberg, L., Hall, J., & Ginns, P. (2015). Academic buoyancy, student's achievement, and the linking role of control: A cross-lagged analysis of high school students. *British Journal of Educational Psychology, 85*(1), 113-130. doi:10.1111/bjep.12066

Department for Education (2015). *Schools, pupils and their characteristics: January 2015.* London, UK: HMSO

Folkman, S., & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, *48*(1), 150-170. doi:10.1037/0022-3514.48.1.150

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109. doi:10.3102/00346543074001059

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. doi:10.1037/a0024917

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. doi:10.1080/10705519909540118

Keith, T. Z. (2013). *Multiple regression and beyond*. Harlow, England: Pearson Education.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Lazarus, R. S. (2006). *Stress and emotion*. New York, NY: Springer.

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. doi:10.1080/01621459.1988.10478722

Luzzo, D. A., Hasper, P., Albert, K. A., Bibby, M. A., & Martinelli, E. A. (1999). Effects of self-efficacy-enhancing interventions on the math/science self-efficacy and career interests, goals, and actions of career undecided college students. *Journal of Counseling Psychology, 46*(2), 233-243. doi:10.1037/0022-0167.46.2.233

Marsh, H. W., & Hau, K. (1996). Assessing goodness of fit. *The Journal of Experimental Education, 64*(4), 364-390. doi:10.1080/00220973.1996.10806604

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in

overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320-341. doi:10.1207/s15328007sem1103_2

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology, 49*(6), 1194-1218. doi:10.1037/a0026913

Martin, A. J., Ginns, P., Brackett, M. A., Malmberg, L., & Hall, J. (2013). Academic buoyancy and psychological risk: Exploring reciprocal relationships. *Learning and Individual Differences, 27*, 128-133. doi:10.1016/j.lindif.2013.06.006

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement-relevant processes prior to task engagement. *Journal of Educational Psychology, 94*(2), 381-395. doi:10.1037//0022-0663.94.2.381

Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical analysis with latent variables: User's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Nicholson, L. J., Putwain, D. W., Nakhla, G., Porter, B., Liversidge, A., & Reece, M. (2019). A person-centered approach to students' evaluations of perceived fear appeals and their association with engagement. *The Journal of Experimental Education, 87*(1), 139-160. doi:10.1080/00220973.2018.1448745

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of School Psychology, 60*, 65-82. doi:10.1016/j.jsp.2016.11.002

Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M., Poikkeus, A., & Nurmi, J. (2018). Variation in situation-specific engagement among lower secondary school students. *Learning and Instruction, 53*, 64-73. doi:10.1016/j.learninstruc.2017.07.007

Putwain, D. W., & Best, N. (2011). Fear appeals in the primary classroom: Effects on test anxiety and test grade. *Learning and Individual Differences, 21*(5), 580-584. doi:10.1016/j.lindif.2011.07.007

Putwain, D. W., Nakhla, G., Liversidge, A., Nicholson, L. J., Porter, B., & Reece, M. (2017). Teachers use of fear appeals prior to a high-stakes examination: Is frequency linked to perceived student engagement and how do students respond? *Teaching and Teacher Education, 61*, 73-83. doi:10.1016/j.tate.2016.10.003

Putwain, D. W., Nicholson, L. J., Nakhla, G., Reece, M., Porter, B., & Liversidge, A. (2016). Fear appeals prior to a high-stakes examination can have a positive or negative impact on engagement depending on how the message is appraised. *Contemporary Educational Psychology, 44-45*, 21-31. doi:10.1016/j.cedpsych.2015.12.001

Putwain, D., & Remedios, R. (2014). The scare tactic: Do fear appeals predict motivation and exam scores? *School Psychology Quarterly, 29*(4), 503-516. doi:10.1037/spq0000048

Putwain, D. W., Remedios, R., & Symes, W. (2015). Experiencing fear appeals as a challenge or a threat influences attainment value and academic self-efficacy. *Learning and Instruction, 40*, 21-28. doi:10.1016/j.learninstruc.2015.07.007

Putwain, D., & Symes, W. (2011a). Perceived fear appeals and examination performance: Facilitating or debilitating outcomes? *Learning and Individual Differences, 21*, 227-232. doi:10.1016/j.lindif.2010.11.022

Putwain, D., & Symes, W. (2011b). Teachers' use of fear appeals in the mathematics classroom: Worrying or motivating students? *British Journal of Educational Psychology, 81*, 456-474. doi:10.1348/2044-8279.002005

Putwain, D. W., & Symes, W. (2014). The perceived value of maths and academic self-efficacy in the appraisal of fear appeals used prior to a high-stakes test as threatening or

challenging. *Social Psychology of Education, 17*, 229-248. doi:10.1007/s11218-014-9249-7

Putwain, D., & Symes, W. (2016). Expectancy of success, subjective task-value, and message frame in the appraisal of value-promoting messages made prior to a high-stakes examination. *Social Psychology of Education, 19*(2), 325-343. doi:10.1007/s11218-016-9337-y

Putwain, D. W., Symes, W., & McCaldin, T. (2019). Teacher use of loss-focused, utility value messages, prior to high-stakes examinations, and their appraisal by students. *Journal of Psychoeducational Assessment, 37*(2), 169-180. doi:10.1177/0734282917724905

Putwain, D. W., Symes, W., & Wilkinson, H. M. (2017). Fear appeals, engagement, and examination performance: The role of challenge and threat appraisals. *British Journal of Educational Psychology, 87*, 16-31. doi:10.1111/bjep.12132

Reschly, A., & Christenson, S. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly & C. Wiley (Eds.), *Handbook of research on student engagement* (pp. 3-19). Boston, MA: Springer US. doi:10.1007/978-1-4614-2018-7_1

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265-278). New York, NY: The Guilford Press.

Skinner, E. A. (2016). Engagement and disaffection as central to processes of motivational resilience and development. In K. R. Wentzel, & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 145-168). New York: Routledge.

Skinner, N., & Brewer, N. (2002). The dynamics of threat and challenge appraisals prior to stressful achievement events. *Journal of Personality and Social Psychology, 83*(3), 678-692. doi:10.1037//0022-3514.83.3.678

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765-781. doi:10.1037/a0012840

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493-525. doi:10.1177/0013164408323233

Skinner, E., & Pitzer, J. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly & C. Wiley (Eds.), *Handbook of research on student engagement* (pp. 21-44). Boston, MA: Springer US. doi:10.1007/978-1-4614-2018-7_2

Soper, D. S. (2019). Significance of the difference between two slopes calculator [software]. Retrieved from http://www.danielsoper.com/statcalc

Wolters, C. A., & Pintrich, P. R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science, 26*(1/2), 27-47. doi:1003035929216

Table 1

Descriptive Statistics for Fear Appeal Appraisals and Student Engagement (T_1 and T_2)

	Mean	SD	Skewness	Kurtosis	ICC_1	ω	Factor Loadings
T_1 Threat appraisal	2.70	1.15	0.19	-0.90	.22	.81	.73 - .79
T_1 Challenge appraisal	3.53	1.10	-0.57	-0.42	.14	.76	.69 - .74
T_1 Behavioural engagement	4.13	0.66	-1.11	2.75	.07	.78	.66 - .80
T_1 Emotional engagement	3.01	1.01	-0.04	-0.54	.09	.88	.79 - .86
T_1 Behavioural disaffection	2.24	0.86	0.44	-0.02	.09	.67	.48 - .82
T_1 Emotional disaffection	2.75	0.96	0.07	-0.46	.10	.78	.60 - .80
T_2 Threat appraisal	2.73	1.16	0.17	-0.93	.21	.84	.76 - .83
T_2 Challenge appraisal	3.44	1.10	-0.48	-0.55	.15	.77	.70 - .76
T_2 Behavioural engagement	4.07	0.66	-0.83	1.66	.07	.74	.64 - .77
T_2 Emotional engagement	3.00	0.99	-0.12	-0.52	.10	.88	.80 - .87
T_2 Behavioural disaffection	2.18	0.87	0.58	0.03	.05	.72	.59 - .82
T_2 Emotional disaffection	2.74	0.96	0.08	-0.55	.09	.78	.60 - .83

Table 2
Tests of Measurement Invariance

	χ^2	RMSEA	SRMR	CFI	TLI	ΔRMSEA	ΔCFL	ΔTLI
Configural invariance	1216.22 (502)	.027	.038	.967	.959			
Metric invariance	1229.05 (514)	.026	.038	.967	.960	-.001	<.001	+.001
Scalar invariance	1291.85 (532)	.027	.040	.965	.959	+.001	-.002	-.001
Residual invariance	1315.42 (550)	.026	.041	.965	.960	-.001	<.001	+.001

Note. All models statistically significant at $p < .001$.

Table 3

Latent Bivariate Correlations between Fear Appeal Appraisals and Student Engagement (T_1 and T_2), and Gender and Age

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. T_1 Threat	-	.66***	.00	-.08*	.08*	.24***	.72***	.48***	-.02	-.15**	.04	.21***	.19***	.05
2. T_1 Challenge		-	.32***	.23***	-.22***	-.11**	.46***	.66***	.21***	.10*	-.20***	-.06	.05	.02
3. T_1 BEH. ENG.			-	.56***	-.61***	-.43***	-.04	.19***	.53***	.30***	-.43***	-.32***	.02	.01
4. T_1 EMO. ENG.				-	-.31***	-.76***	-.11**	.10*	.35***	.63***	-.25***	-.57***	-.11***	.05
5. T_1 BEH. DIS.					-	.62***	.07	-.15***	-.49***	-.21***	.59***	.38***	-.06	.05
6. T_1 EMO. DIS.						-	.17***	-.04	-.34***	-.56***	.42***	.64***	.08**	-.03
7. T_2 Threat							-	.65***	.01	-.12**	.11**	.25***	.15***	-.02
8. T_2 Challenge								-	.32***	.17**	-.17***	-.08	.05	.02
9. T_2 BEH. ENG.									-	.54***	-.62***	-.46***	.06*	.14**
10. T_2 EMO. ENG.										-	-.26***	-.77***	-.10**	.12*
11. T_2 BEH. DIS.											-	.62***	-.13***	-.07
12. T_2 EMO. DIS.												-	.04	-.11*
13. Gender[a]													-	-
14. Year group[b]														-

Notes. **p* <.05. ***p* <.01. ****p* <.001. [a]Gender: 0 = Male, 1 = Female; [b]Year group: 0 = Year 10, 1 = Year 11; BEH. ENG. = Behavioural engagement; EMO. ENG. = Emotional engagement; BEH. DIS. = Behavioural disaffection; EMO. DIS. = Emotional disaffection.

Table 4

Unlagged Standardised Correlation Coefficients from the Cross-Lagged Panel Analyses

	1.	2.	3.	4.	5.	6.
1. Threat appraisal	-	.66***	-.00	-.07	.09*	.23***
2. Challenge appraisal	.62***	-	.32***	.24***	-.22***	-.11**
3. Behavioural engagement	.10	.32***	-	.57***	-.61***	-.43***
4. Emotional engagement	.03	.25***	.54***	-	-.32***	-.76***
5. Behavioural disaffection	.13*	-.08	-.45***	-.14	-	.63***
6. Emotional disaffection	.16**	-.12**	-.34***	-.64***	.59***	-

Notes. *p <.05. **p <.01. ***p <.001. Unlagged coefficients at T_1 and T_2 are presented above and below the diagonal, respectively.

Table 5

Auto- and Cross-Lagged Standardised Path Coefficients from the Cross-Lagged Panel Analyses

	T_2 Threat appraisal	T_2 Challenge appraisal	T_2 Behavioural engagement	T_2 Emotional engagement	T_2 Behavioural disaffection	T_2 Emotional disaffection
T_1 Threat appraisal	.75***	.07	-.06	-.09*	.09	.13*
T_1 Challenge appraisal	-.00	.62***	.05	.05	-.12*	-.04
T_1 Behavioural engagement	.05	-.00	.21*	-.03	-.10	.06
T_1 Emotional engagement	-.20*	-.04	.25	.47***	.06	-.27*
T_1 Behavioural disaffection	.11	-.04	-.41**	.07	.44**	.08
T_1 Emotional disaffection	-.21	.01	.21	-.23	.12	.38**

Notes. *p <.05. **p <.01. ***p <.001.

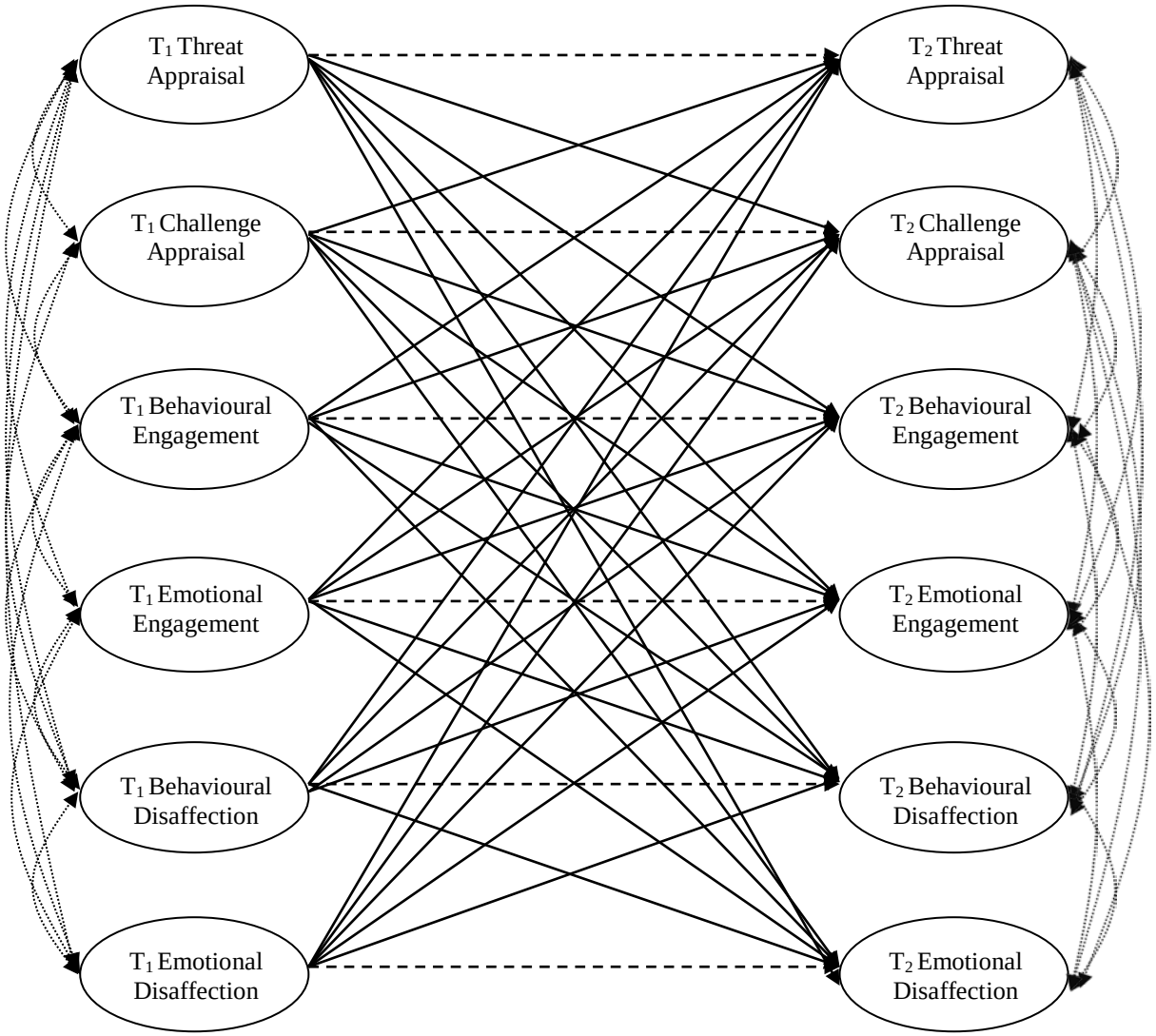

Figure 1. The full cross-lagged panel design testing the relationships between fear appeal appraisal and student engagement. Solid lines represent cross-lagged paths, dashed lines represent auto-lagged paths and dotted lines represent unlagged correlations. Gender and year group were included as covariates in the model but are not shown for simplicity.

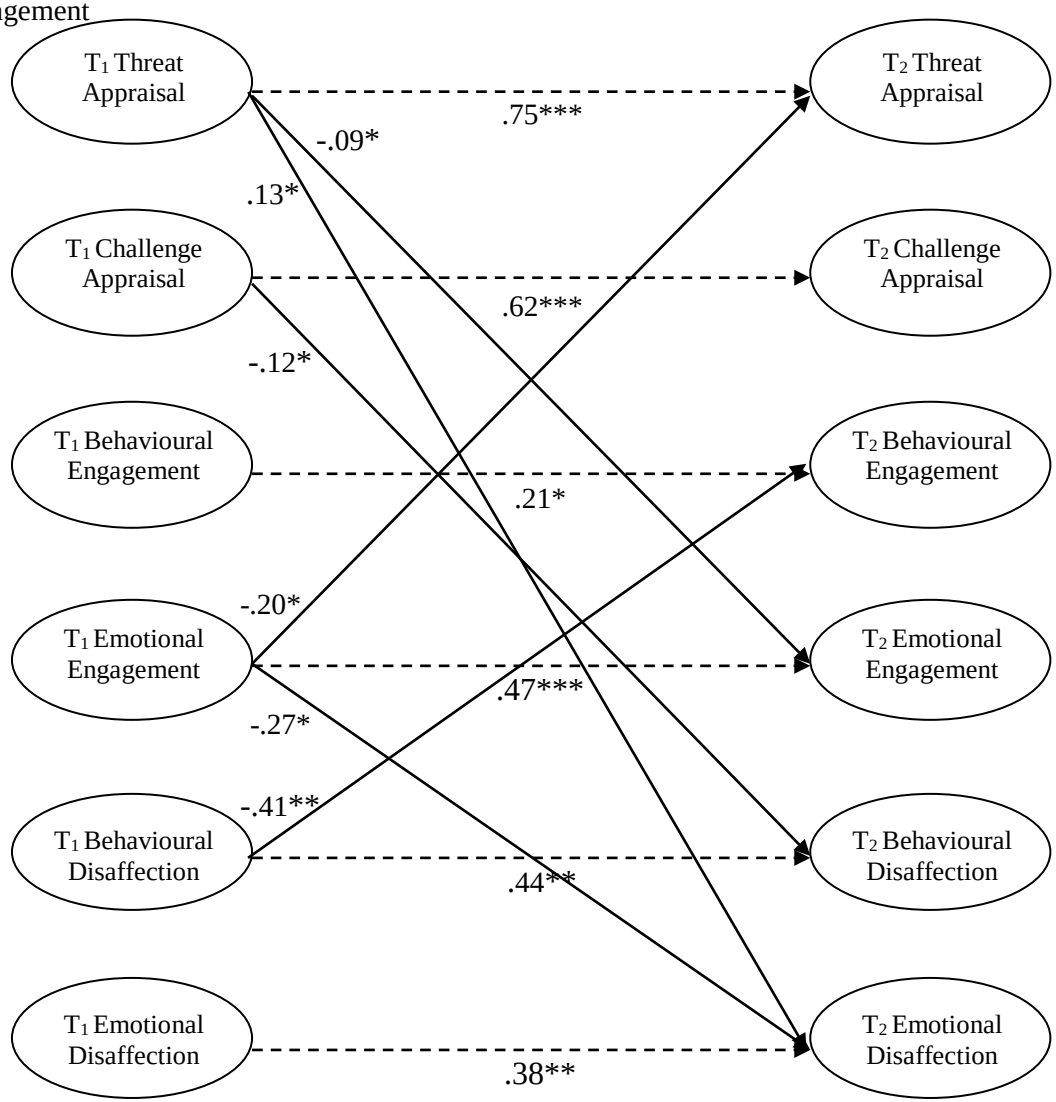

Figure 2. Statistically significant auto-lagged (dashed lines) and cross-lagged (solid lines) standardised path coefficients from the SEM (*p <.05, **p <.01, ***p <.001). Unlagged correlation coefficients are shown in Table 2. Statistically significant paths from the covariates (gender and year group) are not shown for simplicity but are reported in the text.