Inducing Challenge and Threat States: Effects on Learning and Self-Assessment Accuracy after Performance Feedback

Janne Lambermont (6979882)

Faculty Social Sciences, Utrecht University

201500002 Master's Thesis

First Assessor: Dr. Eva Janssen

Second Assessor: Vincent Hoogerheide

June 24, 2022

Word Count: 6856

Abstract

There is much variability of the effects of performance feedback on learning and self-assessment. The Biopsychosocial model of challenge and threat aims to explain the variabilities, through responses to emotions and stress. This paper conducted research to investigate if inducing a challenge or threat state affected one's learning and self-assessment after receiving feedback. A math lesson was created on enlargement and reduction of shapes for students of the pabo Rotterdam. The instruction of the lesson manipulated participants into a challenge or threat state. The t-test results showed no significant results for both learning and self-assessment. A mediation effect was analyzed through a relative challenge measurement. Results of the mediation analysis indicated no direct or indirect effect of relative challenge on learning and self-assessment. Even though no significant results were found, a study has potential if the sample size could be increased, and a motivated performance situation could be set for the participants.

Keywords: BPSM, challenge and threat states, learning from feedback, self-assessment accuracy

In educational settings, students' progress and performance get assessed regularly. This assessment is often accompanied by feedback. Hattie and Timperley (2007) describe feedback as a consequence of one's performance given by an agent (e.g., teacher, peer, parent) regarding aspects of set performance and understanding. For this reason, performance feedback is one of the most significant influences of learning (Carless & Boud, 2018; Hattie & Timperley, 2007). However, not every student benefits equally from performance feedback. The most consistent finding in the meta-analysis of Wisniewski et al. (2020) is the variability of feedback effect between individuals. Reasons for this variability refer to the quality of feedback characteristics (e.g., feedback timing and specificity) (Wisniewski et al., 2020), feedback processes (e.g., maladjustments to receiver's aspirations) (Carless & Boud, 2018), and most essentially, the receivers' response to the feedback (Carless & Boud, 2018; Winstone et al., 2021).

Winstone et al. (2021) claim that no matter how high the quality of the feedback is, students' learning and skill development cannot progress if they do not react to the feedback. In general, for feedback to be effective, students need to perceive it positively (Agricoli et al., 2020; Van der Schaaf et al., 2011). However, not every student perceives feedback as positive. Emotions could explain the variability in students' perceptions toward performance feedback (Brown & Creaven, 2017).

Emotions might also account for the variability in students' self-assessment accuracy. According to Panadero et al (2015), self-assessment is an important skill to have, because students who are trained in self-assessment show an increase in learning and academic performance and an enhanced task-specific self-efficacy. Panadero et al. (2017) conceptualize students' self-assessment as various mechanisms and techniques through which students assess and assign worth to the quality of their learning process and products. In addition,

Panadero et al. (2016) indicate that strong emotions could mislead an individuals' judgment of learning compared to their actual capability.

The biopsychosocial model (BPSM) of challenge and threat is a theoretical model that tries to explain individual differences towards stress and emotions (Blascovich, 2008). This paper aims to research whether the BPSM can explain individual differences in learning and self-assessment accuracy after receiving performance feedback.

Biopsychosocial Model

The biopsychosocial model of challenge and threat explains how individuals evaluate a task through individual demands and situational resources (Blascovich, 2008). In short, one tries to assess if one can cope with the task at hand. Examples of the resources in this evaluative process are skills, knowledge, abilities, and external support. Likewise, demands include the required effort into the task, the familiarity with and uncertainty toward the task, and psychological or physical danger (Moore, et al., 2014). The evaluative process toward a task is complex and mainly occurs unintentionally and automatically, without the individual being aware of this process (Moore et al., 2012; Seery, 2011), although conscious evaluations do occur (Blascovich, 2008). The outcome of this evaluative process is that the individual gets into a challenge or threat state. A challenge state manifests when one's resources meet or exceed the task's situational demands and a threat state occurs when the situational demands outweigh one's resources (Hase et al., 2019; Moore et al., 2014).

An essential premise of the BPSM is that a challenge or threat state occurs when the individuals find themselves in a motivated performance situation (Hase et al., 2019) or motivated goal pursuit (Seery et al., 2010). An example of such a motivated performance situation in education could be a good performance on an assignment or exam, or future career perspectives. BPSM claims that students who encounter relatively high resources and low demands during the (automated) evaluative process of a motivated performance situation

will likely experience this as a challenge. In contrast, students who encounter relatively low resources and high demands will likely experience this as a threat (Seery, 2011). For example, in a math class, an exam is upcoming. The student that studied for the exam might evaluate that his resources (e.g., his knowledge and skills) meet or outweigh the demands of the situation (e.g., familiarity with the topic) and get in a challenge state. In contrast, during the evaluative process of the student who did not study, the demands of the situation (e.g., uncertainty toward the exam) might outweigh the student's resources (e.g., knowledge level) and therefore fall into a threat state.

Challenge versus Threat States and Performance

As indicated, the BPSM of challenge and threat is applicable in many domains if a motivated performance situation is present or can be simulated, such as sports, education, or the medical field (Behnke & Kaczmarek, 2018, Haze et al., 2019). Numerous studies researched the effect of challenge and threat states on performance (Chalabaev et al, 2009; Martin et al, 2021; Scholl et al., 2017). Generally, independent of the domain, a challenge state tends to be beneficial to performance (Blascovich, 2008) and a threat state might obstruct performance (Gildea et al., 2007).

Learning from Feedback

Many studies focused on the effect of challenge and threat states on learning outcomes in the educational domain. For example, the systemic review of Haze et al. (2019) researched 38 studies, 20 of which involved cognitive performance or learning, containing 23 effects. In 15 (65%) of these outcomes researchers found an effect on performance favoring a challenge state, one study found an effect on performance favoring a threat state and seven studies found no effect. In addition, fifteen studies of the original 38 had an experimental design, comparing two groups created by either experimental manipulation or a median split in a continuous challenge and threat measurement. Only four of these experimental design studies

had outcomes involving cognitive performance or learning, namely mathematical and word finding task with related learning outcomes. However, none of these studies included performance feedback as variable the learning process, while according to Duijnhouwer (2012) performance feedback on a task can enhance both performance and motivation.

Feedback and Self-Assessment Accuracy

Feedback also seems to influence the relationship between self-assessment and learning. The meta-analysis of Sitzman et al. (2010) showed that the correlation between self-assessment and learning was stronger in courses with external performance feedback than in courses without performance feedback. There is much research on students' self-assessment having an impact on learning. According to the review of Topping (2003), there is evidence that self-assessment leads to improved effectiveness and quality of learning. Panadero et al. (2017) conceptualize students' self-assessment as various mechanisms and techniques through which students assess and assign worth to the quality of the learning processes and products. In addition to increasing one's learning, a students' self-assessment help them regulate their own learning through metacognitive monitoring of their work and processes against expectations, standards, and goals. (Panadero & Alonso-Tapia, 2013). However, no research was found on the effect of challenge and threat state on self-assessment accuracy as a performance outcome specifically.

Thus, the BPSM could explain why performance feedback is effective for some individuals, but not for others to improve their learning and self-assessment. However, little research has been conducted on how challenge and threat states affect learning from performance feedback and self-assessment. Therefore, the present study examines whether experimentally induced challenge and threat states affect students' learning from feedback and improving self-assessment accuracy. If we know this, we can strengthen theory on when

and why feedback is effective for learning and shape interventions in order to improve learning.

Mediator on Challenge and Threat State and Performance Outcomes

According to Moore et al. (2012), there is limited research on the potential underlying mechanisms through which challenge or threat states influences performance. However, Meijen et al. (2020) propose emotions as a possible underlying mechanism. A challenge state is associated with positive and negative emotions, whereas a threat state is associated with only negative emotions (Jones et al., 2009). A negative emotion, such as frustration, can be perceived as either supportive or discouraging by an individual. When a negative emotion supports the individual, it could place them in a challenge state. On the other hand, when a negative emotion discourages the individual, it could place them in a threat state.

Brown & Craeven (2017) indicate that receiving performance feedback can generate positive or negative emotions toward the task. As an educator, it is important to be aware of the emotions feedback can generate. Negative emotions, such as anxiety, anger, or frustration can hinder learning, whereas positive emotions, such as joy and pride can foster learning (Martin et al., 2021).

In addition, Panadero et al. (2016) indicate that strong emotions influence an individual's judgment, and therefore one's self-assessment. According to Baumeister et al. (2014), individuals make different kinds of judgments using their (task-irrelevant) emotions as a source of information. For example, individuals standing at the foot of a hill judge the hill as steeper as they listen to sad music and less steep as they listen to happy music (Riener et al., 2011).

Thus, if emotions can influence an individual's perception of the task, especially negative emotions can also influence if the induced challenge or threat state is perceived as intended. Therefore, this paper examines if the relative challenge of threat mediates the

relationship between the induced challenge and threat state and both learning from feedback and self-assessment accuracy.

The Present Study

In short, giving performance feedback is a regular occurrence in higher education. However, where performance feedback could help one person in their learning and self-assessment, it could hinder others. The BPSM model could help explain this variability between individuals. To strengthen the theory on how and when feedback is effective, this paper aims to research if an experimentally induced challenge or threat state affects learning from feedback and self-assessment accuracy. In line with the literature, one would expect that the effect on learning from feedback and self-assessment accuracy favors the induced challenge state.

In addition, as negative emotions, such as frustration and stress are linked to both a challenge and a threat state, the induced state might not be the perceived state of an individual. Therefore, this paper aims to examine the relationships between the induced challenge and threat state and the performance outcomes (learning from feedback and self-efficacy) are mediated by relative challenge measure.

In sum, the research questions this paper aims to answer are the following: a) What is the effect of a simulated challenge or threat state on learning from feedback?; b) What is the effect of a simulated challenge or threat state on self-assessment accuracy?; c) Is the effect of a simulated challenge or threat state on learning from feedback mediated by relative challenge?; and d) Is the effect of a simulated challenge or threat state on self-assessment accuracy mediated by relative challenge?

Method

Participants and Design

The initial target group of this study were undergraduate students from the 'pedagogische academie voor basisonderwijs (pabo) in Rotterdam, the Netherlands. In total, 350 first, second-, and third-year students received the invitation to participate. However, due to an unexpected low response rate, and the limited time span for data collection, it was decided to broaden the target group to students in higher education after two weeks. Even though 78 participants started with the data collection, only 39 participants reached the learning phases. Therefore, the final dataset consisted of 39 participants (see figure 1).

Of the 39 participants in the dataset, 82.1% were female, 17.9% were male, and 0% indicated "other". The participants' mean age was 32.51 years (SD = 13.23). Sixteen participants originated from the 'pabo' target audience (M = 26.25, SD = 11.65), and 23 participants came from the other higher education students (M = 34.48, SD = 14.15). Other demographic data were not gathered for privacy reasons. The participants did not receive any financial compensation or study credit. Prior to data collection, participants filled in the informed consent and were notified they could withdraw at any time. If participants finished the study they were debriefed.

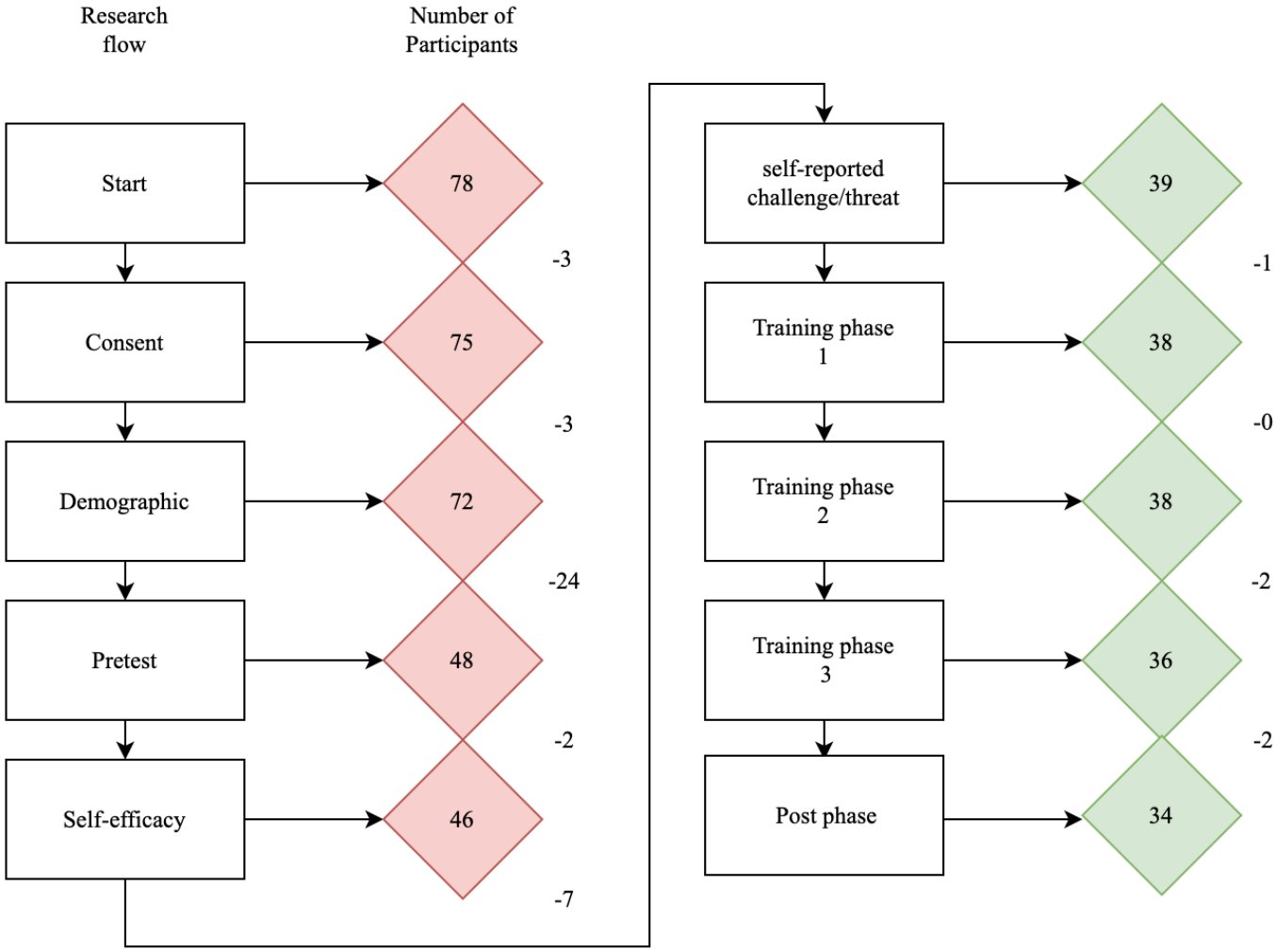
The study was an experimental, quantitative research. All participants were randomly assigned to one of two experimental conditions: a challenge condition or a threat condition. The challenge condition consisted of 19 participants, 18 of whom finished the whole experiment. Of these 19 participants, 79.9% were female, and 21.1% were male. The threat condition consisted of 20 participants, 16 of whom finished the entire experiment. Of these 20 participants, 85% were female and 15% were male.

Measures

The experiment was developed for a math course for pabo-students at the University of Applied Science in Rotterdam. A math lesson on the enlargement and reduction of shapes was given through a video, after which a learning task followed. The entire experiment was

Figure 1

Flow Chart Sample Size and Drop-out during each Research Phase



programmed in Qualtrics, that participants did on their private computer. The course coordinator sent an e-mail with the link to the study, in which it was strongly advised to do participate in the study as preparation for the course.

Instruction Induced Challenge and Threat States

Instruction for the math learning task aimed to place the participants into the manipulation of the experiment. Two written instructions were developed: one for the challenge condition and one for the threat condition. Participants who received the challenge instruction read a positive text. This text tried to stimulate them to view the task as a task they had the resources to accomplish. It contained sentences such as: “We think you are more than

capable of accomplishing this task." and "Even though this task seems difficult, keep reminding yourself you can do it.". Participants who received the threat instruction read a negative text, aiming to outline the high demands of the task. The text contained sentences such as: "These types of tasks can be difficult and frustrating." and "Most students encounter problems with these types of tasks.". In addition, the threat instruction indicated that speed would be included as a performance measure to stimulate a motivated performance situation. An example of such a sentence was: "It is important that you execute the task as fast and efficient and possible.". Previous research formed the basis for both instructions (Feinberg & Aiello, 2010; Moore et al., 2012). These instructions were combined, translated, and modified to fit the math learning task. See Appendix A for both complete instructions.

Math Learning Task

A math learning task on enlargement and reduction of shapes was developed for a math course at the pabo Rotterdam. According to the coordinator of the math course, students at the pabo typically struggle with these types of math problems, hopefully inducing a motivated performance situation for the students.

The task started with a how-to video, which lasted approximately eight minutes. The video explained how to enlarge or reduce shapes in five steps. After watching the video, participants continued with the training phase. The training phase consisted of three math questions, which participants had to solve through five steps explained in the video. Therefore, each question consisted of five sub questions. During the whole training phase, participants could fall back supporting materials. The three main questions were ordered from easy to difficult. After each question, feedback was provided. This feedback informed participants how many steps they had performed correctly and showed the correct answers for each step. If necessary, the calculation toward the right answer was provided.

The final part of the learning task was the post phase. This phase consisted of five questions on enlargement and reduction of shapes that participants had to solve without supporting materials. In addition, they were not guided through the questions by the previously mentioned steps, but they had to fill in the correct answer immediately. After each question, feedback in the form of the correct answer and calculation toward to correct answer were provided.

The learning task was developed with a fellow master's student. After development, and expert (coordinator of the math course) checked and provided feedback on the learning task, ensuring the validity of the task.

Self-Reported Challenge and Threat Questionnaire

To measure if the participants experienced the manipulated challenge and threat instruction as intended, they filled in a questionnaire developed by Scholl et al. (2018), who adopted items from the Stress Appraisal Measure (Peacock & Wong, 1990). Through this 12-item questionnaire, participants self-reported their challenge or threat response to their given instruction. Examples of the items were: "when I think about the learning task ahead, I feel sure I have the skills to master this task." and "when I think about the learning task ahead, I feel I will not meet the expectations of others.". Participants answered the items on a 7-point Likert-scale (with 1 indicating 'totally not applicable to me' and 7 indicating 'totally applicable to me'). According to Dawes (2008) a 5 or a 7-point Likert-scale increased reliability and validity compared to smaller or larger scales. Scholl et al. (2018) validated their own questionnaire through an exploratory factor analysis. All items loaded on one single factor, and after recoding averaged in a relative challenge index (α = .89). For this study, a similar relative measurement was used.

After adaption and translation of items into Dutch, Cronbach's Alpha was used to indicate the reliability of the adapted questionnaire. The Cronbach's Alpha of the translated

the modified challenge and threat questionnaire was α =.77, indicating a high reliability (Trizano-Hermosilla & Alvardo, 2016). See Appendix B for the used relative challenge questionnaire.

Learning from Feedback

The learning from feedback variable was measured through adding the participant's correct answers. This was done for the training phase and the post phase separately. In the training phase participants answered three questions, with five sub questions each. For each correct answer, the participant received one point and for each incorrect answer zero points. Therefore, the total amount of points the participant could earn in the training phase is 15 points. In the post phase participants answered five questions, for which one point could be earned for each correct answer, and zero points for each incorrect answer. In total, the participants could earn five points in the post phase. The learning from feedback variable consisted of the total amount of points the participants had. Both phases were measured separately.

Self-Assessment Accuracy

The participant's self-assessment accuracy was measured through rating one's own work. Participants could fill in how many of the (sub)questions they thought had answered correctly in both the training phase as the post phase. The training phase consisted of three questions, each with five sub questions. Therefore, participant could self-rate to have a maximum of 15 correct answers. The post phase consisted of five questions, for which participants could self-rate a maximum of five correct answers. The absolute difference between the participant's self-rated correct answers and the actual correct answers made the self-assessment accuracy variable. This was measured for the training phase and post phase separately.

Procedure

The experiment's entire procedure is written out for the study's replicability and completeness. However, the experiment was conducted with a fellow master's student who measured different variables. Therefore, not every variable is applicable for this paper. An asterisk highlights the measures or variables that were not relevant for this paper.

Participants performed the experiment entirely from their private computers. The course coordinator asked them to participate through an e-mail with a link. The experiment was divided into three phases. In the first phase, the participants received an introduction into the study, with the intention and duration. In addition, participants read and actively checked the informed consent. Afterward, participants filled in demographic information, such as gender, age, and their study year (if applicable), and conducted a two-question pretest on the subject. This phase lasted approximately eight minutes.

The second phase started with a task-specific *self-efficacy questionnaire. Afterward, participants were randomly assigned to a challenge condition or a threat condition. In their assigned condition participants read either a challenging instruction or a threatening instruction for the learning task, inducing a challenge or threat state (Appendix A). After the instruction, participants filled in the self-reported challenge and threat questionnaire and the *cognitive appraisal ratio. This second phase lasted approximately five minutes.

The third and final phase was the learning task. The learning task consisted of a how-to video, the training phase, and the post phase. The training phase consisted of three questions, each with five sub questions. The post phase consisted of five questions. This third phase took approximately 25 to 30 minutes. After finishing the experiment, participants were debriefed and thanked for their participation. Utrecht University provided ethical approval.

Analyses

Prior to data collection, a power analysis was conducted using G*power 3.1.9.2 (Faul et al., 2007), assuming a fixed alpha level of 0.05 and a power of 0.80. Results indicated that a

sample of $N = 147$ participants would be sufficient to test for a small to medium mediation effect ($f = 0.085$) when considering 0.10, 0.25, and 0.40 as the cut-offs for small, medium, and large effects, respectively (Cohen, 1988). Note that the intended sample size was not reached. This will be discussed further in the discussion section.

Data were analyzed using IBM SPSS Statistics v28. The quantitative data was extracted from Qualtrics in an Excel file, after which all participant ($n = 39$) who did not reach the learning phases were deleted from the file, because they did not have an outcome variable. Next, the file was uploaded tot SPSS. First, the threat measurement in the challenge and threat questionnaire were reversed, to measure the relative challenge score. Second, per phase the amount correct answers per phase were added to calculate the 'learning from feedback'-variable and the absolute difference for the self-assessment was calculated per phase to create the 'self-assessment accuracy'-variable.

An independent t-test examined the effect of induced challenge and threat state on learning from feedback and the effect of induced challenge and threat state on self-assessment accuracy. Violations of assumptions for both t-tests were checked. The data seemed normally distributed and according to Levene's test homoscedastic. There were no outliers in the data.

A PROCESS v4.1 by Andrew F. Hayes analyzed the mediating effect of the participant's relative challenge state of the induced challenge or threat state on learning from feedback and self-assessment accuracy. Violations of the mediation analysis, normally distributed, homoscedastic, and outliers were checked as well and showed no concern.

Results

The descriptive statistics consist of the means and standard deviations of the learning outcomes of the training phase and the post phase per condition (table 1). In addition, table 2 contains the means and standard deviations of the self-assessment accuracy per condition. The pretest revealed low prior knowledge. The challenge condition had an average of $M = 0.89$

(*SD* = 0.81) of the two points in total, and the threat condition averaged *M* = 0.70 (*SD* = 0.73) of the two points.

Table 1

Descriptives Learning from Feedback per Phase

Phase	Condition	*N***	*M**	*SD*
Training phase 1	Threat	20	3.60	1.19
	Challenge	19	4.21	1.27
Training phase 2	Threat	20	4.15	1.09
	Challenge	18	4.22	0.94
Training phase 3	Threat	18	3.28	0.96
	Challenge	18	3.83	1.04
Post phase 1	Threat	16	2.25	1.69
	Challenge	18	3.06	1.39
Total sum all phases	Threat	16	13.75	3.40
	Challenge	18	15.44	3.45

*Note: *per phase a maximum of five points could be earned*

*** explanation differences in sample size, see figure 1*

Table 2

Descriptives Self-Assessment Accuracy per Phase

Phase	Condition	*N***	*M**	*SD*
Training phase 1	Threat	20	1.10	1.21
	Challenge	19	0.79	1.27
Training phase 2	Threat	20	0.95	0.99
	Challenge	18	0.56	0.78
Training phase 3	Threat	18	1.22	1.00

	Challenge	18	0.72	0.96
Post phase 1	Threat	16	1.13	0.89
	Challenge	18	0.78	0.73
Total sum all phases	Threat	16	4.24	2.30
	Challenge	18	2.72	1.99

*Note: *the closer to zero (absolute difference), the more accurate the self-assessment*

*** explanation differences in sample size, see figure 1*

Learning from Feedback

To analyze if the induced challenge and threat state (manipulation) effected learning from feedback, an independent sample t-test was performed. First, the training phase was analyzed because participants received feedback throughout this phase. On average, participants in the challenge condition performed better in the training phase (M = 12.39, SD = 2.40) than participants in the threat condition (M = 11.11, SD = 2.42). Contrary to the first hypothesis, this difference, -1.27, BCa 95% CI [-2.90, 0.19], was not significant, t (34) = -1.58, ρ = .06. This represented an effect of d = .53, which indicated a medium effect size.

Second, a t-test analyzed the results of the post phase. As students went through the training phase and received the feedback, the post phase could indicate if they learned from the feedback. On average, participants in the challenge condition performed better in the post phase (M = 3.06, SD = 1.39), than participants in the threat condition (M = 2.25, SD = 1.69). However, this difference -0.81, BCa 95% CI [-1.88, 0.27], was again not significant, t (32) = -1.52, ρ = 0.14. This result represented an effect of d = .52, which is again a medium effect.

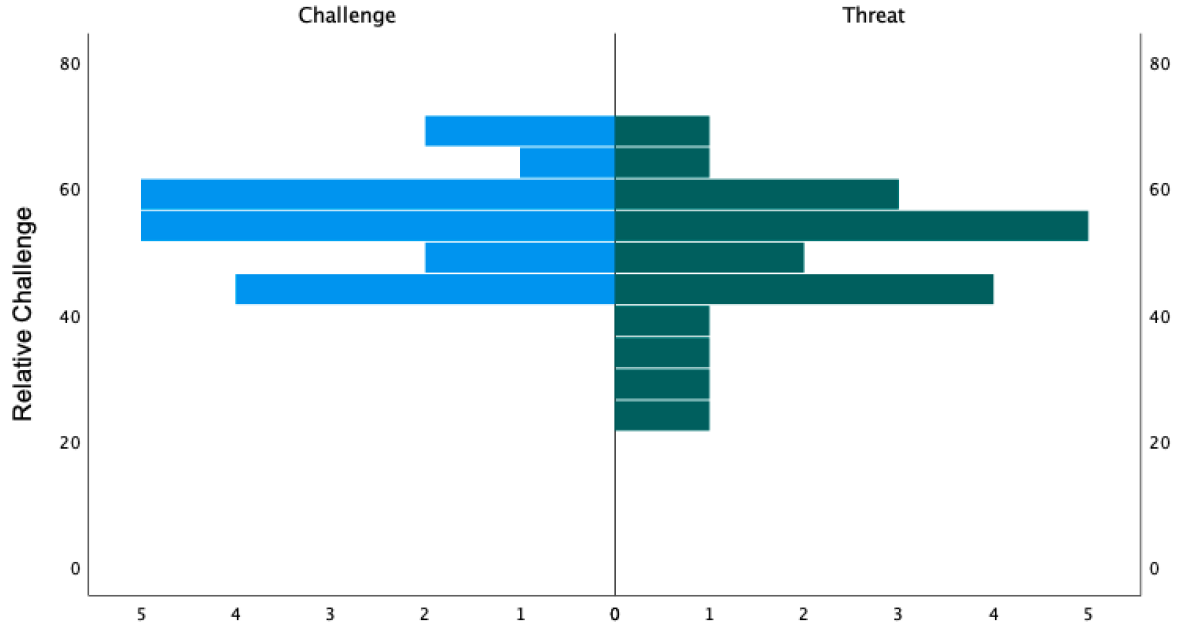
Mediating Effect of Relative Challenge State on Learning from Feedback

First, the relative challenge measure from the challenge and threat questionnaire could indicate if the manipulation of inducing challenge or threat was successful. An independent t-test analyzed the following: on average, participants in the challenge condition reported a higher relative challenge toward the learning task (M = 55.42, SD = 7.99) than participants in

the threat condition (M = 48.75, SD = 11.12). This difference, -6.67, BCa 95% CI [-12.94, -0.40] was significant, t (34) = -2.14, ρ = 0.04 (see figure 2). This resulted in a medium effect of d = -.68. This indicated that the manipulation was successful indeed.

Figure 2

Histogram Relative Challenge Variable per Condition



Note: the higher the self-reported number the more relative challenge

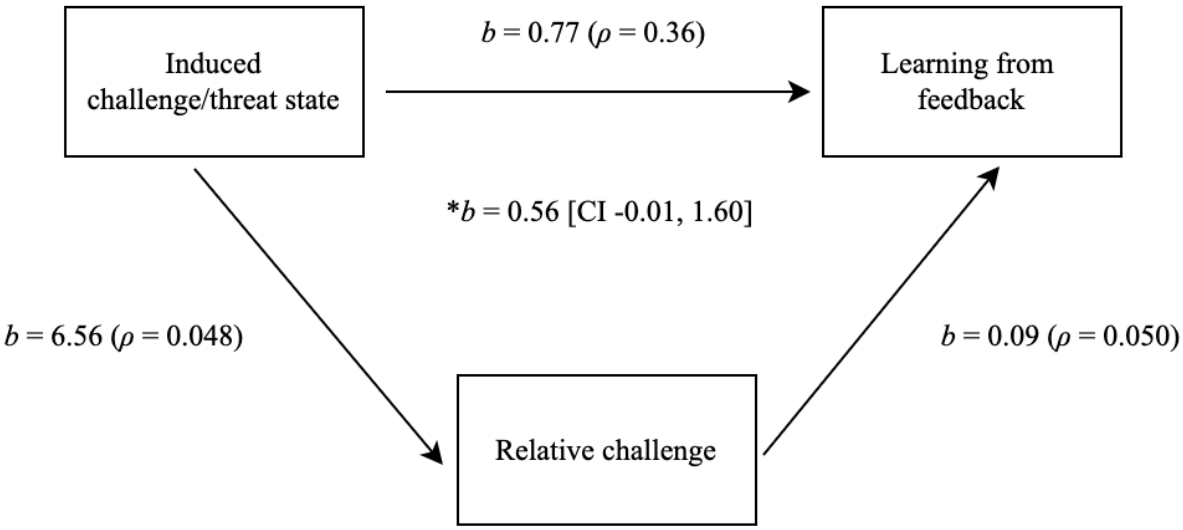
Note: N = 39

Second, a mediation analysis analyzed if the relative challenge variable mediated the relationship of induced challenge and threat state on learning from feedback. The analysis started with the training phase. As expected from the previous t-test, the analysis showed a significant effect of induced challenge and threat state on relative challenge (b = 6.56, 95% CI [0.06, 13.05], t =2.05, ρ = 0.048), indicating that the given instruction directly affected if participants felt relatively more or less challenged. The relative challenge of participants had a direct effect on learning from feedback (b = 0.09, 95% CI [0.00, 0.17], t = 2.03, ρ = 0.050), indicating that a higher relative challenge was related to better learning from feedback outcomes. Finally, the direct effect of induced challenge and threat on learning from feedback was not significant (b = 0.77, 95% CI [-0.93, 2.47], t = 0.92, ρ = 0.36). In addition, no indirect effect of relative challenge was found (b = 0.56, 95% BSa CI [-0.01, 1.60]), indicating that

relative challenge did not mediate a potential effect of induced challenge and threat on learning from feedback in the training phase. Figure 3 displays all effect sizes and ρ-values.

Figure 3

Effect sizes Mediation Analyses Learning from Feedback in Training Phase



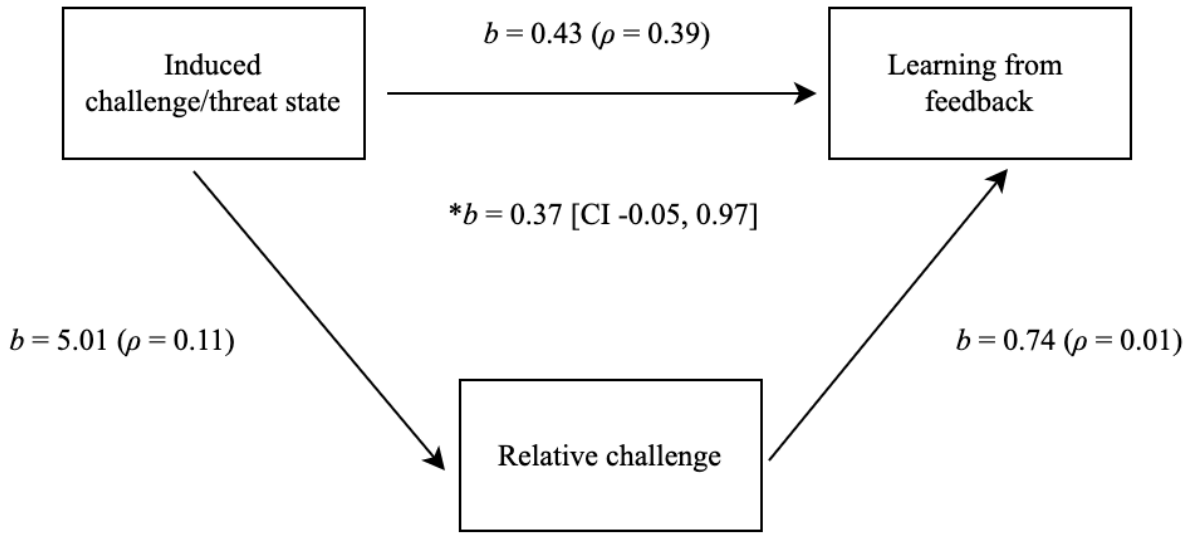
*Note: * indicates the indirect effect; **N = 36*

Next, the post phase was analyzed. Surprisingly, this analysis showed no significant effect of induced challenge and threat state on relative challenge (b = 5.01, 95% CI [-1.14, 11.15], t = 1.66, ρ = 0.11). This result indicates that the given instruction did not directly affect if participants felt relatively more or less challenged, and thus the manipulation failed. This is interesting, as the above-mentioned t-test found a significant result with slightly more participants (this issue will be addressed further in the discussion section). The relative challenge of participants had a significant effect on learning from feedback (b = 0.74, 95% CI [0.02, 0.13], t = 2.59, ρ = 0.01). This indicates that a higher relative challenge was related to better learning from feedback outcomes. Finally, the direct effect of the induced challenge and threat state on learning from feedback was not significant (b = 0.44, 95% CI [-0.60, 1.47], t = 0.86, ρ = 0.40). This is a similar result as the t-test showed. The indirect effect was insignificant as well (b = 0.37, 95% BSa CI [-0.05, 0.98]), indicating that relative challenge

did not mediate the relationship between induced challenge and threat state on learning from feedback during the post phase. Figure 3 displays all effect sizes and ρ-values

Figure 4

Effect sizes Mediation Analyses Learning from Feedback in Post Phase



*Note: * indicates the indirect effect; **N = 34*

Self-Assessment Accuracy

To analyze if the induced challenge and threat state affected self-assessment accuracy another two t-test were performed. First, the results of the training phase indicated that on average, participants in the challenge condition had a better self-assessment accuracy (M = 1.94, SD = 1.83), than participants in the threat condition (M = 3.28, SD = 2.27). This difference, 1.33 BCa 95% CI [-0.05, -2.89] was significant t (34) = 1.93, ρ = .031. This represented a medium effect size of d = .64. This result lines up with the hypothesis that participants in a challenge state are better at self-assessing their own work.

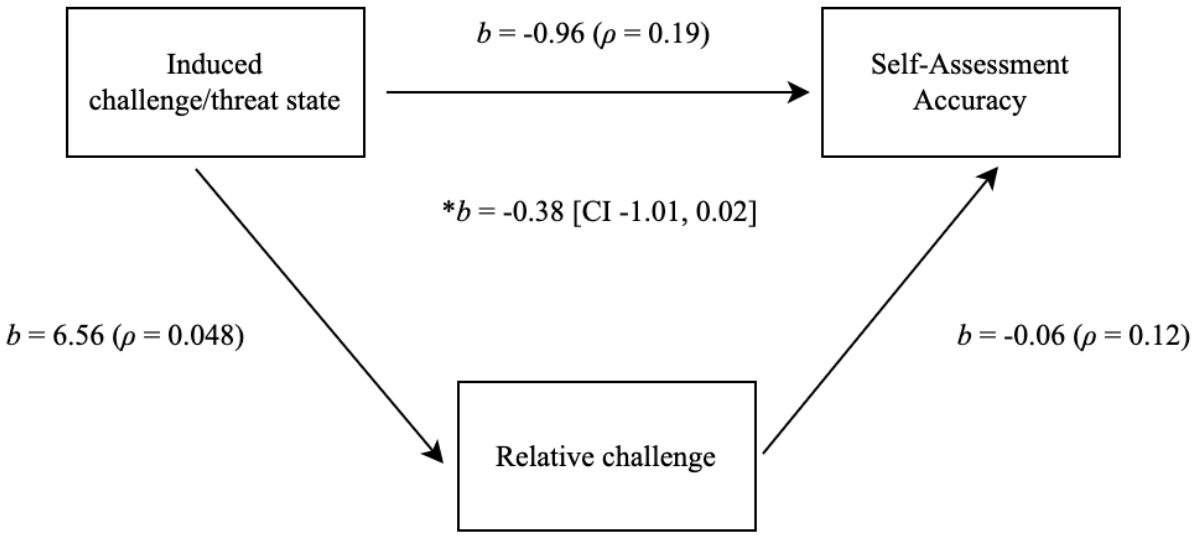
However, the t-test results on the post phase show a different outcome. On average, in the post phase participants in the challenge condition had a better self-assessment accuracy (M = 0.78, SD = 0.73), than participants in the threat condition (M = 1.13, SD = 0.89). This difference, 0.35 BCa 95% CI [-0.18, 0.91] was not significant t (32) = 0.17, ρ = .110. This represented a medium effect of d = .43.

Mediating Effect of Relative Challenge Self-Assessment Accuracy

A second pair of mediation analyses analyzed if the relative challenge mediated the relationship between the induced challenge and threat state on self-assessment accuracy in the training phase and the post phase. In the training phase, direct effect of induced challenge on relative challenge was significant (b = 6.56, 95% CI [0.06, 13.05], t = 2.05, ρ = .048), indicating that the instruction directly affected if participants felt relatively more of less challenged. The direct effect of relative challenge on self-assessment was not significant in the training phase (b = -0.06, 95% CI [-0.13, 0.02], t = -1.57, ρ = .12), indicating that a higher relative challenge did not lead to a better self-assessment accuracy. The direct effect of induced challenge and threat on self-assessment accuracy was not significant (b = -0.96, 95% CI [-2.42, 0.50], t = -1.57, ρ = .19). Furthermore, the indirect effect was not significant (b = -0.38, BSa 95% CI [-1.01, 0.02], again indicating no mediating effect of relative challenge. Figure 5 shows the effect sizes and ρ-values.

Figure 5

Effect Sizes Mediation Analysis Self-Assessment Accuracy in Training Phase

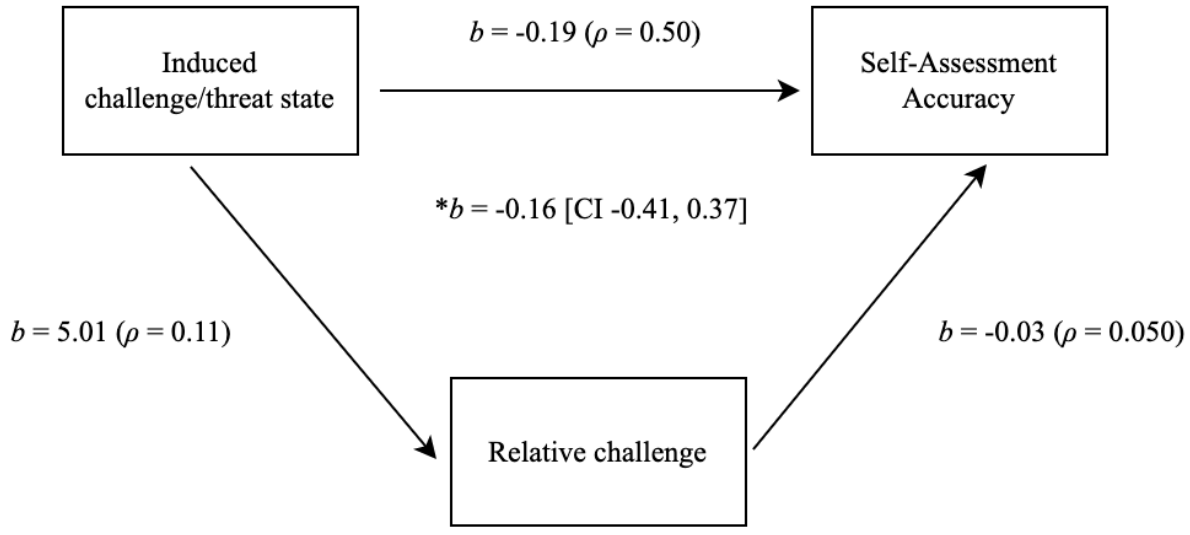


*Note: * indicates the indirect effect; **N = 36*

In the post phase, the direct effect of induced challenge and threat state on relative challenge was not significant, which is the same result as the previous mediation analysis (b = 5.01, 95% CI [-1.14, 11.15], t = 1.66, ρ = .11). The relative challenge of participants on the self-assessment accuracy was significant (b= -0.03, 95% CI [-0.06, 0.00], t = -2.04, ρ = .050). This indicates that a higher relative challenge was related to better self-assessment accuracy. The direct effect (b = -0.19 95% CI [-0.75, 0.37], t = -0.69, ρ = .498) and the indirect effect (b = -0.16 BSa 95% CI [-0.41, 0.37] of the induced challenge and threat state on self-assessment accuracy were both not significant, indicating the relative challenge did not mediate the relationship. Figure 4 displays all effect sizes and ρ-values.

Figure 6

Effect Sizes Mediation Analysis Self-Assessment Accuracy in Post Phase



*Note: * indicates the indirect effect; **N = 34*

Discussion

To complement theory and examine how and when feedback is effective, this paper aimed to research if an experimentally induced challenge and threat state affect learning and self-assessment accuracy after receiving performance feedback. An experiment in the form of a math lesson for pabo-students was created to explore this aim. Participants were randomly

divided into either a challenge condition or a threat condition. The challenge condition read a positive instruction focusing on the individual's resources, and the threat condition read a negative instruction focusing on the task's demands. Afterward, the participants continued the math lesson with a training phase, which included feedback, and a post phase with related questions. In addition to answering the questions, participants filled in how many questions they thought they answered correctly to measure self-assessment accuracy. Based on the theory, it was hypothesized that both the learning from feedback variable and the self-assessment accuracy favored the challenge condition, indicating that the participants in the challenge condition performed better.

T-tests analyzed the training and post-phase results for the variable 'learning from feedback'. The t-test showed that in both the training and the post-phase, the challenge condition, on average, performed better on learning from feedback than the threat condition, but both results were not significant. These results could indicate that, contrary to the hypothesis, participants in a challenge state did not perform better on the outcome variable learning from feedback than participants in a threat state. The t-tests of the training phase and the post phase for the variable self-assessment accuracy again showed that the challenge condition, on average, had a better self-assessment accuracy than the threat condition. The result of the training phase was significant. This result is in line with the hypothesis that participants in a challenge state can self-assess better than participants in a threat state. However, the results of the post-phase were not significant.

As three out of the four t-test results were not significant, this could indicate that placing students in either a challenge or threat state does not affect how performance feedback is perceived and its consequences for learning or self-assessment. One explanation for the insignificant results on the learning from feedback variable and the self-assessment variable could be the participants' lack of motivated performance state during the experiment. As the

literature indicates, an essential premise of the BPSM is that a challenge or threat occurs when individuals find themselves in a motivated performance situation (Hase et al., 2019). And even though the manipulation of the experiment was successful and the participants in the challenge condition felt relatively more challenged than participants in the threat group, the experiment was not an essential part of the math course. Therefore, participants could feel less engaged in the task at hand and did not get into a motivated performance situation. According to Behnke et al. (2021), task engagement can be considered the motivation to engage in the upcoming task successfully. And individuals who exhibit task engagement, the context can be referred to as a motivated performance situation. Therefore, being in a motivated performance situation and the corresponding task engagement is essential for experiencing challenge or threat (Sassenberg & Scholl, 2019).

The lack of social presence is an alternative explanation for the insignificant results for learning from feedback and self-assessment accuracy. Blascovich et al. (1999) believe that the presence of others during a task increases the goal relevance of performance. Participants executed the learning task from their private computers, presumably alone. In addition, the learning task topic (enlargement and reduction of shapes) was something they had not yet learned during the math course. According to Blascovich et al. (1999), individuals should experience greater threat and exhibit characteristic cardiovascular response patterns that accompany threat when they perform unlearned tasks in the presence of others. The reasoning is the lack of resources the individuals bring to the performance situation.

A significant result indicated that inducing a challenge or threat state could affect self-assessment accuracy in the training phase. This result shows that placing individuals in a challenging state could affect performance feedback and more accurately assess their work. Students answered three questions in consequential steps with additional supporting materials. After the first question, participants know the steps to take and can better assess their work for

the following questions. Andrade and Valtcheva (2009) name practice as a feature an individual needs for effective self-assessment. The repetitive nature of the training phase could be considered practice for learning and self-assessment.

The second aim of this study was to examine if the relationship between the induced challenge and threat state and the two outcome variables (learning from feedback and self-assessment accuracy) was mediated by a participants' relative challenge. The relative challenge variable did not mediate the relationship with the induced challenge or threat state for both the training and post-phase outcome variables, indicating no causal effect. The measurement of the relative challenge could explain these insignificant results. Even though the manipulation was successful and the participants in the challenge group felt relatively more challenged than participants in the threat group, a self-reporting measurement could foster bias. According to Hanley and Sikka (2012), self-reporting bias is caused by individuals' desire to avoid providing self-discriminating information about their driving behavior. This bias could therefore confound the observed differences between groups. With the self-reporting bias combined with a small sample size (see limitations for further explanation), the outcome results might not indicate realistic treatment effects.

Limitations and Recommendations

Even though the set-up of the study was strong and could easily be transferred to a live classroom setting, this study had potential limitations. First, the pabo-sample consisted of 16 participants, and the general students' sample consisted of 23 participants. This sample size is too small to generalize the results. In studies with small and varied sample sizes, random variations can cause the possible found effects. According to Kang (2021), studies with inappropriate samples do not provide accurate estimates and could report wrong information on the treatment effect. This is illustrated perfectly by the manipulation check. A t-test analyzing the induced challenge and threat state on relative challenge gave a significant result,

indicating the manipulation was successful. The sample for this t-test was 39 participants. However, when repeated in the mediation analysis, the direct effect of induced challenge and threat on the relative challenge was insignificant. The mediation analysis had a sample of 34 participants. This indicates that even a little variability in sample sizes could result in ambiguous results. An explanation for the small sample size could be that the learning task, particularly after viewing the pretest questions, was too complex for the pabo-students and the 'general' students. The flowchart of participants' drop-out rates (figure 1) showed that approximately one-third (n = 24) of the participants dropped out right before the pretest started. As the motivated performance situation for the participants was not measured and therefore perhaps absent or limited, participants might be faster inclined to drop out of the study due to the difficulty of the topic. Increasing the sample size would make the estimations of effects more accurate, especially if the power is met. Then the study could be more generalizable.

Another limitation of the study was the limited time in which it was conducted. As the study is a Master's thesis, there was limited time to collect the data. However, the BPSM is a complex model, that deserves more research time to accurately measure all aspects involved. For example, relative challenge is best measured through physiological measures, such as heart rate and blood pressure (Moore et al., 2012). However, these measurements are time consuming and in need of expertise. In addition to the relative challenge measurement, more time would affect the response rate of the study.

As a recommendation, it is advised to repeat the study with a clear, motivated performance situation for future research. To replicate the experiment as part of a course by making it a mandatory homework assignment or a classroom assignment. This last example would help with a social presence, an aspect that could create larger threat responses. The set-up of the experiment is strong and could easily be transferred to a live classroom setting. In

addition, if significant results are achieved at replication, current and future educators could benefit by thinking about the way they give performance feedback.

Conclusion

To summarize and answer the main research questions, this study showed that inducing a challenge or threat state on participants did not affect their perception of performance feedback and consequent learning and self-assessment. Explanations for these results could be the lack of a motivated performance situation for the participants in the study and the lack of social presence that could enhance threat responses in participants. If the study is replicated, making the study part of a course through a homework assignment or classroom assignment would be advised, as it would account for both social presence and a motivated performance situation.

In addition to these results, no mediating effect was found in the participants' relative challenge on the relationship between induced challenge threat and the outcome variables in both phases. The measurement of the relative challenge could explain these results, as self-reporting a measurement could create a self-reporting bias. Controlling for this bias in a future study could lead to more significant results.

Even though there are no significant results, the study shows potential. Variance in sample sizes indicates that potential effects can be found when the sample size increases to meet the power for analyses. Furthermore, the set-up of the study is transferrable to a classroom setting, making it easy to replicate. If significant results are found in a replicate study, it could add to existing literature. In addition, it could benefit (future) educators in the way they provide feedback to their students and maximize potential learning.

References

Agricola, B. T., Prins, F. J., & Sluijsmans, D.M. (2020). Impact of feedback requests forms and verbal feedback on higher education students' feedback perception, self-efficacy, and motivation. *Assessment in Education: Principles, Policy & Practice, 27*(1), 6-25. https://doi.org./10.1080/0969594X.2019.1688764

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into practice*, *48*(1), 12-19

Baars, M., & Wijnia, L. (2018). The relation between task-specific motivational profiles and training of self-regulated learning skills. *Learning and Individual Differences, 64,* 125-137. https://doi.org/10.1016/j.lindif.2018.05.007

Baumeister, R.F., Alquist, J. L., & Vohs, K.D. (2015). Illusions of learning: Irrelevant emotions inflate judgments of learning. *Journal of Behavioral Decision Making, 28*(2), 149-158.

Behnke, M., & Kaczmarek, L. D. (2018). Successful performance and cardiovascular markers of challenge and threat: A meta-analysis. *International Journal of Psychophysiology,* 130, 73-79. https://doi.org/10.1016/J.ijpsycho.2018.04.007

Behnke, M., Hase, A., Kaczmarek, L. D., & Freeman, P. (2021). Blunted cardiovascular reactivity may serve as an index of psychological task disengagement in the motivated performance situations. *Scientific Reports*, *11*(1), 1-10

Blascovich, J. (2008). Challenge and threat. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 431-445). New York: Psychology Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., & Salomon, K. (1999). Social" facilitation" as challenge and threat. *Journal of personality and social psychology*, *77*(1), 68

Brown, E.G., & Creaven, A. (2017). Performance feedback, self-esteem, and cardiovascular adaption to recurring stressors. *Anxiety, Stress, & Coping, 30*(3), 290-303. https://doi.org/10.1080/10615806.2016.1269324

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education. 43*(8), 1315-1325. https://doi-org.proxy.library.uu.nl/10.1080/02602938.2018.1463354

Chalabaev, A., Major, B., Cury, F., & Sarrazin, P. (2009). Physiological markers of challenge and threat mediate the effect of performance-based goals on performance. *Journal of Experimental Social Psychology, 45*(4), 991-994. https://doi.org/10.1016/j.jesp.2009.04.009

Chalabaev, A., Major, B., Sarrazin, P., & Cury, F. (2012). When avoiding failure improves performance: Stereotype threat and the impact of performance goals. *Motivation and Emotion, 36*(2), 130-142. https://doi.org/10.1007/s11031-011-9241-x

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2. ed., reprint). Psychology Press

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research, 1*(50), 61-104.

Duijnhouwer, H., Prins, F.J., & Stokking, K.M. (2012). Feedback providing improvement strategies and reflection on feedback use: Effects on students' writing motivation, process, and performance. *Learning and Instruction, 22*(3), 171-184. https://doi.org/10.1016/j.learninstruc.2011,10.003

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Feinberg, J. M., & Aiello, J. R. (2010). The effect of challenge and threat appraisals under evaluative presence. *Journal of Applied Social Psychology, 40*(8), 2071-2104. https://doi.org/10.1111/j.1559-1816.2010.00651

Fong, C. J., Schallert, D. L., Williams, K. M., Williamson, Z. H., Lin, S., Kim, Y. W., & Chen, L. H. (2021). Making feedback constructive: the interplay of undergraduates' motivation with perceptions of feedback specificity and friendliness. *Educational Psychology,* 1-19. https://doi.org/10.1080/01443410.2021.1951671

Gildea, K. M., Schneider, T.R., & Shebilske, W.L. (2007). Stress appraisals and training performance on a complex laboratory task. *Human factors, 49*(4), 745-758. https://doi.org/10.1518/001872007X215818

Hanley, P. F., & Sikka, N. (2012). Bias caused by self-reporting distraction and its impact on crash estimates. *Accident Analysis & Prevention*, *49*, 360-365.

Hase, A., O'Brien, J., Moore, L. J., & Freeman, P. (2019). The relationship between challenge and threat states and performance: A systemic review. *Sport, Exercise, and Performance Psychology, 8*(2), 123. Doi.10.1037/spy0000132.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi-rg.proxy.library.uu.nl/10.3102/003465430298487

Jones, M., Meijen, C., McCarthy, P. J., & Sheffield, D. (2009). A theory of challenge and threat states in athletes. *International Review of Sport and Exercise Psychology, 2*(2), 161-180. https://doi.org/10.1080/17509840902829331

Kang, H. (2021). Sample size determination and power analysis using the G* Power software. *Journal of educational evaluation for health professions*, *18*.

Martin, A. J., Kennett, R., Pearson, J., Mansour, M., Papworth, B., & Malmberg, L. E. (2021). Challenge and threat appraisals in high school science: Investigating the roles of

psychological and physiological factors. *Educational Psychology, 41*(5), 618-639. https://doi.org/10.1080/01443410.2021.1887456

Meijen, C., Turner, M., Jones, M. V., Sheffield, F., & McCarthy, P. (2020). A theory of challenge and threat states in athletes: A revised conceptualization. *Frontiers in psychology, 11*(126). https://doi.org/10.3389/fpsyg.2020.00126

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2012). The effect of challenge and threat states on performance: An examination of potential mechanisms. *Psychophysiology, 49*(10), 1417-1425. https://doi.org/10.111/j.1469-8986.2012.01449.x

Moore, L. J., Vine, S. J., Wilson, M. R., & Freeman, P. (2014). Examining the antecedents of challenge and threat states: The influence of perceived required effort and support availability. *International Journal of Psychophysiology, 93*(2), 267-272. https://doi.org/10.1016/j.ijpsycho.2014.05.009

Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation, 39*(3), 125-132.

Panadero, E., Brown, G. T., & Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review, 28*(4), 803-830. https://doi.org/10.1007/s10648-015-9350-2.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74-98. https://doi.org/10.1016/j.edurev.2017.08.004

Peacock, E. J., & Wong, P. T. (1990). The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress medicine, 6*(3), 227-236.

Riener, C. R., Stefanucci, J. K., Proffitt, D. R., & Clore, G. (2011). An effect of mood on the perception of geographical slant. *Cognition and Emotion, 25*(1), 174-182.

Sassenberg, K., & Scholl, A. (2019). Linking regulatory focus and threat–challenge: Transitions between and outcomes of four motivational states. *European Review of Social Psychology*, *30*(1), 174-215

Scholl, A., Moeller, K., Scheepers, D., Nuerk, H. C., & Sassenberg, K. (2017). Physiological threat responses predict number processing. *Psychological Reseach, 81*(1). 278-288. https://doi.org/10.1007/s00426-015-0719-0

Scholl, A., de Wit, F., Ellemers, N., Fetterman, A. K., Sassenberg, K., & Scheepers, D. (2018). The burden of power: Construing power as responsibility (rather than as opportunity) alters threat-challenge responses. *Personality and Social Psychology Bulletin, 44*(7), 1024-1038. https://doi.org/10.1177/0146167218757452

Seery, M. D. (2011). Challenge or threat? Cardiovascular indexes of resilience and vulnerability to potential stress in humans. *Neuroscience & Biobehavioral Reviews, 25*(7), 1603-1610. https://doi.org/10.1016/j.neubiorev.2011.03.003

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*(2), 169-191.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In *Optimising new modes of assessment: In Search of Qualities and Standards* (pp. 55-87). Springer, Dordrecht.

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's Alpha reliability in realistic conditions: Congeneric and asymmetrical measures. *Frontiers in Psychology 7*(1), 769. https://doi.org/10.3389/fpsyg.2016.00769

Van der Schaaf, M. F., Baartman, L. K. J., Prins, F. J., Oosterbaan, A., & Schaap, H. (2011). Feedback dialogues that stimulate students' reflective thinking. *Scandinavian Journal of Educational Research, 57*(3), 227-245. https://doi.org/10.1080/00313831.2011.628693

Winstone, N.E., Hepper, E. G., & Nash, R. A. (2021). Individual differences in self-reported use of assessment feedback: The mediating role of feedback beliefs. *Educatinal Psychology, 41*(7), 844-862. https://doi-org.proxy.library.uu.nl/10.1080/01443410.2019.1693510

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisted. A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087.

Appendix A

Challenge and Threat Instructions (in Dutch)

Challenge Instruction

Probeer voor jezelf de opdrachten te zien als een uitdaging die je aan moet gaan en zal overwinnen. Zie jezelf als iemand die in staat is om deze uitdaging aan te gaan. Wij denken in ieder geval dat je meer dan in staat bent om deze uitdaging aan te gaan en overwinnen. Ons eerdere onderzoek heeft aangetoond dat de meeste deelnemers opdrachten aankunnen die lijken op diegene die jij zo gaat maken. En hoewel sommige studenten soms verwachten dat de opdrachten moeilijk zijn, ondervonden zelfs studenten met weinig ervaring dat ze heel goed in staat waren om de opdrachten uit te voeren en zij voelden zich goed over hun prestaties. Nogmaals, het klinkt misschien als een moeilijke opdracht, maar blijf jezelf eraan herinneren dat je in staat bent om het goed te doen. Doe je best!

Threat Instruction

De opdrachten kunnen moeilijk en frustrerend zijn. De meeste studenten ondervinden problemen op de kennisbasistoets bij dit type opdrachten. Het is belangrijk dat je het goed doet tijdens deze opdrachten en de opdrachten zo snel en efficiënt mogelijk uitvoert. Zowel de snelheid als de nauwkeurigheid zullen beoordeeld worden. Sommige studenten vinden het heel frustrerend, stressvol en moeilijk om zowel op snelheid en nauwkeurigheid te letten. Onthoud dus dat het belangrijk is dat je niet alleen goed, maar ook snel de vragen beantwoordt. Nogmaals, ook al is het een lastige rekentaak die sommigen stressvol vinden, probeer toch gefocust te blijven op de opdracht.

Appendix B

Challenge and Threat Questionnaire (in Dutch)

	Item	Factor
Als ik denk aan de opdrachten die ik straks ga maken, dan…	… voel ik mij een beetje bedreigd	threat
	… ben ik bang dat ik de opdrachten niet kan beheersen	threat
	… denk ik dat ik de verwachtingen van anderen niet kan waarmaken	threat
	… ben ik er helemaal klaar voor om te beginnen	challenge
	… voel ik mij erg gestimuleerd	challenge
	… ben ik er zeker van dat ik het ga beheersen	challenge
Hoe ervaar je de situatie op dit moment?	Overweldigend	threat
	Bedreigend	threat
	Stimulerend	challenge
	Als een uitdaging die ik aankan	challenge
	Intimiderend	threat
	Enthousiasmerend	challenge