

The Achievement Emotions Questionnaire: Validation And Implementation For Undergraduate Physics Practicals

Aesha Bhansali^a, Manjula D. Sharma^a

Corresponding author: Aesha Bhansali (abha0871@uni.sydney.edu.au)

^aSchool of Physics, The University of Sydney, Sydney NSW 2006, Australia

Keywords: Achievement Emotions Questionnaire, Confirmatory Factor Analysis, measuring emotions, physics education

Abstract

Physics is a discipline associated with diverse emotions; some enjoy it, others don't. Yet students' emotional engagement, which is important for students' continued interest and learning outcomes, is under researched. This study adapts the Achievement Emotions Questionnaire (AEQ) to measure the emotions of students with first year physics undergraduate practicals. The aims of this research are to validate the AEQ in our context and to probe students' emotions towards two practicals; the control which is of standard format and the intervention which incorporates colour and historical aspects seeking to produce more positive emotions. Confirmatory Factor Analysis and descriptive statistics conducted with a sample of 320 students confirm the reliability and internal validity of the adapted AEQ (AEQ-PhysicsPrac) for the purposes of this study. Differences in emotions between the control and intervention are detected indicating that the AEQ-PhysicsPrac has utility in physics education.

Introduction

Researchers have referred to student engagement as a resource that 'once established, builds on itself, thereby contributing to increased improvements in more distal outcomes of interest' (Fredricks, Blumenfeld, & Paris, 2004). Students' emotional engagement is required for continued interest and considerable success in a subject. It has been shown that emotions influence students' learning, motivation, and achievement outcomes (Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008; Pekrun, 2006; Robinson, Lloyd, & Rowe, 2008; Svanum & Bigatti, 2009; Weiner, 2010). For the majority of students, positive emotions can be beneficial and negative emotions can be detrimental for the academic learning (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). The question then arises, how can positive emotions be invoked when it comes to subjects such as physics which are often associated with particularly diverse emotions and can be difficult for students to find interesting (Pekrun, Goetz, Titz, & Perry, 2002; Williams, Stanisstreet, Spall, Boyes, & Dickson, 2003). Specifically, for females, emotions they feel are pertinent to the way they perceive physics (Gokalp, 2018).

The interpretation and measurement of emotions is quite a debated and difficult process (Scherer, 2005). Researchers differ in their understandings and definitions of 'emotion' as well as in its measurement (Lewis & Haviland-Jones, 2000; Gendron & Barrett, 2009; Linnenbrink-Garcia & Pekrun, 2011). The Achievement Emotions Questionnaire (AEQ) developed by Pekrun, Goetz, & Perry (2005), in line with Scherer's model (2009), sees emotions as interrelated psychological processes. Emotions associated with the achievement activity as well as with achievement outcomes are considered as achievement emotions. Pekrun's control-value theory (2006) organises emotions according to three dimensions:

valence (positive vs negative), activity level (activating vs deactivating), and object focus (activity vs outcome). The AEQ measures and represents nine emotions separating them in four quadrants; positive activating emotions are enjoyment, hope and pride; positive deactivating emotion is relief; negative activating emotions are anger, anxiety and shame; and negative deactivating emotions are hopelessness and boredom. These emotions are measured in three contexts: class attendance, learning and taking tests. The three positive activating emotions are thought to promote both motivation and self-regulation; thus positively affecting students' academic performance. The two negative deactivating emotions reduce motivation, implying negative effects on performance. The effect of the positive deactivating and negative activating emotions are more complex, mostly affecting achievement performance negatively (Pekrun, 2006).

The AEQ, initially developed for college students, has been used in multiple languages and contexts (Pekrun, Goetz, & Perry, 2005). Lichtenfeld, Pekrun, Stupnisky, Reiss, & Murayama (2012) constructed the AEQ—Elementary School which assesses three emotions (enjoyment, anxiety and boredom). Pekrun, Goetz, & Frenzel (2005) developed the AEQ—Mathematics (AEQ-M) which measures seven emotions (enjoyment, pride, anger, anxiety, shame, hopelessness, boredom) for pre-adolescents. The AEQ has been tested and trialled in sufficient studies to suggest that it is robust and a good starting point for investigating emotions in more subjects and in different contexts.

Looking at physics, the subject of this study, there is research on students learning in undergraduate laboratories ranging from metacognition (Kung & Linder, 2007) to perceptions of learning experiences (Barrie et al., 2015). Furthermore, a range of instruments are available for measuring other affective and attitudinal aspects of studying physics, for example, Physics Self-Efficacy Questionnaire (Lindstrøm & Sharma, 2011), Physics Goal Orientation survey (Lindstrøm & Sharma, 2010) and the Maryland Physics Expectations (MPEX) Survey (Redish, Saul, & Steinberg, 1998). Of note, is work done three decades ago by Watson, Clark, & Tellegen (1988) developing and validating the Positive Affect and Negative Affect Schedule (PANAS). More recently, in a qualitative study involving 19 first year physics students, Kahu (2014) found that students' emotions are important for explaining the links between their interest and engagement with physics. Lehtamo, Juuti, Inkinen, & Lavonen (2018) in researching situational emotions with 36 school students found that reduced stress is a significant predictor for retaining students in physics. Test anxiety has received some attention, see for example Weiner (2010). A Physics Anxiety Rating Scale (PARS) has been developed for both school and university students' by Sahin, Caliskan, & Dilek (2015). However, the research on achievement emotions when undertaking physics is limited, particularly as defined by the AEQ. The AEQ not only assesses the individual emotions but provides an organised and hierarchical structure for correlating emotions. It enables us to measure emotions in a systematic manner. The AEQ was suitable for our research purposes because it is a valid tool, has been implemented for various contexts, and contains the emotions relevant to us.

Current Research

The goal of this research was to adapt, validate, and implement the AEQ to measure undergraduate students' achievement emotions towards physics practicals. Our focus was on those emotions that occur in our context, allowing us to have a relatively short instrument measuring discrete emotions addressed by the AEQ. The modified AEQ, called the AEQ-PhysicsPrac, has been trialled in an authentic teaching and learning context within the School of Physics at a research-intensive metropolitan university in Australia. The undergraduate

physics subjects have a laboratory component with the first-year subjects containing practicals designed to be carried out within three hours. Two of these three-hour practicals are used in this study.

This study has created a novel approach for distinguishing the two practicals with the intent of discerning differences in achievement emotions using the AEQ-PhysicsPrac. The first practical, the control, is a long-standing practical which was taught with no modifications. The second practical, the intervention utilised a ‘science story’ with history and colour in the student notes to strive for achievement emotions.

The representation of content in student notes is important for engaging students and capturing their interest. Stinner (1995) demonstrated the positive impact of presenting content as ‘science stories’ which include the history of science. It also mentions that appropriately designed contexts attract students’ interest and creates great motivation to learn science. Students readily relate to the human aspiration and thought process behind particular discoveries and science understandings. Furthermore, Höttecke (2012) and Monk & Osborne (1997) have demonstrated that science stories lead to the development of scientific thinking skills. Together with science stories, colour has long been associated with emotion. Elliot (2015) noted how Goethe (1810) related different colours with different emotions, and the further elaboration of this theory by Goldstein (1942). In this study, the student notes for the intervention feature science stories and include colour.

Method

The intervention and the control

The intervention was a newly developed practical on ‘heat and thermodynamics’ focusing on measurements of temperature and the feeling of hot and cold, measuring the heat conduction of different metals using real time data collection, and modelling heat conduction on an Excel spreadsheet. The colourful historical science story is presented on the first page of the student notes in the form of a cyclic depiction of how human’s understanding of ‘heat’ has evolved over time and some of the images reappear as thumbnails in the student notes. The rest of the content in the student notes is black and white.

The control was an established practical on ‘ultrasound’ which focused on measuring the speed of sound in air using ultrasound waves, measuring ultrasound wave properties using the signals displayed on a digital oscilloscope, and modelling the relationship between amplitude/intensity and distance for ultrasonic waves. The first few pages contained the theoretical background without any historical context or science story. The student notes were entirely in black and white.

The common features were that both the practicals were for the same cohort of students. For both the practicals, the lab notes presented to the students had the same length in terms of number of pages and had comparable structures. They were of equal complexity in terms of subject matter and skills, as well as measurement and analysis requirements. Both of the practicals were intended to be completed in three hours.

Sample and procedure

The cohort are predominantly straight from school, 18 to 21-year-old, with around 25% females. Three-hour practical sessions run multiple times in the week. Each student is assigned to one practical session for one topic per week. Four practicals for four topics,

named ‘microwaves’, ‘vibrating wires’, ‘heat and thermodynamics’, and ‘ultrasound’ were available over three weeks, with students undertaking three of the four practicals. Students work in teams of three, seeking assistance from the tutors who facilitate student activities and learning.

For the control, 136 students returned surveys out of 193 present; 71% response. For the intervention, 184 students returned surveys out of 213 present; 86% response. So, this was a non-biased and non-random sampling. The study has approval from the institutional Human Ethics Committee.

The AEQ-PhysicsPrac: Development, data collection and analysis

Survey selection and item development

A literature search was undertaken to specifically find surveys that could measure aspects of students’ emotional engagement. The AEQ emerged as the most pertinent for ~~three~~ these reasons: it has been implemented with sound statistical results; it has all the emotions relevant to us. Mathematics is allied to physics, likely to arouse similar emotions. Therefore, we decided to use items from AEQ-M (Pekrun et al., 2005) as the basis of our AEQ-PhysicsPrac survey. The emotions selected were pride, enjoyment, anger, anxiety, hopelessness and boredom. These emotions cover all three relevant quadrants of the control-value theory (Pekrun, 2006), considering their valence (positive vs negative) and activity level (activating vs deactivating).

Table 1 shows the categories of changes. In the first category, T1, the tense of the item was changed to past tense to capture students’ retrospective emotions that they felt during the laboratory, and/or made specific reference to practical work which is referred to as experiment in physics. The next category, T2, included adding or removing qualifiers which introduces relative subjectivity without significantly changing the meaning. Category T3 simplifies double barrelled items, sometimes inserting an alternative word. We considered several emotion wheels that originate from the same theoretical background as The Junto Emotion Wheel (Chadha, 2020), one by Plutchik (1980), and the Geneva Emotion Wheel (Scherer, Shuman, Fontaine, & Soriano, 2013) to identify alternative words which come from the natural language. Category T4 created new items avoiding extreme emotions. A total of 19 Likert scale items were selected, see Appendix for the AEQ-PhysicsPrac.

Table 1: Categories of adaptation of the items for the AEQ-PhysicsPrac

| Category | Explanation | Example |
|----------|--|--|
| T1 | The tense was changed and/or explicit reference was made to physics practicals/experiments. | ‘I am happy that I could cope with this experiment.’ |
| T2 | Words which introduce relative subjectivity were removed, while maintaining the sentiment of the item e.g. ‘fairly’. | ‘After a math test, I am fairly annoyed’ was changed to ‘I felt annoyed by this experiment.’ |

| | | |
|----|--|--|
| T3 | Double barrelled items were simplified. In some cases suitable words from the Geneva Emotion Wheel were used. | ‘Because I take pride in my accomplishments in mathematics, I am motivated to continue’ was changed to ‘I felt elated by my accomplishments during this experiment.’ |
| T4 | New items were constructed remaining congruent with the original AEQ items. The items avoid any extreme sentiment. | I resented doing this experiment |

Administration, data collection and analysis

Tutors provide a brief introduction prior to students starting their practicals. At this point, the AEQ-PhysicsPrac survey was introduced, the purpose of the study read from a script, emphasising that completion was voluntary and that marks would not be affected. The surveys were administered after the students completed their practical and collected as students left the laboratory. The surveys were kept and coded in bundles for each session so that data could be examined for variations between sessions.

The data were curated by removing the responses which had three or more items left blank. The data from the 320 responses were entered into EXCEL. The Likert scale was interpreted as: Strongly Agree=5, Agree=4, Neutral=3, Disagree=2, Strongly Disagree=1. The data were then exported into SPSS Version 24. Extensive data exploration was carried out, including checking the distribution of individual items for normality. Exploratory Factor Analysis produced two factors; one containing items from positive emotions of pride and enjoyment, and the other containing items from negative emotions of anger, anxiety, hopelessness and boredom as theorised by Goetz, Pekrun, Hall, & Haag (2006), and Russel (1980). Factor loadings were greater than 0.4 and were accepted as per Field (2000) and reliability of each factor ascertained by Cronbach’s alpha was 0.75, indicating acceptable reliability (Pekrun et al., 2011; Peixoto, Mata, Monteiro, Sanches, & Pekrun, 2015). The assumptions were adequate; Kaiser-Meyer-Olkin (KMO) Sampling Adequacy was .69, Bartlett’s sphericity test for correlations had significance $p=.00$, the determinant was .06. These satisfy the recommended criteria. (Dziuban & Shirkey, 1974; Lindstrøm & Sharma, 2010). The inter-item correlations were in the range 0.3-0.8, medium to high, indicating that multicollinearity was not an issue (Cohen, 1988; Peixoto et al., 2015; Sharma, Stewart, Wilson & Gokalp, 2013).

The Exploratory Factor Analysis indicated that the data were appropriate for Confirmatory Factor Analysis (CFA) to test the internal structure of the emotions. CFA was carried out using Amos Version 24.0. Analysis was conducted using maximum-likelihood estimation. The parameters and associated criteria used to assess the goodness of fit of the model are shown in Table 2.

Table 2: Confirmatory Factor Analysis: Parameters, Criteria, and Reference

| Parameter | Criteria | Reference |
|-----------|----------|-----------|
|-----------|----------|-----------|

| | | |
|---|-------------------------|---|
| Comparative fit index (CFI) | good fit at $\geq .90$ | Hooper, Coughlan, & Mullen, 2008; Sharma et al., 2013 |
| Root Mean Square Error of Approximation (RMSEA) | good fit at $< .08$ | Hooper, Coughlan, & Mullen, 2008; Sharma et al., 2013 |
| Relative χ^2 (χ^2/df) | reasonable fit at < 3 | Pekrun et al., 2011; Sharma et al., 2013; Marsh & Hocevar, 1985 |
| Average Variance Extracted(AVE) | Should be $> .4$ | Huang, Wang, Wu, and Wang, 2017 |
| Composite Reliability | Should be $> .6$ | Sharma et al., 2013 |

Means and correlations were computed for each emotion using SPSS. In order to compare the scores Independent Samples *t*-test was conducted.

Results

Validity and Reliability of the AEQ-PhysicsPrac

Component structure of emotions

Figure 1 shows the model containing six interrelated factors in a multi-dimensional structure as per Peixoto et al. (2015). The factor loadings, the left most digits in Figure 1, are > 0.6 for all, except two emotion items which had reasonable factor loadings of 0.52 and 0.56. The items load appropriately on their respective emotions; hence all items were retained. The latent factor correlations show positive relationship between emotions of the same valence and negative relationship between the emotions of opposite valence as in Pekrun et al. (2011).

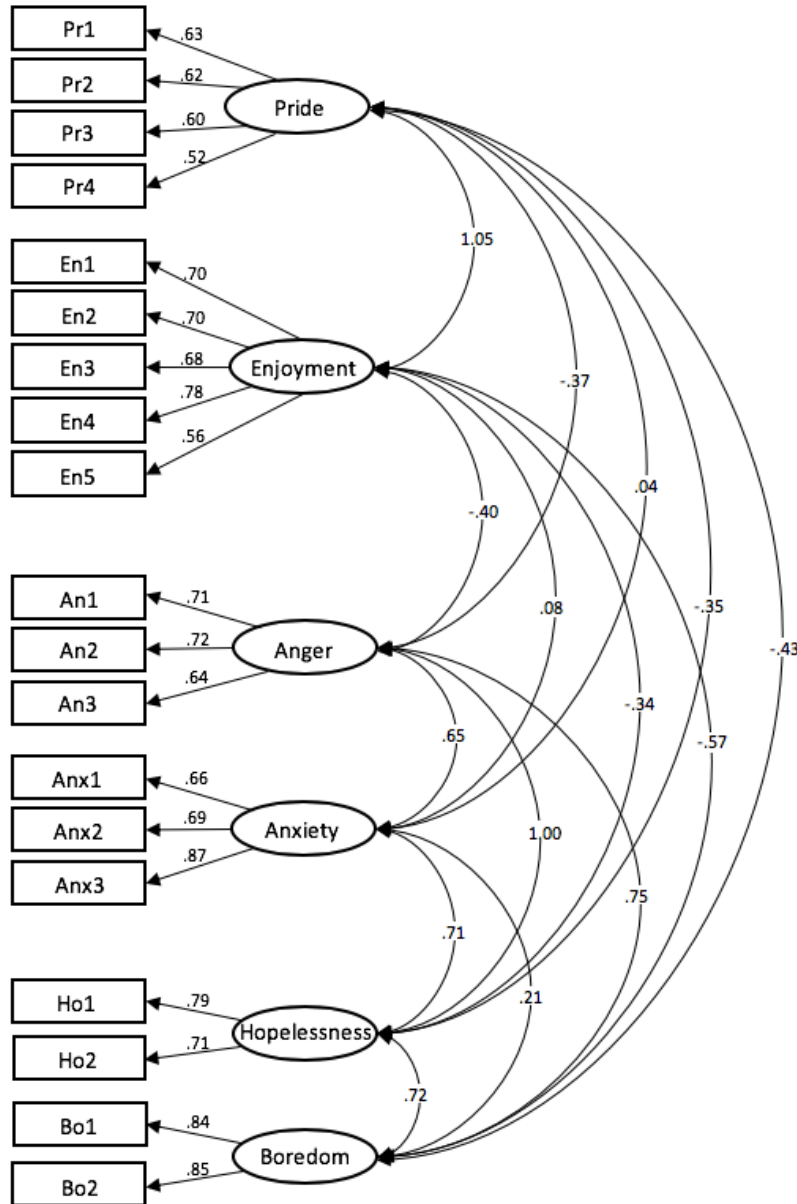


Figure 1: Six-factor model for physics experimental activity related emotions. Pr, pride; En, enjoyment; An, anger; Anx, anxiety; Ho, hopelessness; Bo, boredom. Each box represents a separate item. (Read from left to right the digits represent factor loadings, and latent factor correlations.)

The acceptable goodness-of-fit indexes validates the six interrelated factors in a multi-dimensional model of the AEQ-PhysicsPrac survey used in our study. Since reliability scores were >0.6 for all emotion scales, The AEQ-PhysicsPrac model is seen to be reliable. Furthermore, the model leads to the idea that the emotions are differentiated and discrete, hence can be probed separately, compared across treatments and their relationships investigated.

Table 3: Validity and Reliability measures of the model

| Validity Measures - Goodness-of-fit parameters of the model | Score |
|--|------------------------------|
| CFI | good fit at .902 |
| Root Mean Square Error of Approximation (RMSEA) | good fit at .077 |
| Relative χ^2 (χ^2/df) | reasonable fit at 2.89 |
| Average Variance Extracted (AVE) | Acceptable at .5 |
| Reliability measures for each Emotion | Composite Reliability |
| Pride | 0.68 |
| Enjoyment | 0.82 |
| Anger | 0.73 |
| Anxiety | 0.79 |
| Hopelessness | 0.72 |
| Boredom | 0.83 |

Correlational analysis of the emotions

Table 4 shows that for both the intervention and control, there are strong positive correlations between pride and enjoyment; and between anger, boredom and hopelessness. In addition, there are moderate negative correlations between those positive and negative emotions. Anxiety is seen to be having strong positive correlation with anger and hopelessness, yet it does not seem to be correlated with positive emotions and with boredom. This exception is seen in earlier research by Pekrun et al. (2011). So, for AEQ-PhysicsPrac the like valenced correlations, as well as opposite valence correlations are as per the AEQ model. The correlations show that the emotions are clearly separable. The strongest correlations were found between pride and enjoyment, and between anger and hopelessness, as expected for the like valenced trait-like emotions (Pekrun et al, 2004). The strongest negative correlation was found to be between enjoyment and boredom. It should be noted that enjoyment is positive activating and boredom is negative deactivating emotion, thus our results indicate that not only valence, but the activity level also matters while considering relationships of emotions.

Table 4: Correlations between the emotions for the intervention (above the diagonal) and control (below the diagonal in Italics)

| | Pride | Enjoyment | Anger | Anxiety | Hopelessness | Boredom |
|--------------|----------------|----------------|----------------|---------------|----------------|----------------|
| Pride | - | .654** | -.268** | -.028 | -.160* | -.287** |
| Enjoyment | <i>.834**</i> | - | <i>-.293**</i> | <i>.075</i> | <i>-.252**</i> | <i>-.474**</i> |
| Anger | <i>-.164</i> | <i>-.188*</i> | - | <i>.537**</i> | <i>.711**</i> | <i>.576**</i> |
| Anxiety | <i>.083</i> | <i>.070</i> | <i>.472**</i> | - | <i>.547**</i> | <i>.139</i> |
| Hopelessness | <i>-.243**</i> | <i>-.227**</i> | <i>.707**</i> | <i>.517**</i> | - | <i>.541**</i> |
| Boredom | <i>-.284**</i> | <i>-.375**</i> | <i>.502**</i> | <i>.171*</i> | <i>.522**</i> | - |

** . Significant at the 0.01 level (2-tailed).

* . Significant at the 0.05 level (2-tailed).

Implementation

Comparing the means for intervention and control

Table 5 shows statistics relevant for comparing the intervention with the control. Independent samples *t*-test shows statistically significant differences between the intervention and control for five of the six emotions. For positive emotions, pride and enjoyment, the means are statistically significantly higher for the intervention when compared with the control; For negative emotions, anger, hopelessness, and boredom, the means are statistically significantly lower for the intervention when compared with the control. There is no significant difference for anxiety.

Table 5: Comparing intervention with control for each emotion: descriptive statistics and *t*-statistics

| Emotion | Mean (SD) | | Range | <i>t</i> | p |
|--------------|-----------------------------|------------------------|-------|----------|------|
| | <i>Intervention</i> (n=187) | <i>Control</i> (n=133) | | | |
| Pride | 15 (2) | 14 (3) | 4-20 | 2.25 | <.05 |
| Enjoyment | 18 (3) | 17 (4) | 5-25 | 3.66 | <.01 |
| Anger | 8 (2) | 9 (2) | 3-15 | -4.08 | <.01 |
| Anxiety | 7 (3) | 8 (3) | 2-15 | -1.85 | >.05 |
| Hopelessness | 5 (2) | 6 (2) | 1-10 | -3.88 | <.01 |
| Boredom | 5 (2) | 6 (2) | 1-10 | -3.55 | <.01 |

Trends in student responses to each item

Here we seek to discern patterns in student responses for each item for the intervention and control. We combine the Likert % responses: Strongly Agree is combined with Agree to give ‘% agreement’, neutral stays as ‘% neutral’, Strongly Disagree is combined with Disagree to give ‘% disagreement’. Table 6 shows these data for the intervention and control for all the items, sorted according to valence, i.e. positive emotions followed by negative emotion. Four

points are noteworthy. First, for both the intervention and control, the % agreement is higher than % disagreement with the positive valence items. And vice versa is true for the negative valence items.–Students’ emotional responses indicate both practicals are positive learning experiences. Second, the % neutral responses are pretty similar across the items.–Third, there is higher % agreement with positive valence items for the intervention than for the control.–So, the intervention attracts more positive emotions compared to the control. Fourth, now considering the negative emotions items, we get higher % agreement for control than for the intervention. Specifically, hopelessness and boredom show double the % agreement for control compared to the intervention. So, the control attracts more negative emotions compared to the intervention.

Table 6: Percentage Responses for AEQ-PhysicsPrac for Intervention and Control. Combined as: ‘% agreement’ (SA/A), ‘% neutral’ (N) and ‘% disagreement’ (SD/D)

| Item | Summary | % agreement | | % neutral | | % disagreement | |
|-----------------------|------------------------|--------------|---------|--------------|---------|----------------|---------|
| | | Intervention | Control | Intervention | Control | Intervention | Control |
| Positive Items | | | | | | | |
| Pr1 | Satisfied | 80 | 69 | 16 | 23 | 2 | 8 |
| Enj5 | Cope with | 72 | 62 | 24 | 32 | 4 | 6 |
| Pr4 | Important contribution | 72 | 59 | 24 | 29 | 4 | 12 |
| Enj3 | Efforts paid off | 66 | 56 | 29 | 35 | 4 | 9 |
| Enj1 | Enjoyable challenges | 63 | 50 | 32 | 32 | 6 | 17 |
| Enj4 | Enjoyed | 62 | 39 | 29 | 36 | 10 | 25 |
| Pr3 | Keep up | 59 | 52 | 34 | 36 | 7 | 12 |
| Enj2 | Exciting | 45 | 32 | 35 | 38 | 19 | 30 |
| Pr2 | Elated | 42 | 37 | 46 | 45 | 10 | 18 |
| Negative Items | | | | | | | |
| An1 | Annoyed | 28 | 32 | 24 | 27 | 48 | 41 |
| An2 | Irritating | 26 | 40 | 28 | 27 | 46 | 32 |
| An3 | Resented | 19 | 38 | 30 | 32 | 50 | 30 |
| Anx1 | Scared | 19 | 33 | 35 | 35 | 46 | 32 |
| Anx2 | Nervous | 18 | 35 | 37 | 36 | 44 | 29 |
| Anx3 | Panicky | 18 | 38 | 38 | 36 | 42 | 26 |
| Ho1 | Give up | 17 | 23 | 24 | 25 | 58 | 52 |
| Ho2 | Resigned | 15 | 19 | 25 | 22 | 61 | 59 |
| Bo1 | Dull | 12 | 23 | 31 | 42 | 57 | 35 |
| Bo2 | Bored | 10 | 17 | 24 | 26 | 66 | 56 |

Discussion

This study set out to develop, validate and implement a survey for measuring emotions in physics labs. Our findings indicate that the AEQ-PhysicsPrac is a valid and reliable tool, using the criteria of Peixoto et al. (2015). It supports the model that the emotions are differentiated, discrete, and can be probed separately. When investigating relationships, we find that the positive emotions pride and enjoyment are highly positively correlated with each other, and negatively correlated with negative emotions anger, hopelessness and boredom. Negative emotions are mostly positively correlated with each other, except for anxiety which has also been noted by Pekrun et al. (2011) to be an exception. It is noteworthy to mention that earlier studies have found anxiety to have complex effects on students’, sometimes affecting them positively, sometimes negatively and sometimes not affecting them at all

(Pekrun, 2009). When comparing the intervention which uses ‘science story’ and colour with the control which is a standard practical, we find that the intervention attracts more positive emotion and less negative emotion. This demonstrates that our survey can be administered and utilised by practitioners. We note that despite sporadic attempts at highlighting importance of colour and historical stories, see for example Elliot (2015) and Höttecke (2012), such strategies are not widely used for teaching science. Our study suggests that this be actively pursued by teachers and practitioners.

The AEQ-PhysicsPrac can be a useful tool for further research in assessing and understanding undergraduate students’ emotions towards physics in laboratory situations. AEQ-PhysicsPrac also opens up the possibility of obtaining different profiles of emotions experienced by students for different practicals. Thus, the learning resources can be designed for a better student engagement. Some limitations need to be considered. There is a large and complex dimension arising from what the student brings to learning. Factors here will include: the expertise of the learner, the sophistication of their approach to learning (maturity), learner expectations and prior experience of similar learning tasks. As this adaptation was made with a sample of the University of Sydney first year students in physics, more research is needed to test the AEQ-PhysicsPrac over time, with different cohorts and across different practicals.

Acknowledgements

We acknowledge earlier work on the topic. We would like to thank students, tutors, Alexander Samarian and the technical team who participated in the research and supported the data collection. We thank the School of Physics and the SUPER group at the University of Sydney.

References

- Barrie, S., Bucat, R., Buntine, M., Burke Da Silva, K., Crisp, G., George, A., ... Yeung, A. (2015). Development, Evaluation and Use of a Student Experience Survey in Undergraduate Science Laboratories: The Advancing Science by Enhancing Learning in the Laboratory Student Laboratory Learning Experience Survey. *International Journal of Science Education*, 37(11), 1–20.
- Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2), 230–258.
- Browne, M.W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K.A. & Long, J.S. [Eds.] *Testing structural equation models*. Newbury Park, CA: Sage, 136–162.
- Chadha, R. (2020). The Junto Emotion Wheel: Why and How We Use It. Retrieved from <https://www.thejuntoinstitute.com/blog/the-junto-emotion-wheel-why-and-how-we-use-it>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dziuban, C., & Shirkey, E. (1974). When is a Correlation Matrix Appropriate for Factor Analysis? Some Decision Rules. *Psychological Bulletin*, 81(6), 358–361.
- Elliot, A. J. (2015). Color and psychological functioning: a review of theoretical and empirical work. *Frontiers in Psychology*, 6(368).
- Field, A. (2000). *Discovering Statistics Using SPSS for Windows*. London: Sage Publications Ltd.
- Fredricks, J., Blumenfeld, P., & Paris, A. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, 74 (1), 59-109.
- Gendron, M., & Barrett, L. (2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review*, 1(4), 316-339.
- Goethe, W. (1810). *Theory of Colors*. London: Frank Cass.
- Goetz, T., Pekrun, R., Hall, N.C., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: antecedents and domain specificity of students' affect in the context of Latin instruction. *The British journal of educational psychology*, 76(2), 289-308.

- Gokalp, M. (2018). Multidimensional perceptions of physics. *International Journal of Innovation in Science and Mathematics Education*, 26 (1), 27-45.
- Goldstein, K. (1942). Some experimental observations concerning the influence of colors on the function of the organism. *Occupational Therapy*, 21, 147–151.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods*, 6, 53-60.
- Höttecke, D. (2012). HIPST—History and Philosophy in Science Teaching: A European Project (Book review). *Science & Education*, 21(9), 1229–1232.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Huang, C., Wang, Y., Wu, T., & Wang, P. (2013). An Empirical Analysis of the Antecedents and Performance Consequences of Using the Moodle Platform. *International Journal of Information and Education Technology*, 3(2), 217-221.
- Kahu, E. (2014). Increasing the emotional engagement of first year mature-aged distance students: Interest and belonging. *International Journal of the First Year in Higher Education*, 5(2).
- Kline, R. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.
- Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the Effects of Student Engagement on First-Year College Grades and Persistence. *The Journal of Higher Education*, 79(5), 540-563.
- Kung, R. L., & Linder, C. (2007). Metacognitive activity in the physics student laboratory: is increased metacognition necessarily better?. *Metacognition and Learning*, 2(1), 41–56.
- Lehtamo, S., Juuti, K., Inkinen, J., & Lavonen, J. (2018). Connection between academic emotions in situ and retention in the physics track: applying experience sampling method. *International Journal of STEM Education*, 5(1), 1–6.
- Lewis, M., & Haviland-Jones, J. M. (Eds.). (2000). *Handbook of emotions* (2nd edition). New York, NY: Guilford.
- Lichtenfeld, S., Pekrun, R., Stupnisky, R. H., Reiss, K., & Murayama, K. (2012). Measuring students' emotions in the early years: the achievement emotions questionnaire-elementary school (AEQ-ES). *Learning and Individual Differences*, 22(2), 190-201.
- Lindstrøm, C. A., & Sharma, M. D. (2011). Self-Efficacy of First Year University Physics Students: Do Gender and Prior Formal Instruction in Physics Matter?. *International Journal of Innovation in Science and Mathematics Education*, 19(192), 1-19.
- Lindstrøm, C., & Sharma, M. (2010). *International Journal of Innovation in Science and Mathematics Education*, 18(2), 10-20.
- Linnenbrink-Garcia, L., & Pekrun, R. (2011). Students' Emotions and Academic Engagement: Introduction to the Special Issue. *Contemporary Educational Psychology*, 36(1), 1–3.
- Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, 97(3), 562-582.
- Monk, M., & Osborne, J. (1997). Placing the history and philosophy of science on the curriculum: A model for the development of pedagogy. *Science Education*, 81(4), 405-424.
- Peixoto, F., Mata, L., Monteiro, V., Sanches, C., & Pekrun, R. (2015). The Achievement Emotions Questionnaire: Validation for Pre-Adolescent Students. *European Journal of Developmental Psychology*, 12(4), 1–10.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315 –341.
- Pekrun, R. (2009). Emotions at school. In K. R. Wenzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (p. 575–604). Routledge/Taylor & Francis Group.
- Pekrun, R., Goetz, T., & Frenzel, A. C. (2005). Achievement Emotions Questionnaire—Mathematics (AEQ-M). User's manual. Munich: University of Munich, Department of Psychology.
- Pekrun, R., Goetz, T., & Perry, R. P. (2005). Achievement Emotions Questionnaire (AEQ)—User's manual. Munich: University of Munich, Department of Psychology.
- Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36–48.
- Pekrun, R., Goetz, T., Titz, W., & Perry R. P. (2002). Academic Emotions in Students' Self-Regulated Learning and Achievement: A Program of Qualitative and Quantitative Research. *Educational Psychologist*, 37(2), 91-105.
- Plutchik, R. (1980). A general pschoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.),

Emotion: Theory, research, and experience: Vol. 1. Theories of emotion (pp. 3-33). New York: Academic.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Educational Administration Quarterly*, 44(5), 635–674.

Russel, A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.

Sahin, M., Caliskan, S., & Dilek, U. (2015). Development and Validation of the Physics Anxiety Rating Scale. *International Journal of Environmental and Science Education*, 10(2), 183-200.

Scherer, K. (2005). What Are Emotions? And How Can They Be Measured?. *Social Science Information*, 44(4), 695–729.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, 23, 1307-1351.

Scherer, K., Shuman, V., Fontaine, J., & Soriano, C. (2013). The GRID meets the Wheel: Assessing emotional feeling via self-report. In *Components of Emotional Meaning*.

Sharma, M., Stewart, C., Wilson, R., & Gokalp, M. (2013). Student Approaches to Learning in Physics - Validity and Exploration Using Adapted SPQ. *International Journal of Environmental & Science Education (IJESE)*, 8(2), 241-253.

Stinner, A. (1995). Contextual settings, science stories, and large context problems: Toward a more humanistic science education. *Science Education*, 79(5), 555-581.

Svanum, S., & Bigatti, S. (2009). Academic course engagement during one semester forecasts college success: Engaged students are more likely to earn a degree, do it faster, and do it better. *Journal of College Student Development*, 50(1), 120-132.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.

Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist*, 45, 28–36.

Williams, C., Stanisstreet, M., Spall, K., Boyes, E., & Dickson, D. (2003). Why aren't secondary students interested in physics?. *Physics Education*, 38, 324.