Australian Journal of Educational & Developmental Psychology. Vol 14, 2014, pp. 93-116

Sustaining young people's enrolment intentions in relation to physics: Development and validation of a tool

Jessy Abraham[1] & Katrina Barker,

University of Western Sydney

ABSTRACT

Currently there is a need for measures to examine the issue of sustaining students' enrolment intentions over an extended period of study in physics, a subject which is generally perceived as hard and demanding by students. This paper addresses this gap in research by describing the development and the assessment of psychometric properties of the Physics Motivation Questionnaire, which examines the predictive relations among senior secondary physics students' achievement motivation, sustained engagement and sustained enrolment intentions. The theoretical framework of the instrument largely draws on the Expectancy-Value theory of achievement motivation and the latent variables are assessed through six measures. Data shows that the Physics Motivation Questionnaire is a theoretically sound and psychometrically valid instrument which has utility in examining physics at a topic-specific level. The questionnaire makes a unique contribution to the physics enrolment literature and has significant implications for educational practitioners. These implications are discussed in the context of the findings.

Key words: Physics education, Motivation, Engagement, Expectancy-Value theory.

INTRODUCTION

Falling enrolment trends and the relatively lower participation of females are two major concerns that are prominent in Australian and international physics education research literature (e.g. Ainley, Kos, & Nicholas, 2008; Barnes, 1999; Lyons & Quinn, 2010). The persistent pattern of declining student participation has significant consequences. Low enrolment in physics could lead to a shortage of physicists, physics teachers and qualified persons in physics related careers. Also, the increasing alienation of females from physics could exacerbate the gender-differentiated pattern within physics related careers in future years (DEST, 2006).

The need for the development of a new instrument

To explain the poor student participation in physics, researchers examine when and why students, particularly females, are leaving physics. The majority of physics enrolment studies focus on the critical exit point from physics that is the transition from junior secondary schools to senior high school and, instruments that examine students' enrolment motivation are readily available (e.g. Barnes, 1999; Eccles, Barber, Updegraff, & O'Brien, 1998; Wigfield & Eccles, 2000). At this point, students are simply choosing to do or not to do physics in senior high school, although they have not

[1] Contact
Jessy Abraham,
School of Education,
University of Western Sydney,
Locked Bag 1797, Penrith, NSW 2751, Australia
Phone: +61(2) 97726342
Fax: +61(2) 97726738
Email: jessy.abraham@uws.edu.au

studied this subject before, but just completed a preparatory course in general sciences. Thus, these enrolment studies mainly examine the factors that contribute to students' intention to enrol in an *imagined* physics curriculum (Vickers & Ha, 2007) at the senior secondary level. Nevertheless, once students have enrolled in an elective physics course at senior secondary level, their sustained intentions to enrol in further physics will be shaped by their actual experience of the *enacted* physics curriculum (Vickers & Ha, 2007). The growing body of physics enrolment literature fails to examine this exit point of senior students. This could be attributed to the fact that in most countries (notably in the United States) there is only one year of physics in the senior high school program, normally in grade 12. Thus, while the declining trend in physics enrolment and the widening gender imbalance in physics participation have been explored widely, the retention of students who are enrolled in senior year physics and their decisions to continue with this subject for the final years remains largely unexplored.

If the aim is to determine the motivational factors that predict students' sustained engagement and enrolment intentions in relation to physics and the strength of the relationships among these factors after an extended period of study of the subject, then the Australian senior secondary school curricular structure offers a unique opportunity. This is possible because Australian high school students may study physics in Year 11, and then have the option to discontinue or continue with physics into Year 12 provided they retain the minimum units of study for their final year. Furthermore, this curricular structure provides an opportunity to examine whether there is a gender gap in physics in the senior school context. Given the underrepresentation of females in physics classes, this issue could be of special research interest. Thus, there exists a need and an opportunity to shift the focus from exclusively examining students' initial enrolment intentions in physics to examining sustained enrolment intentions in the subject once they have started studying physics. Consequently, there is a requirement for an instrument which can measure sustained enrolment intentions of students in physics. The Physics Motivation Questionnaire (PMQ) is a self-reporting instrument that was developed for this purpose.

The purpose of this paper is to report on the construction and validation of PMQ which was designed to measure achievement motivation, sustained engagement and enrolment intentions of senior secondary students in New South Wales (NSW) in relation to the enacted physics curriculum. Understanding the motivations and intentions of students studying physics can inform educators of what factors influence young people to continue with studying this important subject. Factors that can boost motivation and increase a student's commitment to studying physics will positively affect enrolments in this critical area. Furthermore, it is vital to determine whether there are gender difference in physics motivation and enrolment intentions so that support for students is targeted appropriately.

Furthermore, if the aim is to provide practical support to educators, then a fine-grained exploration of these factors as relevant to various topics in the relevant curriculum is the essential first step. Physics curricula at higher education levels consist of various topics with varying characteristics. For example, some topics are descriptive, while others are more problem oriented; some are theory oriented, while some have more practical utility in everyday life. However, previous studies have tended to measure motivational patterns in physics as a school subject as a whole, treating it as a single unit of study (e.g. Barnes, 1999; Eccles et al., 1998; Wigfield & Eccles 2000) and, therefore the current measures that examine the motivation, engagement and choice in relation to physics are domain specific. PMQ was developed acknowledging the possibility of varying levels of motivational predictors and the varying levels of further enrolment intentions measured at the completion of the individual topics in the physics curriculum. Hence the psychometric properties of PMQ were examined at a more sensitive topic-specific level in a multi-occasional study among the senior secondary students in their first year of electing to study physics.

Theoretical foundation of PMQ

PMQ largely draws on the Expectancy-Value (EV) theory of achievement motivation. EV theory was developed to explain student motivation and its influence on choice, persistence, and performance in achievement-related tasks (Wigfield & Eccles, 2000). This theory has been successfully applied to explore student choices in various subject domains, such as mathematics and sciences including physics (e.g. Barnes, 1999; Eccles et al., 1998; DeBacker & Nelson, 1999; Watt, Eccles & Durik, 2006; Wigfield, 1994; Wigfield & Eccles, 2000; Woods, 2008). However, initial motivation to enrol in a subject that a student has not studied previously, does not necessarily imply sustained engagement with the subject. Although highly interrelated, motivation and engagement are distinct constructs because "while motivation could be described as the energy and direction, the reasons for behaviour, why we do what we do; engagement describes energy in action; the connection between person and activity" (Ainley, 2004, p. 2). Therefore, it is possible that some students' engagement with a subject can decline as they become more familiar with it and ultimately lead them to a decision to discontinue the subject. Consequently, the extent to which students' sustained engagement with physics influence their sustained enrolment intentions in relation to the subject also needs to be examined. Thus, potential interactive relations among motivation, engagement and the sustained choices of students can be more clearly explained.

An extensive search of the physics enrolment literature on the EV theoretical perspective identified that enrolment motivation has four major underlying EV latent variables and that any analysis should proceed to examine the sustained enrolment intentions in physics predicted by these variables. These latent variables comprise interest value of the subject, students' performance perceptions in the subject, their gender role beliefs in relation to the subject and the utility value of the subject.

Interest value. Interest value *(interest*) refers to the "inherent enjoyment or pleasure one gets from engaging in an activity" (Eccles, O'Neill & Wigfield, 2005, p. 239). Interest in science is a major predictor of science enrolment including physics (e.g. Feder, 2002; Frost, Reiss & Frost, 2005; Ivie & Stowe, 2000; Osborne, Driver & Simon, 1998; Ramsden, 1998; Stokking, 2000). Low enrolment in physics is often linked to the lack of interest in the subject. For example, physics is identified as the least interesting science among United Kingdom (UK) secondary school students (e.g Barmby & Defty, 2006). This has also been found to be the case for female students across the UK, United States (US) and Australia (Kelly, 1986; Speering & Rennie, 1996; Woolnough, 1996).

Performance perceptions. The foundational work of Eccles, Adler, Futterman, Goff, Kaczala, Meece, and Midgely (1983) demonstrated that, students' expectancies of success in a subject and their task-specific beliefs regarding the subject (which subsume self-concept of ability and perceptions of task difficulty) significantly influence their enrolment plans in relation to the subject. In this investigation, both constructs together are represented by a single construct—namely, *performance perceptions (perfperc)*—for the physics module. Numerous studies have linked this factor to the declining levels of enrolment in the subject (e.g. Barnes, 1999; Barmby & Defty, 2006; Carlone, 2004; Duckworth & Entwistle, 1974; Fullarton, Walker, Ainley & Hillman, 2003; Osborne & Collins, 2000).

Sex-stereotyped attitudes. Gender role beliefs in physics measure the extent to which an individual perceives that physics is a male domain and these beliefs have been found to significantly affect the enrolment plans of students. The term sex-stereotyped attitude towards physics (*sexstereo*) represents the construct in PMQ. Gender-biased stereotyping is ubiquitous in science education, particularly in physics education. Physics is portrayed as a "masculine" domain by society and the media, propagating a gender stereotyped concept among adolescent students that females are less suitable to or capable in physics, inhibiting females from choosing physics (Baker & Leary, 1995; Kahle & Meece, 1994). A large body of literature based on both EV and non EV frameworks demonstrates that boys' and girls' beliefs and values in physics are represented in gender stereotyped patterns (e.g. Barmby & Defty, 2006; Häussler & Hoffmann, 1998; 2000; Murphy, 1990). Students, parents and peers have been found to possess stereotyped attitudes on the ability of males and females in sciences and mathematics, generally favouring males. This perception is more profound in physics (Baker & Leary, 1995; Kessels, 2005).

Utility value. Utility value (*utility*) is defined as the "value a task acquires because it is instrumental in reaching a variety of long-and short-range goals" (Eccles & Wigfield, 1995, p. 216). The EV theoretical framework establishes the *utility* of a subject as a significant predictor of academic choice in a given subject (Eccles, 1994; Jozefowiczs, Barber & Eccles, 1993), and it is identified as the strongest predictor of physics enrolment given the perceived instrumental value of the subject for career plans (Barnes, 1999; Stokking, 2000; Whitelegg, Murphy & Hart, 2007; Woods, 2008; Zhu, 2008). In PMQ, the *utility* for a particular physics module was related to students' long term and short-term goals, such as career plans and future study plans.

Based on the literature review, it was hypothesised that these four EV variables would interact over time and influence students' sustained engagement with physics (*engage*), which in turn would predict their sustained enrolment intentions (*choicein*) in physics. The relationships among the variables are depicted pictorially as a structural equation path model where the path coefficients indicate the strength of the relationships among the variables.

The conceptual diagram depicting the latent constructs and the hypothesised relationships among them is presented in Figure 1. The model is referred to as Sustained Enrolment Model for Physics (SEMP) in this paper.

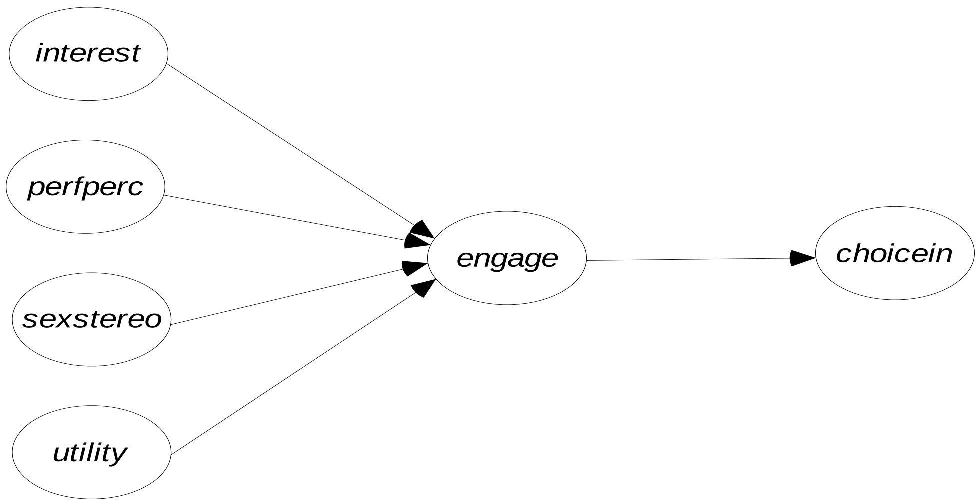

Figure 1: Conceptual diagram of the hypothesised SEMP

PMQ was developed to measure the six constructs (four predictor EV variables and the two outcome variables, *engage* and *choicein*), of the SEMP. The theoretical development of the subscales of the PMQ is first discussed in this paper. This is followed by the validation of the following psychometric properties the PMQ: (a) congenerity of the PMQ subscales; (b) reliability of the PMQ subscales (Cronbach's alpha estimates for corresponding factors); (c) factorial structure of the PMQ; (d) construct validity (convergent and discriminant validity) of the PMQ and ; (e) factorial invariance across critical groups (male vs female).

Method

Participants

The participants of this study were senior secondary school physics students in Year 11 from nine government and catholic schools in Western and Northern Sydney in NSW. Despite these schools being drawn from Western and Northern Sydney, their demographic profiles were similar. For instances all nine schools share a similar Index of Community Socio-Educational Advantage (ICSEA) score. ICSEA scores in New South Wales Department of Education and Communities (DEC) schools enable fair and meaningful comparisons of academic performance of students in a given school with that of similar schools serving students with statistically similar backgrounds. Furthermore, survey data collected from the senior physics students across the schools showed that the majority of participants came from families with high parental education and 80% of the reported parental occupations belonged to the highest twenty categories of the Socio-economic Scale (ANU4; Jones & McMillan, 2001). These findings from the sample are consistent with the linear relationship found between parental education and physics participation of children (Fullarton et al., 2003). Consequently school effects were not specifically examined.

Data were collected by administering the module-specific questionnaires on four occasions during the 2009 academic year (Time points 1 – 4). These four time points corresponded with the completion of each of the four physics modules of the Year 11 physics curriculum. These four physics modules comprise: The World Communicates (commonly referred to and used in this paper as the waves module); Moving About (motion); Electrical Energy at Home (electricity); and The Cosmic Engine (cosmic engine). The sample size across all 9 schools varied at each time point (T1 = 270, T2 = 280, T3 = 239, T4 = 222). Consistent with previous research reports (Fullarton et al., 2003) there were more males than females in the sample at each time point. The percentage of males across the four data collection points were 66% (n= 178) at time 1 (T1), 64% (n= 180) at time 2 (T2), 62% (n= 147) at time 3 (T3) and 63% (n= 140) at time 4 (T4).

Procedure

The questionnaires were administered in classrooms at the completion of each module under the supervision of the classroom physics teacher. There were no exclusion criteria for the sample selection, and students participated voluntarily each time.

Development of the Instrument

The newly developed PMQ originally consisted of 33 items constructed on the basis of the six subscales (*interest*, *utility*, *perfperc*, *sexstereo*, *engage* and *choicein*). Item development was guided by theory and existing well-established scales (see Table 1 to find the items and the scales they were adapted from). All items were measured on a six point Likert scale (1 = strongly disagree to 6 = strongly agree). Some of the items were reverse-coded, to avoid a misleading response bias in student responses. Each of the subscales was inferred from four or five indicator items, except one construct, namely *choicein*, was measured using a single indicator. A single indicator item was used since multiple items would be redundant in measuring the sustained enrolment plans of students.

Table 1: Subscales and items of original PMQ

Interest Value Subscale

Numerical identifier on PMQ	Item	Scale adapted from
1	I think what I learnt in . . . (name of the module) lessons was interesting	Eccles, O'Neill, and Wigfield (2005)
6	This module was more interesting to me than other modules *	
11	I have a real desire to study more about this module	Motivated Strategies of Learning Questionnaire (MSLQ; Pintrich & De Groot, 1990)
16	I look forward to learning more about this module	Test of Science-Related Attitudes (TOSRA; Fraser, 1981)
21	I would enjoy physics more if this module was not included in the physics curriculum	TOSRA
26	I am interested in learning more about this module *	
31	This was a fun module and I was really into it	MSLQ

Note. * = original items

Utility Value Subscale

Numerical identifier on PMQ	Item	Scale adapted from

5	This module is the core of physics, in my view *	
10	This module is a great module for my career interests	Eccles, O'Neill & Wigfield, 2005
15	I can get on perfectly well in life without knowing this module	ATSI
20	Knowing this module well would be helpful to do well in the course I want to do at Uni/TAFE*	Eccles & Wigfield, 1995
25	I see no reason why anyone needs to know about this module *	
30	Knowing this module would be very useful if you want to get a good job	Eccles & Wigfield, 1995

Note. * = original items

Performance Perceptions Subscale

Numerical identifier on PMQ	Item	Scale adapted from
3	I know I am able to do well in this module	Self and task perceptions questionnaire (Eccles et al., 2005)
8	I know that my classmates regard me as being capable in this module	ATSI
13	Compared to the majority of students in my class, I found this module easy	The self and task perception questionnaire (Wigfield & Eccles, 1995)
18	I was very good at this module	The self and task perception questionnaire (Wigfield &Eccles, 1995)
23	No matter how hard I tried, I did not understand this module	ATSI
28	This module was very hard for me*	
32	I am happy with my performance in this module*	

Note. * = original items

Sex-stereotyped Attitudes Subscale

Numerical identifier on PMQ	Item
4	I think this module was more useful for boys than for girls*
9	I think boys are naturally better than girls in this module*
14	I think this was a module for boys*
19	I think boys can perform better than girls in this module*
24	I think this module was more interesting for boys than for girls*
29	The activities associated with this module were more relevant to the life experiences of boys than that of girls*

Note. * = original items

Sustained Engagement Subscale

Numerical identifier on PMQ	Item
2	I would like to continue physics to Year 12 if the other modules are similar to this one
7	I found this module so exciting that I would like to spend more time on it.
12	This module is irrelevant to my real life experiences
17	If I have a chance, I would drop physics from my curriculum
22	I think I made a wrong decision choosing physics
27	I was enthusiastic to participate in the activities associated with this module

Sustained Enrolment Intentions Subscale

Numerical identifier on PMQ	Item
33	I do not want to continue physics to Year 12

Data Analysis Procedures

Data screening. Data screening, preliminary analyses and Cronbach's alpha reliability estimates were achieved using SPSS 15.0. Test for congenerity, CFA and invariance testing were conducted with LISREL 8.72 and using Maximum Likelihood Estimation (Byrne, 1998). Considering that missing values did not exceed 5% for any one item, the guidelines of Tabachnik and Fidell (2007) and Schafer and Graham (2002) for dealing with missing values (i.e., Expectation Maximization (EM) algorithm) and univariate and multivariate outliers were utilised. Despite EM being "well regarded" (Olinsky, Chen, & Harlow, 2003, p.58), future research should consider employing Full Information Maximum Likelihood (FIML) as the preferred method of dealing with missing data.

Congenerity of the PMQ Subscales. The one-factor congeneric models of the five multiple indicator factors of the original version of PMQ were examined. For the single indicator variable *choicein* this examination was deemed unnecessary, as the associated model for such a factor would be a saturated (just-identified) model. Furthermore, the reliability of the single indicator was a priori estimated following the guidelines set by Marsh (1990), Kelloway (1998) and Kline (1998) when employing a single indicator construct. In the models, the loadings of items on the factors they were intended to measure were also examined, and loadings > 0.5 were regarded as "high" (Kline, 1998). A model-trimming strategy (Kline, 1998) which entails eliminating paths that cause the model misfit was adopted, to improve model fit. However, the strategy was not employed purely on statistical criteria as this could have inflated Type 1 errors. Therefore, although finding the possible sources of model misspecification was executed on the basis of statistical criteria, removing such paths was guided by theoretical considerations.

Reliability tests. Internal consistency reliability was estimated using Cronbach's alpha. Indicators with a Cronbach's alpha above 0.70 were accepted as indications of an adequate measure (Hills, 2008).

Confirmatory factor analysis. Analyses proceeded to full multi-factor confirmatory factor analysis (CFA) to test construct validity of the theoretically driven *apriori* factors and, to test the hypothesised factorial structure of PMQ (see Figure 2). All items were constrained to load on only the factor it was intended to measure, while all cross-loadings were forced to zero. Following recommendations that multiple fit indices should be examined across both incremental and absolute indices, (Marsh, Balla & Hau, 1996), the χ^2 /df ratio, Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI) were reported. For the

χ^2/df less than 3 and for the CFI and TLI values greater than 0.90 are deemed acceptable. For RMSEA, a value of 0.05 indicates a close fit, values near 0.08 indicate a fair fit, and values above 0.10 indicate a poor fit (Byrne, 1998).

Additionally, the factor loadings of the items were scrutinised to confirm the hypothesised six-factor structure of the PMQ. The proportion of variance explained by the model was calculated, where 25% is considered a "large" effect (Hills, 2008, p. 72).

Invariance testing. The purpose of invariance testing is to compare across distinct groups to assess whether the instrument may hold the same meaning for all groups (Byrne, 1998). Invariance testing was conducted to determine whether the instrument measures the same components of motivation, as well as the outcome variables, with equal validity for males and females. To determine whether the invariance testing supplied a satisfactory result, the goodness of fit indices of the nested models with progressive increments of parameter restrictions were examined (Byrne & Watkins, 2003). In this investigation, these sets of parameters were identified as the *factor loading, factor variances, covariances* and, *factor uniqueness* (Byrne, 1998; Marsh, 1994). The first three models in the five nested model series (the completely-free model, factor loading invariant model and, the factor loading, variance and covariance invariant model) were considered in determining whether the particular scale was invariant across the genders. The fourth (factor loadings and factor uniquenesses invariant model) and fifth (where all parameters are held invariant across the groups) models are considered to be excessively restrictive criteria of invariance (Byrne, 1998; Marsh, 1994).

Variations in the CFI index less than 0.01 were set as the minimal requirement for factorial invariance, with any variation in RMSEA or TLI considered arbitrary in nature (Marsh, Tracey & Craven, 2006). Since the instrument included two measurement models (namely, the models for the EV predictor variables and the outcome variables), both models were subjected to invariance testing across the two critical groups of this investigation.

However, before the invariance testing, separate CFAs were conducted for male and female data, to test the likelihood of invariance. This is because, if the separate analyses reveal minimal differences between the models for males and females, then it is likely that the multi-group test, where parameters are constrained to be equal between males and females, will be invariant.

RESULTS

Congenerity of the PMQ Subscales

One-factor congeneric models of the five multiple indicator latent factors were conducted across for modules. The procedure of model re-specification for the waves module, which was conducted as per model-trimming strategy (Kline, 1998) is outlined below.

One-factor congeneric models the waves module

Interest. The fit indices of the one-factor congeneric model for *interest* with all items included indicated fair fit (RMSEA =0 .084, CFI = 0.978, TLI = 0.967). Large correlation residuals were found between PQ01 and PQ21, and PQ01 and PQ 06. PQ21 also reported the lowest factor loading among the indicators (0.37). Upon inspection of the items, it was noted that PQ01 was the single item on the subscale that measured *interest* as a past event, while the other items were oriented to future action. Likewise, PQ21 was the only negatively worded item in the model. It was reasonable to assume that these items had been subjected to differing interpretations by the participants from what was originally intended. Therefore, PQ01 and PQ21 were removed. The respecified model showed improved fit indices to the data (see Table 2).

*perfperc***.** The model with all items included indicated a poor fit (RMSEA = 0.148, CFI = 0.925, TLI = 0.888). The possible sources of model misspecification were identified in the largest correlation residual between PQ23 and PQ28. Both items also displayed the lowest factor loadings among the indicators (0.48 and 0.52 respectively). Closer inspection of PQ23 and PQ28 revealed that they were the negatively phrased items on the subscale and this might have contributed to misinterpretation by the respondents. The model with PQ23 and PQ28 removed showed an improved fit (see Table 2). It should be noted that RMSEA value does not mean a perfect fit in this instance

(Kline, 2011) but generated due to the smaller χ^2 statistic (2.48) of the model than its degrees of freedom (df = 5) (Kline, 2011; Steiger, 2000).

Sexstereo. The fit indices of the model for *sexstereo* with all items indicated a poor fit (RMSEA =0.117, CFI = 0. 983, TLI = 0. 972), Large correlation residual was reported between PQ09 and PQ19. The respecified model with PQ09 removed showed improved fit (see Table 2). The zero value for RMSEA was generated due to the smaller χ^2 statistic (3.95) of the model than its degrees of freedom (df = 5) (Steiger, 2000). A closer inspection of PQ09 suggested that this item might have been ambiguous for the respondents, as the term *better* is very general; perhaps too general to allow participants to attribute a precise meaning to it. This is a possible reason why PQ09 failed to fully measure the construct it was intended to measure. Hence, this item was removed.

Utility. The original model for *utility* for waves showed a poor fit (RMSEA = 0.112, CFI = 0.931, TLI = 0.885). PQ05 reported a low factor loading (0.25), suggesting the item was not measuring the construct. This item gave emphasis to the relative usefulness of the module rather than specifically emphasising the usefulness of the module for further study (e.g. PQ20) and the career plans of participants (e.g. PQ10), perhaps causing misinterpretation. Also the large correlation residual between PQ15 and PQ25 suggested the removal of one of the items. PQ25 was found to be different from the other item which focused on the specific usefulness of physics (e.g. securing a good job or fulfilling future study plans) while PQ25 was more general and negatively phrased. Given that it was less specific, the respondents may have found this item open to different interpretations. Therefore, PQ05 and PQ25 were removed. The refined model displayed excellent fit indices (Table 2).

The factor loading of item PQ15 was just below the accepted cut-off value (0.30) postulated by Kline (1998). However, examination of the content validity of the indicator suggested that the removal of this item (PQ15) merely on statistical criteria was not justifiable theoretically, as it was considered a suitable measure of the general instrumental value of the module. Nevertheless, the negative phrasing might have caused misinterpretation of the intended meaning. Although it was decided to retain this item tentatively, it was noted that interpretation of data from this measure that included a low loading item should be done with caution, as this is indicative that the item is not measuring the factor fully (Kline, 1998). Therefore, this item was scrutinised for its reliability further in the next step of analysis.

Engage. The fit indices of the original *engage* model with all items indicated a poor fit (RMSEA = 0.161, CFI = 0.830, TLI = 0.716). Large correlation residuals were observed between PQ02 and PQ07. Furthermore, PQ07 displayed a low factor loading (0.15). Closer inspection of PQ07 indicated a difference from the other items in the subscale. The other items related to disengagement with the Year 11 physics school curriculum more explicitly while PQ07 was referring to a more general behavioural expectation. Another item with a low factor loading was PQ12 (0.20). On further analysis it was revealed that while the other items of the scale were more specifically phrased to address engagement with classroom physics, PQ12 referred to experiences outside of the classroom. This may have caused participants difficulty in relating this item to the *engage* construct. Therefore, these items were deemed not to be measuring *engage* as intended and hence, they were removed. The fit indices of the refined model displayed an improved fit (see Table 2).

Cronbach's Alpha Reliability Estimation of the PMQ

Reliability tests were conducted on the refined 24-item PMQ. Although each latent factor demonstrated strong levels of internal consistency (0.7- 0.9), the removal of PQ15 and PQ27 suggested a significant improvement in the alpha coefficient (see Table 3). Therefore these items were analysed further**.**

It was assumed that the general nature of the phrasing of the item PQ15 might have prevented participants from linking it to the specific and immediate utilitarian value of studying the subject. Furthermore**,** unlike the other items which focused specifically on the Year 11 physics curriculum, PQ27 comprised an affective component which related to physical engagement with specific class activities. Hence these items were deemed to not fully measure the constructs they were was intended to and, were removed from respective subscales. Thus the items of the PMQ were further reduced from 24 to 22.

Item reliability estimations were conducted for the three remaining modules of the Year 11 curriculum. They demonstrated good estimates (Kline, 1998) of subscale reliability (see Table 4).

Table 2: Fit Indices of One-Factor Congeneric Models of Interest Value for Waves

interest	Item number	RMSEA with item removed	CFI with item removed	TLI with item removed
	1	.021	.999	.998
	6	.855	.981	.969
	11	.101	.968	.947
	16	.091	.972	.954
	21	.077	.985	.974
	26	.089	.974	.957
	31	.099	.969	.948
perfperc	3	.191	.877	.794
	8	.182	.905	.841
	13	.181	.900	.833
	18	.190	.877	.795
	23	.000	1.000	1.008
	28	.030	.998	.997
	32	.190	.897	.828
	4	.154	.978	.955
sexstereo	9	.000	.994	1.002
	14	.122	.984	.968
	19	.000	1.000	1.003
	24	.133	.983	.966
	29	.151	.979	.959
utility	5	.130	.884	.942
	10	.140	.855	.710
	15	.073	.980	.961
	20	.143	.866	.732
	25	.051	.991	.981
	30	.149	.874	.749
engage	2	.108	.912	.824
	7	.075	.972	.944
	12	.218	.816	.633
	17	.113	.878	.756
	22	.102	.896	.793
	27	.173	.841	.681

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

Table 3: Item Analysis of PMQ for the Waves Module

Construct	Item	Alpha with item removed	Alpha with all items present
interest	PQ06	.850	.832
	PQ11	.786	
	PQ16	.770	
	PQ26	.780	
	PQ31	.798	
perfperc	PQ03	.815	.858
	PQ08	.849	
	PQ13	.822	
	PQ18	.807	
	PQ32	.851	
sexstereo	PQ03	.815	.918
	PQ08	.849	
	PQ13	.822	
	PQ18	.807	
	PQ32	.851	
utility	PQ10	.584	.738
	*PQ15	.824	
	PQ20	.598	
	PQ30	.656	
engage	PQ02	.685	.729
	PQ17	.607	
	PQ22	.596	
	*PQ27	.754	
choicein	PQ33	.98**	

Note. interest = interest value of the physics module, *perfperc* = performance perceptions, *sexstereo* = sex-stereotyped attitudes to module, *utility* = utility value of the module, *engage* = sustained engagement with the module, *choicein* = sustained intention to continue in physics; * = item removed following Cronbach's alpha test; ** = estimated value.

Confirmatory Factor Analysis

The full CFA was conducted on the refined 22 item PMQ. This model resulted in a good fit to the data for the waves module (RMSEA =0.058; χ^2/df = 1.89; TLI =0. 958; CFI =0 .964). All indicators specified to measure a common underlying factor demonstrated high loadings (> 0.5) on that factor except PQ06, which reported a comparatively modest value of 0.45 (see Table 5). In addition, the analysis of Modification Indices (MI) of the model did not show any justification for further modification (Byrne, 1998) to the factorial structure of the PMQ.

Table 4: Reliability Estimates of PMQ subscales across the Modules

Construct	Coefficient Alpha		
	Electricity	Motion	Cosmic engine
Interest	.884	.874	.909
Perfperc	.850	.887	.865
Sexstereo	.926	.934	.934
Utility	.759	.785	.841
Engage	.702	.770	.718
Choicein	.98*	.98*	.98*

Note. interest= interest value of the physics module, *perfperc*=performance perceptions in the module, *sexstereo*= sex-stereotyped attitudes to the module, *utility*= utility value of the module, *engage*= sustained engagement with the module, *choicein*= sustained intention to continue in physics, * = estimated value.

Convergent validity and discriminant validity of the indicator sets were statistically measured through scrutinising the correlation matrices. Indicators designed to measure the same construct displayed moderate correlations among themselves, exhibiting convergent validity of the construct. Furthermore, the inter-item correlations among indicators that were intended to measure dissimilar constructs were low, displaying discriminant validity (Kline, 1998). The estimated high factor loadings and the large average variance explained for each set of indicators further demonstrated the convergent validity of items on each subscale of the PMQ (Kline, 1998).

Between-factor correlations of the six constructs of PMQ were verified to check whether the factors are measuring distinct constructs rather than different facets of a higher order latent variable (Kline, 1998). This was indicated by the correlation values that show that the constructs within the instrument were discriminant ($r < 0.85$; Kline, 1998, p. 60) with those they were dissimilar to (see Table 6). This means that the subscales were measuring distinct constructs and hence met the condition of discriminant validity (see Table 7 for the results for the three remaining modules).

Table 5: Factor Loadings of the PMQ for the Waves Module and the variance explained

	interest	*perfperc*	*sexstereo*	*utility*	*engage*	*choicein**
Factor loadings	.45	.81	.76	.88	.56	.98*
	.77	.65	.89	.78	.83	
	.81	.76	.85	.70	.79	
	.78	.84	.86			
	.72	.67	.80			
Average variance explained	52%	56%	69%	63%	36%	95%*

Note. interest = interest value of the physics module, *perfperc* = performance perceptions, *sexstereo* = sex-stereotyped attitudes to module, *utility* = utility value of the module, *engage* = sustained engagement with the module, *choicein* = sustained intention to continue in physics; * = error variance fixed *a priori*.

Table 6: Correlations among the Six Latent Variables of PMQ for the Waves Module

	interest	*perfperc*	*sexstereo*	*utility*	*engage*	*Choicein*
interest	1.000					
perfperc	.396	1.000				
sexstereo	-.109	.146	1.000			
utility	.678	.260	-.022	1.000		
engage	.360	.504	-.069	.405	1.000	
choicein	.213	.424	.041	.239	.805	1.000

Note. interest = interest value of the physics module, *perfperc* = performance perceptions, *sexstereo* = sex-stereotyped attitudes to the module, *utility* = utility value of the module, *engage* = sustained engagement with the module, *choicein* = sustained intention to continue in physics.

Table 7:Correlations among the Six Latent Variables of PMQ across the Remaining Modules

Module		*interest*	*perfperc*	*sexstereo*	*utility*	*engage*	*choicein*
Electricity	*interest*	1.000					
	perfperc	.516	1.000				
	sexstereo	.035	.092	1.000			
	utility	.652	.354	.035	1.000		
	engage	.338	.515	-.263	.265	1.000	
	choicein	.182	.319	-.223	.161	.824	1.000
Motion	*interest*	1.000					
	perfperc	.710	1.000				
	sexstereo	-.019	.019	1.000			
	utility	.626	.484	.023	1.000		
	engage	.494	.568	-.288	.427	1.000	
	choicein	.343	.415	-.290	.258	.838	1.000
Cosmic engine	*interest*	1.000					
	perfperc	.761	1.000				
	sexstereo	-.174	-.072	1.000			
	utility	.668	.539	.079	1.000		
	engage	.382	.439	-.418	.277	1.000	
	choicein	.170	.309	-.285	.117	.777	1.000

Note. interest= interest value of the physics module, *perfperc*=performance perceptions in the module, *sexstereo*= sex-stereotyped attitudes to the module, *utility*= utility value of the module, *engage*= sustained engagement with the module, *choicein*= sustained intention to continue in physics.

Assessment of factorial structure across the other modules

The six-factor CFAs were conducted for the remaining three physics modules (see Table 8 for the fit indices of the models, Table 9 for the factor loadings and Table 10 for the average variance explained for the factors). The results show that CFAs for the PMQ demonstrated acceptable fits to the data across the remaining physics modules, thus showing support to the six-factor structure of the PMQ.

The analysis of construct validity supported the six-factor structure of the PMQ for all four modules. The inter-item correlations demonstrated that the indicators were convergent within the corresponding construct, and discriminant from other factors. The factor correlation coefficients among factors demonstrated the PMQ six first-order scales as being discriminant among all the six constructs for all the modules (see Table 7).

Table 8: Fit indices of CFA Models across the Physics Modules

Module	χ^2	df	χ^2/df	TLI	CFI	RMSEA	Type of fit
electricity	484.08;	195	2.48	.942	.951	.073	Fair
motion	457.06	195	2.34	.947	.956	.075	Fair
cosmic engine	567.47	195	2.91	.928	.939	.092	Fair

Note. χ^2 = Chi-square; df = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

Table 9: Factor Loadings of PMQ subscales across the Modules

Module	Factors					
	interest	*perfperc*	*sexstereo*	*utility*	*engage*	*choicein**
Electricity	.631	.789	.781 .846	.801	.392	.98*
	.781	.639	.886	.743	.838	
	.841	.665	.884	.614	.828	
	.815	.874	.828			
	.814	.708				
Motion	.650	.762	.836	.762	.548	.98*
	.786	.706	.847	.824	.827	
	.780	.767	.916	.648	.866	
	.824	.858	.873			
	.768	.830	.835			
Cosmic	.750	.796	.729	.852	.453	.98*
engine	.832	.733	.907	.836	.827	
	.885	.688	.925	.712	.847	
	.827	.803	.831			
	.798	.724	.904			

Note. interest= interest value of the physics module, *perfperc*=performance perceptions in the module, *sexstereo*= sex-stereotyped attitudes to the module, *utility*= utility value of the module, *engage*= sustained engagement with the module, *choicein*= sustained intention to continue in physics; * = error variance fixed *a priori***.**

Table 10: Average Variance Explained for the PMQ subscales across the Modules

Module	Factors					
	interest	*perfperc*	*sexstereo*	*utility*	*engage*	*choicein*
Electricity	61%	55%	72%	52%	51%	.96%
Motion	58%	62%	74%	56%	58%	.96%
Cosmic engine	68%	56%	74%	64%	54%	.96%

Note. interest= interest value of the physics module, *perfperc*=performance perceptions in the module, *sexstereo*= sex-stereotyped attitudes to the module, *utility*= utility value of the module, *engage*= sustained engagement with the module, *choicein*= sustained intention to continue in physics.

Factorial Invariance Testing of the PMQ by Gender Groups

Prior to commencement of factorial invariance testing, the sample (n = 270) was split by gender and a separate CFA for males (n = 178) and females (n = 92) was conducted to ensure satisfactory goodness of fit estimates were identified. Results indicate that both separate analyses provided acceptable fits to the data, and suggested potential consistency across gender (see Table 11).

Factorial Invariance of the first measurement model. The first measurement model represented the exogenous variables of the SEMP. They included the four EV variables, namely *interest, perfperc, sexstereo* and *utility*. Five models were tested for factorial invariance of the first measurement model across gender (see Table 12). Examining the variation in the CFI for the nested models, as recommended by Cheung and Renvold (2002), it was identified that invariance across the male and female groups was achieved for the factor loading (Model 2) and for the factor loading, variance and covariance (Model 3). Further invariance testing of the parameters of the uniqueness (Models 4 and 5) also achieved invariance. The results of the invariance testing support that the factor structure of the EV measurement model of the PMQ is consistent across the male and female student sample.

Factorial invariance of the second measurement model. The second measurement model tested included the representation of the endogenous variables (outcome variables) included in the hypothesised SEMP, namely, *engage* and *choicein*. Examination of the variation in the CFI of the five nested models for the second measurement model (see Table 13) provided evidence for the invariance across the male and female groups for the factor loading (Model 2) and, factor loading, factor variance, and covariances (Model 3). The invariance testing for factor loading and uniqueness and, all the parameters taken together (Models 4 and 5) fell outside the acceptable range; however, they represent overly restrictive practice (Byrne, 1998), and therefore did not affect the invariance testing. Therefore, it was concluded that the measurement model of the outcome variables of the PMQ was also invariant across gender for the waves module.

Table 11: CFAs for Males and Females for the Waves Module

Model	χ^2	df	χ^2/df	TLI	CFI	RMSEA
Hypothesised model Males	334.19	195	1.71	.942	.951	.064
Hypothesised model Females	276.06	216	1.27	.941	.950	.068

Note. χ^2 = chi-square; df = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

Table 12: Factorial Invariance across Gender for the First Measurement Model for the Waves Module

	Model	χ^2	df	χ^2/df	TLI	CFI	RMSEA
1	Completely free	371.21	258	1.44	.910	.966	.057
2	Invariant factor loading	383.33	272	1.41	.962	.967	.055
3	Invariant factor loading, variance & covariance	391.76	282	1.39	.965	.967	.054
4	Invariant factor loadings and uniqueness	420.14	290	1.44	.957	.960	.058
5	Completely Invariant	428.83	300	1.43	.959	.960	.057

Note. χ^2 = chi-square; df = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

Table 13: Factorial Invariance across Gender for the Second Measurement Model for the Waves Module

	Model	χ^2	df	χ^2/df	TLI	CFI	RMSEA
1	Completely free	4.81	4	1.20	.995	.998	.039
2	Invariant factor loading	6.43	6	1.07	.998	.999	.023
3	Invariant factor loading, variance & covariance	7.88	9	.87	1.03	1.000	.000
4	Invariant factor loadings and uniqueness	15.85	9	1.76	.970	.977	.075
5	Completely Invariant	17.51	12	1.45	.981	.981	.059

Note. $\chi^{2\,=}$ Chi-square; df = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation.

Analysis of results of factorial invariance demonstrated that the PMQ was gender invariant for the data for the other three modules as well (see Table 14). This meant that specific facets of the constructs included in the PMQ, or correlations among these facets, did not qualitatively differ markedly between males and females in relation to these modules (Martin, 2004).

Table 14: Factorial Invariance across Gender for the Measurement Models across the Modules

	Model	CFI	
		First Measurement model	Second Measurement Model
Electricity	1	.899	.999
	2	.901	1.000
	3	.899	1.000
	4	.875	.967
	5	.884	.973
Motion	1	.903	.995
	2	.899	.986
	3	.899	.981
	4	.869	.941
	5	.869	.929
Cosmic engine	1	.930	.997
	2	.930	.999
	3	.929	.996
	4	.915	.965
	5	.914	.944

Note. CFI=Comparative Fit Index; 1= Completely free, 2= Invariant factor loading, 3= Invariant factor loading, variance & covariance, 4= Invariant factor loadings and uniqueness, 5= Completely Invariant.

DISCUSSION

Psychometric properties of the instrumentation

The results suggest PMQ is a psychometrically sound, valid, and a robust measure of the constructs included in the SEMP, and is appropriate for use with male and female adolescent physics students. Across the physics modules, the subscales were congeneric measures of the specified constructs included in the PMQ. In addition, the internal consistency reliability for each of the subscales was acceptable across all of the physics modules, verifying that the indicators were reliable measures of what they were intended to measure. The validity and the gender invariance of PMQ were also verified.

Implications for theory

Consistent with Eccles and Wigfield's investigations (1995; Eccles, Wigfield, Harold & Blumenfeld, 1993), the *perfperc* subscale of the PMQ measured both *self-concept of ability* and *task difficulty*. Since these two constructs have been shown to correlate highly and since both significantly influence students' enrolment plans, researchers have proposed that they are empirically indistinguishable (Eccles & Wigfield, 1995). The construct validity of *perfperc* in the PMQ provides further support for Wigfield's adapted version of General Model of Academic Choice (GMAC; 1994), where the two constructs are in a similar manner represented by a single construct, namely, task-specific beliefs.

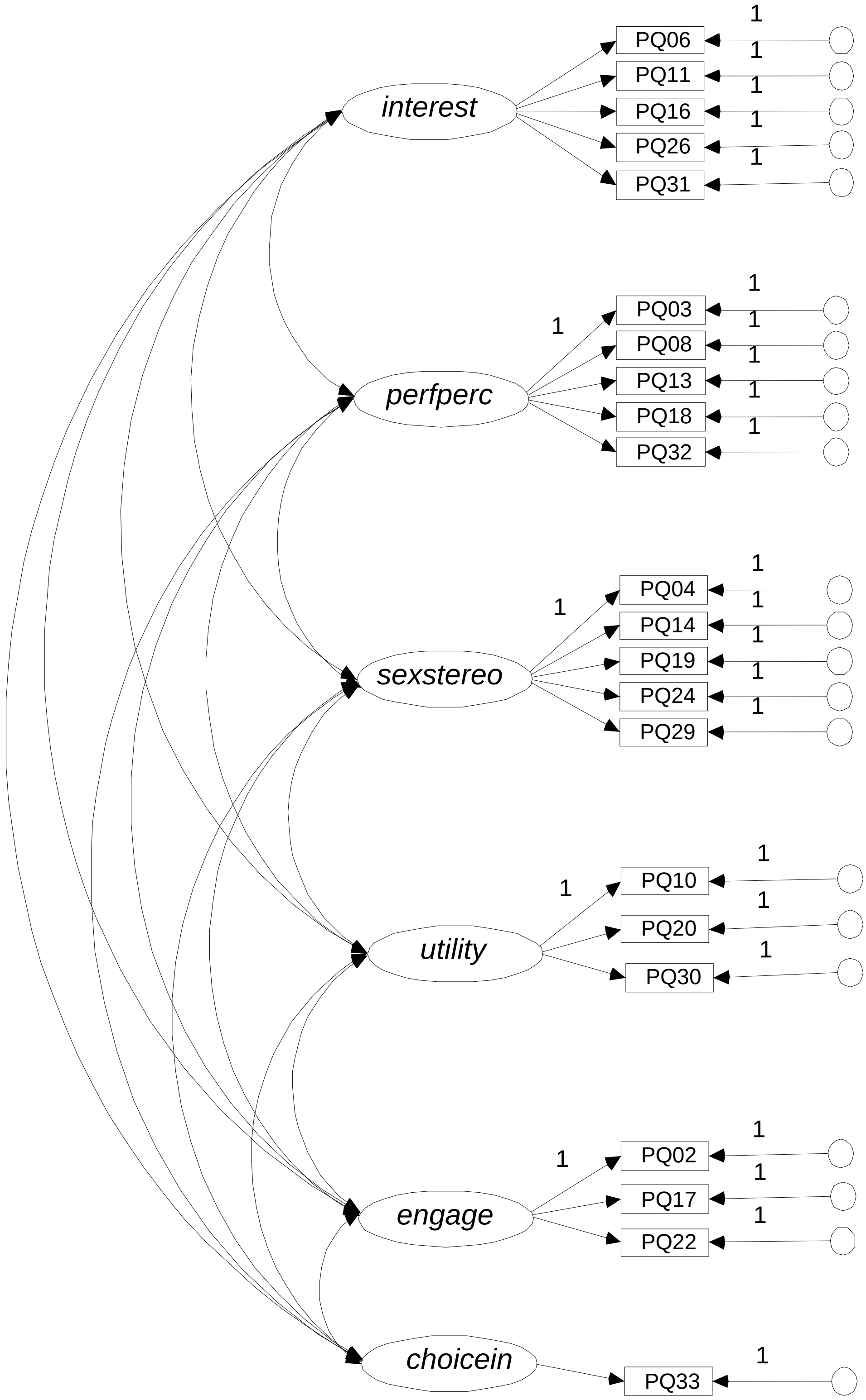

Figure 2. The CFA model of the refined PMQ for the waves module**.**

*Note. interes*t = interest value of physics, *utility* = utility value of physics, *sex-stereo* = sex-stereotyped attitudes towards physics, *perfperc* = performance perceptions in physics, *engage* = sustained engagement with physics**,** *choicein* = sustained enrolment intentions in further physics**.**

The validation of the factorial structure of PMQ and the discriminant validity of EV constructs across the four physics topics support the existence of the expectancy and value constructs of achievement motivation as separate entities in the data sets even at a topic-specific level. The operational distinction between expectancy and value constructs has been previously established for physics as a domain (Barnes, 1999; Eccles et al., 1998; Woods, 2008) and in other subjects, such as mathematics and English (Eccles et al., 1983; Eccles et al., 1998) but this is one of the first studies to go beyond domain-specific EV measures to topic-specific EV measures.

The six-factor structure of the PMQ also supports the breakdown of subjective task values into *interest value* and *utility value*, as represented in the GMAC by Eccles et al. (1983). Moderate factor correlations between these value constructs indicate that they are measures of module-specific subjective task values, yet are distinct enough to indicate that they measure separate aspects of task values. This result supports previous findings (Barnes, 1999; Woods, 2008) at a domain specific level for physics.

The factor analysis of the PMQ further adds to the research evidence supporting the discriminant validity of the subjective motivation and engagement constructs. In addition, the low correlations observed between motivation and engagement constructs of the PMQ demonstrate that the constructs were distinct for the four physics modules. This finding offers support for the delineation of these constructs proposed by Russell, Ainley, and Frydenberg (n.d) (see also Martin, 2005).

Implications for practitioners

PMQ was validated as a psychometrically sound measure to examine student motivation, engagement and intentions to continue with physics at a topic-specific level. This meant that this instrument has a greater sensitivity than the existing domain specific instruments in measuring and therefore, fine-grained analysis of students' physics enrolment motivation is possible with this instrument. This finding has significant implications for educational practitioners including school counsellors. For example, this instrument could be used by educational practitioners to evaluate the effectiveness of interventions designed to increase student enrolment intentions.

The gender invariance in the factor structure of the PMQ suggests that there are no fundamental differences in the way in which males and females perceive the motivation and engagement constructs, although the degree to which they are motivated and engaged with the subject can be different (see Abraham & Barker, 2014a). This holds implications not only for researchers employing this scale, but also for educational practitioners employing intervention programs and classroom strategies aimed at enhancing students' motivation and engagement with physics. As there was no fundamental difference observed in the student perceptions, teachers can be confident that male and female students hold equal values and expectancies of success in relation to physics. This knowledge is essential first step to consider initiatives to bolster numbers to address the current attrition of female students from physics.

Given the perceived 'masculinity' of the subject, the understanding of the gender difference in motivation and engagement will provide a strong foundation for effective practices in physics classroom. For example, according to the Institute of Physics report (2012), "for the best practice in physics classes differences between girls and boys, and the teaching styles that suit each should be recognised and followed" (p.8). However, findings from several studies suggest that teaching practices may contribute to the enhancement of the gender stereotyping problem, even though teachers are unaware of it (e.g. Zohar & Bronshtein, 2005). Likewise, based on related studies Hoffmann (2002) asserts that male students generally receive more attention and have more interaction with teachers in science classes (e.g. Spender, 1982, cited in Hoffmann, 2002; Taber, 1992); while Crossman (1987, cited in Taber, 1992) argues that the differential treatment of boys and girls is found to be most prominent in physics classes. Also, teachers are found to set higher expectations for boys' achievement in science than for girls' (Hoffmann, 2002; Millar & Toscano, 2006) and often pay more attention to contributions from boys than from girls in physics classrooms. Educational practitioners who hold gender-biased notions should be alerted to the equivalence of males' and females' motivation, engagement and sustained enrolment plans while setting instructional strategies and learning goals. However, there is a possibility that the degree to which males and females are motivated can be different. Further exploration of comparative analysis is possible with PMQ (see Abraham & Barker, 2014a). It is recommended that for future studies to examine mean differences, a test of intercept invariance should be conducted with the PMQ. Demonstrating invariant intercepts shows group differences in the means of observed PMQ items stem from differences in the means underlying constructs. Examining mean differences will enable educational practitioners to identify gender differences in the given motivational variables more sensitively at a topic-specific level and test whether they can vary with topic. This has significant implications for researchers and educational practitioners who operate in the contexts of improving females' motivation in physics, given their underrepresentation in the subject. Although ICSEA scores demonstrate that the nine schools identified for this study are statistically similar and therefore we did not specifically account for school effects in our models, it would however, be fruitful for future research to examine school effects utilising multilevel models to test the PMQ since individuals are nested within schools.

Data collected employing PMQ could test the empirical viability of the theoretically developed SEMP, which examines the predictive power of the major latent factors in sustaining students' engagement and enrolment intentions in the enacted senior secondary physics curriculum. The strength of the relations among the constructs can indicate the importance that students place on the constructs in respect of sustaining enrolment decisions in physics (Abraham & Barker, in press). This would enable the educators to develop and implement intervention plans to sustain students' intentions to continue the study of physics. PMQ could be modified for other physics topics, although it would require revalidation.

Conclusion

The findings from this study are important to share with teachers and counsellors to inform them about the motivations students espouse for physics while they are currently engaged in the subject. The sound measurement properties of PMQ provide accurate conclusions regarding factor structure and the interrelationships among the variables the scale is intended to measure. Furthermore, the gender invariance of PMQ suggests that male and female students' perception of the key facets of motivation, sustained engagement and choice intentions in relation to physics seemed to be qualitatively the same. Further studies employing this instrument can examine the strongly held view that these perceptions differ quantitatively between genders and, the relative influence of the motivational precursors of students' retention plans in relation to physics.

REFERENCES

Abraham, J., & Barker, K. (2014a). Exploring Gender Difference in Motivation, Engagement and Enrolment Behaviour of Senior Secondary Physics Students in New South Wales. *Research in Science Education*, 1-15. doi: 10.1007/s11165-014-9413-2

Abraham, J., & Barker, K. (in press). An Expectancy-Value Model for Sustained Enrolment Intentions of Senior Secondary Physics Students. *Research in Science Education*.

Ainley, M. (2004). What Do We Know about Student Participation and Engagement? Paper presented at the Australian Association for Research in Education Melbourne, Australia.

Ainley, J., Kos, J., & Nicholas, M. (2008). Participation in Science, Mathematics and Technology in Australian Education (Research Monograph No 63). Camberwell, Victoria, Australia: The Australian Council for Educational Research.

Baker, D., & Leary, R. (1995). Letting Girls Speak out about Science. Journal of Research in Science Teaching, 32(1), 3–27. doi: 10.1002/tea.3660320104

Barmby, P., & Defty, N. (2006). Secondary School Pupils' Perceptions of Physics. Research in Science & Technological Education, 24(2), 199–215.

Barnes, G. R. (1999). A Motivational Model of Enrolment Intentions in Senior Secondary Science Courses in New South Wales Schools. Doctoral dissertation, University of Western Sydney, Macarthur.

Byrne, B. M. (1998). Structural Equation Modelling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications and Programming. Mahwah, NJ: Erlbaum.Byrne, B. M. (2009).

Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming. CRC Press.

Byrne, B. M., & Watkins, D. (2003). The Issue Of Measurement Invariance Revisited. Journal of Cross-Cultural Psychology, 34(2), 155–175. doi: 10.1177/0022022102250225

Carlone, H. B. (2004). The cultural production of science in reform-based physics: Girls' access, participation, and resistance. Journal of Research in Science Teaching, 41(4), 392–414. doi: 10.1002/tea.20006

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Structural Equation Modeling: A Multidisciplinary Journal, 9(2), 233–255.

DeBacker, T. K., & Nelson, R. M. (1999). Variations on an Expectancy-Value Model of Motivation in Science. Contemporary Educational Psychology, 24(2), 71–94. doi: DOI: 10.1006/ceps.1998.0984

DEST. (2006). Audit of Science, Engineering and Technology Skills Summary Report. Department of Education Science and Training Canberra. Retrieved from http://www.dest.gov.au/NR/rdonlyres/AFD7055F-ED87-47CB-82DD-3BED108CBB7C/13065/SETSAsummaryreport.pdf

Duckworth, D., & Entwistle, N. J. (1974). The Swing from Science: a Perspective from Hindsight. Educational Research, 17(1), 48–53.

Eccles, J. S. (1994). Understanding Women's Educational and Occupational Choices: Applying the Eccles et al. Model of Achievement-Related Choices. Psychology of Women Quarterly 18(4), 585-609. doi: 10.1111/j.1471-6402.1994.tb01049.x

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgely, C. (1983). Expectancies, Values and Academic Behaviors. In J.T. Spence (Ed.), Achievement and Achievement Motivation (pp. 75–146). San Francisco, CA: Freeman.

Eccles, J., Barber, B., Updegraff, K., & O'Brien, K. (1998). An Expectancy-Value Model of Achievement Choices: The Role Of Ability Self-Concepts, Perceived Task Utility, and Interest in Predicting Activity Choice and Course Enrolment. In A. Krapp, A. Renninger & J. Baumert (Eds.), Interest and learning: Proceedings of the Seeon Conference on Interest and Learning (pp. 267–279): Institute for Science Education at the University of Kiel (IPN).

Eccles, J., O'Neill, S., & Wigfield, A. (2005). Ability Self-Perceptions and Subjective Task Values in Adolescents and Children. In K. A. Moore & L. H. Lippman (Eds.), What Do Children Need to Flourish? (Vol. 3, pp. 237–249): Springer US.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and Gender Differences in Children's Self- and Task Perceptions during Elementary School. Child Development, 64(3), 830–847. doi: 10.1111/j.1467-8624.1993.tb02946.x

Eccles, J. S., & Wigfield, A. (1995). In the Mind of the Actor: The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs. Personality and Social Psychology Bulletin, 21(3), 215–225. doi: 10.1177/0146167295213003

Feder, T. (2002). Women, and Some Men, Ask Why Women Don't Flock to Physics. Physics Today, 55(5), 24–26.

Fraser, B. J. (1981). TOSRA Test of Science Related Attitudes Handbook. Hawthorn, Vic., Australia: Australia Council for Educational Research.

Frost, S., Reiss, M., & Frost, J. (2005). Count me in! Gender and Minority Ethnic Attainment in School Science. School Science Review, 86(316).

Fullarton, S., Walker, M., Ainley, J., & Hillman, K. (2003). Longitudinal studies of Australian youth: Research report 33: Patterns of participation at Year 12. Melbourne: Australian Council for Educational Research: Camberwell, Victoria.

Häussler, P., & Hoffmann, L. (1998). Qualitative Differences in Students' Interest in Physics and the Dependence Onon Gender and Age. Paper Presented at the Proceedings of The Seeon-Conference on Interest and Gender.

Häussler, P., & Hoffmann, L. (2000). A Curricular Frame for Physics Education: Development, Comparison with Students' Interests, And Impact on Students' Achievement and Self-Concept. Science Education, 84(6), 689–705. doi: 10.1002/1098-237x(200011)84:6<689::aid-sce1>3.0.co;2-l

Hills, A. M. (2008). Foolproof Guide to Statistics Using SPSS. Frenchs Forest, Australia: Pearson Education.

Hoffmann, L. (2002). Promoting Girls' Interest and Achievement in Physics Classes For Beginners. Learning and Instruction, 12(4), 447–465. doi:10.1016/s0959-4752(01)00010-x

Institute of Physics (2012). It's Different for Girls. London, UK, IOP. http://www.iop.org/publications /iop/2012/page_58292.html (accessed 19th January, 2014)

Ivie, R., & Stowe, K. (2000). Women in Physics, 2000: AIP.

Jones, F. L., & McMillan, J. (2001). Scoring Occupational Categories for Social Research: A Review of Current Practice, with Australian Examples. Work, Employment & Society, 15(3), 539–563. doi: 10.1177/09500170122119147

Jozefowicz, D. M., Barber, B. L., & Eccles, J. S. (1993). Adolescent Work-Related Values and Beliefs: Gender Differences and Relation to Occupational Aspirations. Paper presented at the Biennial Meeting of the Society for Research on Child Development, New Orleans.

Kahle, J. B., & Meece, J. (1994). Research on Gender Issues in the Classroom. In D. L. Gabel (Ed.), Handbook of Research on Science Teaching and Learning (pp. 542–557). Toronto: Macmillan.

Kelloway, K. E. (1998). Using LISREL for Structural Equation Modelling. Thousand Oaks: SAGE Publications.

Kelly, A. (1986). The Development of Girls and Boys Attitudes to Science—A Longitudinal Study. European Journal of Science Education, 8(4), 399–412. doi:10.1080/0140528860080407

Kessels, U. (2005). Fitting into the Stereotype: How Gender-Stereotyped Perceptions of Prototypic Peers Relate to Liking for School Subjects. European Journal of Psychology of Education, 20(3), 309–323.

Kline, R. B. (1998). Principles and Practice of Structural Equation Modelling. New York: The Guilford Press.

Kline, R. B. (2011). Principles and Practice of Structural Equation Modeling. (3 ed.).New York: Guilford Press.

Lyons, T., & Quinn, F. (2010). Choosing Science: Understanding the Declines in Senior High School Science Enrolments. Research Report to the Australian Science Teachers Association: UNE.

Marsh, H. (1990). A Multidimensional, Hierarchical Model of Self-Concept: Theoretical and Empirical Justification. Educational Psychology Review, 2(2), 77–172. doi: 10.1007/bf01322177

Marsh, H. W. (1994). Confirmatory Factor Analysis Models of Factorial Invariance: A Multifaceted Approach. Structural Equation Modeling: A Multidisciplinary Journal, 1(1), 5 –34

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An Evaluation of Incremental Fit Indices: A Clarification of Mathematical and Empirical Processes. In G. A. Marcoulides & R. E. Schumacker (Eds.), Advanced Structural Equation Modeling: Issues and Techniques. Hillsdale, NJ: Erlbaum.

Marsh, H. W., Tracey, D. K., & Craven, R. G. (2006). Multidimensional Self-Concept Structure for Preadolescents With Mild Intellectual Disabilities. Educational and Psychological Measurement, 66(5), 795–818. doi: 10.1177/0013164405285910

Martin, A. J. (2004). School Motivation of Boys and Girls: Differences of Degree, Differences of Kind, or Both? Australian Journal of Psychology, 56(3), 133–146. doi: 10.1080/00049530412331283363

Martin, A. J. (2005). Exploring the effects of a youth enrichment program on academic motivation and engagement. Social Psychology of Education, 8(2), 179-206. doi: 10.1007/s11218-004-6487-0

Millar, V., & Toscano, M. (2006). Girls in Physics. Australian School Innovation in Science,Technology and Mathematics (ASISTM) Project.: Retrieved from :http://www.vicphysics.org/documents/events/ stav2006/A2.ppt

Morgan, V., & Dunn, S. (1988). Chameleons in the Classroom: Visible and Invisible Children in Nursery and Infant Classrooms. Educational Review, 40(1), 3–12.

Murphy, P. (1990). Gender Gap in the National Curriculum. Physics World, 3(1), 11-12.

Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. European Journal of Operational Research, 151, 53–79.

Osborne, J., Driver, R., & Simon, S. (1998). Attitudes to Science: Issues and Concerns. School Science Review, 79(288), 27–33.

Osborne, J. F., & Collins, S. (2000). Pupils' and Parents' Views of the School Science Curriculum: London: King's College London.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and Self-Regulated Learning Components of Classroom Academic Performance. Journal of Educational Psychology, 82(1), 33–40.

Ramsden, J. M. (1998). Mission Impossible?: Can Anything be Done about Attitudes to Science? International Journal of Science Education, 20(2), 125–137.

Russell, V. J., Ainley, M., & Frydenberg, E. (n. d.). Schooling Issues Digest: Student Motivation and Engagemen. Retrieved from http://www.dest.gov.au/sectors/school_education/publications_resources/schooling_issues_digest/schooling_issues_digest_motivation_engagement.htm

Schafer, J. L., & Graham, J. W. (2002). Missing Data: Our View of the State of the Art. Psychological Methods, 7(2), 147–177.

Speering, W., & Rennie, L. (1996). Students' Perceptions about Science: The Impact of Transition from Primary to Secondary School. Research in Science Education, 26(3), 283–298. doi: 10.1007/bf02356940

Steiger, J. H. (2000). Point Estimation, Hypothesis Testing, and Interval Estimation Using the RMSEA: Some Comments and a Reply to Hayduk and Glaser. Structural Equation Modeling, 7(2), 149–162. doi: 10.1037/0033-2909.103.3.391

Stokking, K. M. (2000). Predicting the Choice of Physics in Secondary Education. International Journal of Science Education, 22(12), 1261–1283.

Tabachnik, B. G., & Fidell, L. S. (2007). Using Multivariate Statistics (4th ed.). Needham Heights, MA: Allyn and Bacon

Taber, K. S. (1992). Girls' Interactions with Teachers in Mixed Physics Classes: Results of Classroom Observation. International Journal of Science Education, 14(2), 163–180. doi: 10.1080/0950069920140205

Vickers, M., & Ha, M., T. (2007). The Imagined Curriculum: Who Studies Computing and Information Technology Subjects at the Senior Secondary Level? In J. Lynch (Ed.), Gender and IT (pp. 27–39). Altona,Victoria: Common Ground Publishing.

Watt, H., M,G, Eccles, J. S., & Durik, A., M. (2006). The Leaky Mathematics Pipeline For Girls: A Motivational Analysis Of High School Enrolments in Australia and the USA. Equal Opportunities International, 25(8), 642–659. doi: 10.1108/02610150610719119

Whitelegg, E., Murphy, P., & Hart, C. (2007). Girls and Physics: Dilemmas and Tensions. In R. Pintó & D. Couso (Eds.), Contributions from Science Education Research (pp. 27–36): Springer Netherlands.

Wigfield, A. (1994). Expectancy–value Theory of Achievement Motivation: A Developmental Perspective. Educational Psychology Review, 6(1), 49–78. doi: 10.1007/bf02209024

Wigfield, A., & Eccles, J. S. (2000). Expectancy-Value Theory of Achievement Motivation. Contemporary Educational Psychology, 25(1), 68–81. doi: DOI: 10.1006/ceps.1999.1015

Woods, C. (2008). Factors Predicting High Ability Girls' Enrolment in Year 11 and 12 Physics and Intermediate Mathematics in Regional and Rural North Queensland. Paper presented at the AARE Conference, Brisbane.

Woolnough, B. E. (1996). Changing Pupils' Attitudes to Careers in Science. Physics Education, 31(5), 301.

Zhu, Z. (2008). A Comparison Study of Students' Reasons for Taking Physics between Two Australian High Schools and Two Chinese High Schools. Paper presented at the AARE 2008 International Education Conference Brisbane. http://www.aare.edu.au/08pap/zhu08544.pdf

Zohar, A., & Bronshtein, B. (2005). Physics Teachers' Knowledge and Beliefs Regarding Girls' Low Participation Rates in Advanced Physics Classes. International Journal of Science Education, 27(1), 61–77. doi: 10.1080/0950069032000138798.