Self-Efficacy of First Year University Physics Students: Do Gender and Prior Formal Instruction in Physics Matter?

Christine Lindstrøm[a, b] and Manjula D. Sharma[a]

Corresponding author: Christine.Lindstrom@hioa.no
[a]School of Physics, University of Sydney, Sydney NSW 2006, Australia
[b]Educational Development Unit, Program for Teacher Education, The Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Keywords: self-efficacy, gender, prior knowledge, tertiary physics education

International Journal of Innovation in Science and Mathematics Education, *19*(2), 1-19, 2011.

Abstract

Self-efficacy represents a person's belief that he or she can perform a particular task. It has been found to correlate with academic achievement and people's choice of subjects and career. While relatively widely studied, self-efficacy has not received much attention in tertiary physics. Therefore, we adapted and validated a short Physics Self-Efficacy Questionnaire before administering it four times in one year to the first-year physics cohort at the University of Sydney (*N* between 122 and 281). Investigating whether gender and prior formal physics instruction mattered to students' physics self-efficacy, we found that both showed a significant effect. Females consistently reported lower self-efficacy than males, and males with no prior formal physics instruction showed the highest self-efficacy of *any* subgroup, suggesting a 'male overconfidence syndrome'. Investigating correlations between students' physics self-efficacy and end-of-semester physics examination scores, these only seemed to develop after a relatively long time of physics study (of the order of a year or more); females developed such a correlation faster than males. We conclude that gender and prior formal instruction in physics do matter when studying physics self-efficacy, which may have important consequences both for the study of self-efficacy itself, and for the way tertiary physics is taught.

Introduction

It is a commonly held belief that physics is a particularly challenging subject. To understand this belief, we need to first understand its various facets amongst current students. One important facet is self-efficacy, which represents a person's belief that he or she can perform a certain task (Bandura, 1994), physics in this study. Although measures of self-efficacy show certain consistent features, there are important variations that make studying self-efficacy across different subjects and student groups critical for understanding which variables influence self-efficacy. Whereas undergraduate physics students' attitudes and beliefs have received much attention in recent years (Gire & Jones, 2009; Gray, Adams, Wieman & Perkins, 2008; Otero & Gray, 2008), the research conducted on self-efficacy in tertiary physics education is sparse (Dalgety & Coll, 2006; Fencl & Scheel, 2005; Shaw, 2004). Consequently, this study focuses on tertiary students' self-efficacy in undergraduate physics, which has only recently begun receiving attention (Fencl & Scheel, 2004, 2005; Gungor, Eryilmaz & Fakioglu, 2007; Sawtelle, Brewe & Kramer, in press; Shaw, 2004).

This research had three overall purposes. First, to develop a questionnaire to measure students' *physics* self-efficacy in tertiary education. Second, to investigate differences in self-efficacy for males and females, both with and without prior formal senior high school physics instruction, in their first year of university physics studies at different times of the year. Third, to carry out a study of first year students' physics self-efficacy and academic achievement across one academic year.

Background

Self-efficacy is defined as 'people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives' (Bandura, 1994, p. 71). It has consistently been found to be a good predictor of academic achievement, study strategies, and persistence in the face of difficulty (Cavallo, Rozman & Potter, 2004; Pajares, 2002), and of choice of academic major and career (Hackett, 1995).

There are different levels of self-efficacy ranging from global life skills ('When I make plans, I am certain I can make them work'), through general academic self-efficacy, domain specific self-efficacy (e.g., a specific university course), down to task-specific self-efficacy (e.g., personal belief in ability to perform uncertainty calculations within a physics course) (Choi, 2005; Lent, Brown & Gore, 1997). Of importance is that the correlation between a self-efficacy measure and the achievement measure is greatest when the two measures are matched in their level of specificity (Choi, 2005; Lent et al., 1997).

Self-efficacy and academic tasks

Self-efficacy is a dynamic construct that can be influenced and changed by feedback. Bandura (1997) identifies four sources of self-efficacy: 1) mastery experience, 2) vicarious experience, 3) verbal (or social) persuasion, and 4) physiological and affective states. Mastery experiences refer to situations in which students master a task, which in turn influences their belief in their capability to achieve their potential (Cervone, 2000; McInerney & McInerney, 2002; Palmer, 2006). In physics those tasks could be solving problems, leading to solving more challenging problems, or understanding new concepts or how concepts are linked. Vicarious experiences refer to judgement of personal capabilities through observation of others. In particular, when one observes a peer of similar ability mastering a task, this reinforces the belief that one can also perform the same task. Verbal or social persuasions refer to messages from significant others (such as family, teachers, and peers) (Zeldin, Britner & Pajares, 2008). When positive appraisal based on actual performance is provided, emphasizing that the student is making progress (McInerney & McInerney, 2002; Palmer, 2006), it can boost their self-belief in personal achievement potential. The final source of self-efficacy is physiological and affective states such as stress and irrational devaluating personal convictions. The way these four sources of self-efficacy combine to produce an overall self-efficacy belief varies between individuals as well as between different domains (Bandura, 1997).

In subjects with which students are familiar, firm beliefs about performance capabilities are developed, and students show fairly stable self-efficacy (Cervone & Palmer, 1990). A certain internal resistance to change is necessary to avoid being greatly affected by temporary anomalies in performance, but there is a fine line between a healthy and unhealthy resistance. It has been found that it is not uncommon for students to keep an unrealistic self-efficacy in the face of repeated counter-evidence (Cantor & Kilhstrom, 1987). In such cases of poor performances the correlation between self-efficacy and performance is reduced. Furthermore, students who do not respond to feedback increase their risk of failure.

Unlike students who are familiar with the subject, novices are not expected to have formed stable self-efficacy beliefs related to that subject. Their belief in their potential to achieve *should* be tentative only and easily changed in response to feedback (Cervone & Palmer, 1990). However, evidence exists that initial self-efficacy can be surprisingly resistant to change, even in the face of clear counter-evidence (Lepper, Ross & Lau, 1986). Cervone and Palmer (1990) showed that people require several rounds of feedback before a stable and well-calibrated self-efficacy is established. These findings were in agreement with Tversky and Kahneman's (1974) description of the 'anchoring and adjustment' strategy where, upon receiving feedback, people adjust their self-efficacy to yield a final value that is *biased in the direction* of the original self-efficacy (anchor), rather than adjusted *to* the performance value.

Measures of self-efficacy depend on when they are made. One construct related to temporal variations in an individual's self-efficacy is 'test anxiety' about assessments such as assignments, quizzes, group presentations and the final examination. By far students get most anxious over higher stake tests, such as end of semester examinations (Zoller & Ben-Chaim, 1989). In a large meta-analysis of 562 studies, Hembree (1988) concluded that test anxiety is inversely related to self-efficacy, a finding more recently confirmed by Ruthig, Perry, Hall and Hladkyj (2004). In addition, in another meta-analysis of 151 studies, Hembree (1990) found that with respect to causality, it is test anxiety that causes poor performance rather than previous poor performance causing test anxiety. Short and long time scale changes are also evident in test anxiety (Hembree, 1988). Spielberger, Gorsuch, Lushene, Vagg, and Jacobs (1983) found that students studying to become science teachers experienced a decrease in overall test anxiety from their first to second year at university, but still had increased levels of test anxiety before tests.

Self-efficacy and gender

Generally females report lower academic science self-efficacy than males (Pajares, 2002), including in physics (Cavallo et al., 2004). This difference emerges in middle to late primary school (Andre, Whigham, Hendrickson & Chambers, 1999; Pajares, 2002), but there is no consensus in the literature on what causes such gender differences (Dalgety & Coll, 2006). Some studies have found that many gender differences in self-efficacy disappear when previous academic achievement is controlled for (Pajares, 2002). However, Cervone and Palmer (1990) observed that in the absence of prior knowledge, males still reported a statistically significantly higher self-efficacy than females. As experience was gained, this difference declined but was not eliminated by the end of the study. An interesting point to note is that Arch (1987) found that females tended to devalue their performance and in general were more self-critical.

Although Bandura (1997) does not attribute a difference between genders with respect to the relative importance of the four sources of self-efficacy, more recent research suggests that such differences exist. In particular, an extended multiple case study by Zeldin and colleagues (Zeldin et al., 2008; Zeldin & Pajares, 2000) with cross case comparison interviewing 25 males and females in careers in science, technology, engineering, or mathematics uncovered some interesting gender differences. Their key findings were that mastery experiences emerged as the most significant source of self-efficacy for males, whereas vicarious experiences and verbal and social persuasion had been the most important sources for women.

Another indicator of gender difference is suggested to occur in students' physiological and affective states. Most research related to this factor has been conducted in the area of test anxiety. Females self-report higher test anxiety levels in mathematics and science than males, observed from year 3 of primary school (Hembree, 1988). In addition, the 'harder' the subject, the higher the associated test anxiety (i.e., in order of increasing anxiety: biology < physics ≈ chemistry < mathematics) (Zoller & Ben-Chaim, 1989). In a meta-analysis of 30 studies Becker (1989) found that males consistently outperformed females on academic achievement tests in the 'harder' sciences (biology, general science, and physics), but not in the softer sciences (geology and earth sciences). However, she also found that this effect was on average greater for studies that focussed on gender and science, suggesting experimenter effects or publication bias. It should be emphasised, however, that the gender difference occurs both for test anxiety (Hembree, 1988, 1990) and for self-efficacy (Anderman & Young, 1994; Andre et al., 1999; Cervone & Palmer, 1990) even when there is no difference in academic achievement. When interpreting such data it is useful to be aware of Hembree's (1990) meta-analysis of 151 studies in which he found that for high school students (year 7-12) males with high test anxiety were less likely to take more maths courses than females with high test anxiety, thus skewing the gender differences even further.

Pajares (2002) discusses the gender difference in terms of males and females operating with different 'metrics' when self-reporting both test anxiety and self-efficacy. Along similar lines Wigfield, Eccles, and Pintrich (1996) suggest that males and females have different self-reporting standards. If males and females indeed use different metrics, then analyses of self-efficacy and test anxiety need to consider gender in order to provide meaningful interpretation.

Self-efficacy in undergraduate physics

In terms of work done on self-efficacy in first year university physics, two recent studies are of particular interest. Fencl and Scheel (2004, 2005) have investigated how different types of teaching environments (both traditional and non-traditional) affect students' self-efficacy. They found that collaborative learning had the greatest impact, but this finding was stronger for majors than non-majors. Self-efficacy was also found to correlate with expected grades (Fencl & Scheel, 2005) ($r = 0.57$, $p < 0.001$, $N = 218$), not separating by gender.

Shaw (2004) also found gender differences when investigating first year undergraduate students' self-efficacies. She reports on findings separated by gender for three different courses: Concepts of Physics ($N = 365$), a conceptual course assuming little mathematical background; College Physics ($N = 83$), primarily taken by science education and biological science majors, which assumes some familiarity with trigonometry; and University Physics ($N = 80$), which assumes some calculus background and is mostly populated by engineering majors. The self-efficacy questionnaire was administered mid-semester for the concepts of Physics course, at the end of both semesters for the College Physics course, and at the end of second semester for the University Physics course. Only the Concepts of Physics course showed a significant difference between genders, with males having higher self-efficacies than females. In terms of correlations between self-efficacy and course grade, in the Concepts of Physics course only weak correlations were found; in College Physics statistically significant correlations were found in both semesters for females, but at no point for males; and in the University Physics course there were no statistically significant correlations.

Purpose of the study

Current literature strongly suggests that self-efficacy instruments provide better measures if they are specifically aligned with the subject of study (Choi, 2005; Lent et al., 1997). Further, the length of the instrument is critical for two reasons. First, we consider longer questionnaires that aim to identify several different constructs (or factors). Each construct is a combined measure of several items that all need to exhibit a statistical relationship with the intended construct across different administrations of the questionnaire. If that does not happen, then individual scores (called factor scores) cannot be systematically calculated for each construct for comparison across the different administrations of the questionnaire. That is why it is more complex to develop longer instruments than shorter ones. Dalgety and Coll (2006), in their administration of a 17-item chemistry self-efficacy questionnaire to first-year tertiary chemistry students, found that the factor structure varied significantly across three separate administrations of the study, so composite factor scores could not be calculated. Second, the practical length of an in-class questionnaire is constrained by the time allotted to it and by the duration of students' interest. Therefore, if the aim is not to measure different constructs within self-efficacy, a short self-efficacy instrument can be combined with other constructs on one questionnaire. In addition, such a multi-item questionnaire can be used to provide answers to other interesting research questions of which self-efficacy is only one aspect. While we were developing our questionnaire, Gungor, Eryilmaz, and Fakioglu (2007) in Turkey were also developing a short practical self-efficacy scale. Their physics self-efficacy instrument also has five items and is quite general, but has several other constructs being measured in parallel. It is pleasing to see two physics self-efficacy scales developed in parallel studies in Australia and Turkey.

Hence, the first aim of this study was to develop and evaluate a short, one-factor instrument for physics self-efficacy that would result in a single score per individual. No such instrument was found in the literature at the inception of this study. The second aim was to investigate physics self-efficacy of males and females with and without prior formal senior high school physics instruction across one academic year. Such a study was considered to be of great interest since very little self-efficacy research has been carried out on tertiary physics students to date. The third aim was to observe the relationship between physics self-efficacy and academic achievement at different times of the year. A correlational relationship has been reported in literature, but comments on how such a correlation varies with time were not found.

Part I: Development of the Physics Self-Efficacy Questionnaire

The development of the questionnaire went through five distinct phases, as summarised in Table 1. Each phase is briefly described below.

Table 1: Overview of the development process of the Physics Self-Efficacy Questionnaire.

<u>a. Selection of items</u> Existing items and scales were surveyed. Four items were chosen from Jerusalem and Schwarzer's (Schwarzer, 1993) General Self-Efficacy Scale and one item was written based on general perception of all items perused.
↓
<u>b. Validation by experts</u> Three physics education experts critiqued the chosen items. Minor changes were made upon their feedback.
↓
<u>c. Initial trial</u> The questionnaire was trialled with 111 first year physics students in June 2006. Principal components analysis was carried out to check the validity and reliability of the questionnaire.
↓
<u>d. Confirmatory trial</u> The questionnaire was completed by 379 first year physics students in March 2007. A confirmatory factor analysis was carried out to check the validity and reliability of the questionnaire.
↓
<u>e. Invariance and stability</u> The questionnaire was administered to different groups of first year physics students four times throughout 2007. Exploratory factor analysis provided a consistent factor structure.

a. Selection of items

An extensive survey of self-efficacy scales was carried out (Dalgety & Coll, 2006; Redish, Saul & Steinberg, 1998; Schwarzer, Bassler, Kwiatek, Schroder & Zhang, 1997; Tuan, Chin & Shieh, 2005). The chemistry self-efficacy scale by Dalgety and Coll (2006) was given serious consideration since teaching and learning in chemistry and physics have many parallels. However, the scale was not utilised because it was particularly specific on individual aspects of the authors' first year chemistry course and, as stated earlier, the factor structure was inadequate. The Maryland Physics Expectations (MPEX) survey (Redish et al., 1998) and the Colorado Learning Attitudes about Science Survey (CLASS) (Adams et al., 2006) were also considered as they are specific to tertiary physics education. However, CLASS has been 'designed to measure student beliefs about physics and about learning physics' and the MPEX survey has been designed to study students' expectations, attitudes and beliefs towards learning physics. As these do not focus on self-efficacy, they were not used. The style in which the items on the chemistry self-efficacy scale, the MPEX survey, and CLASS are written and the content they cover influenced the development of the Physics Self-Eefficacy Questionnaire.

Both individual items and whole scales found in the literature were perused and critically evaluated to develop a broad base of possible items. Together with Bandura's definition and theories of self-efficacy (Bandura, 1994), we developed a sound understanding of the construct of self-efficacy and how it is measured. Ultimately Jerusalem and Schwarzer's ten-item General

Self-Efficacy Scale (Schwarzer, 1993) was chosen to form the basis of the Physics Self-Efficacy Questionnaire, based on four reasons. First, this short scale is established and is translated into 30 different languages. Second, the items are appropriate to be made specific to our local teaching and learning context. The third reason for choosing the General Self-Efficacy Scale was that it has a focus on student agency, in that all of the items emphasise how students have the ability to act in ways that allow them to improve their performance. Lastly, the General Self-Efficacy Scale had consistently yielded satisfactory internal consistencies across several research projects as measured by Chronbach's alpha between 0.75 and 0.90 (Schwarzer et al., 1997) as well as adequate factor loadings (Schwarzer, 1993).

All of the items in the General Self-Efficacy Scale were scrutinised for adaptability and appropriateness of use in our specific situation and the local teaching and learning context. Four items were chosen: two were subjected to minor changes, and two underwent extensive changes where the items were made relevant to the local context, but still conserved their original intent. One example of an extensive change is as follows: item 7 in the General Self-Efficacy Scale was 'I can remain calm when facing difficulties because I can rely on my coping abilities', which was changed to 'I will remain calm in my physics exam because I know I will have the knowledge to solve the problems'. One additional item was designed, based on the understanding the authors had developed for the concept of self-efficacy (item 5 in Table I). For each item students were asked to indicate on a five-point Likert scale whether they strongly disagreed (1), disagreed (2), were neutral (3), agreed (4), or strongly agreed (5). With a total of five items, the draft questionnaire was short, as intended. As a factor requires at least four items with factor loadings greater than 0.6 (Field, 2000), it was decided to present this version for validation by experts. If they considered the length an issue, more items would be considered.

b. Validation by experts

Jerusalem and Schwarzer's (Schwarzer, 1993) original questionnaire and the five proposed items were given to three experienced physics education experts, one of whom is also an expert in self-efficacy and related constructs. They were asked to comment on the validity of the items. The experts were satisfied with the items suggesting only minor changes, which were incorporated in the final version of the questionnaire (see Table 2).

Table 2: The Physics Self-Efficacy Questionnaire.

	Physics self-efficacy items	*Factor loadings*
1.	I generally manage to solve difficult physics problems if I try hard enough	0.704
2.	I know I can stick to my aims and accomplish my goals in physics	0.821
3.	I will remain calm in my physics exam because I know I will have the knowledge to solve the problems	0.775
4.	I know I can pass the physics exam if I put in enough work during the semester	0.737
5.	The motto 'If other people can, I can too' applies to me when it comes to physics	0.694

c. Initial trial

The questionnaire was administered in-class to first year physics students at the end of the first semester in June, 2006. The authors introduced the voluntary questionnaire, emphasising that responses would not affect results. One hundred and eleven students completed the questionnaire. The first author, who handled the original data files, was not involved in student assessment. The data were analysed by principal components analysis (a type of exploratory factor analysis), using the Statistical Package for the Social Sciences (SPSS) version 15.0. With only five statements in the questionnaire, the sample size was satisfactory (Floyd & Widaman, 1995), and analysis of the data found it suitable for exploratory factor analysis. The condition for factor extraction was based on a combination of Kaiser's criterion of eigenvalue > 1 and an investigation of the Scree plot, both of which clearly indicated one factor only. As there was only one factor, factor rotation did not apply. The five items had factor loadings in the range 0.694 to 0.821 (see Table 2), confirming the intended factor structure (at least four factor loadings over 0.6) (Field, 2000). The factor explained 56% of the variance (values over 50% are acceptable according to Streiner (1994)), and evidence for the reliability of the questionnaire was provided by a Cronbach's α of 0.796.

d. Confirmatory trial

A factor structure needs to be confirmed by a confirmatory factor analysis on an independent data set. The questionnaire was therefore administered to a fresh first year physics cohort in March 2007. The same procedures were carried out as in the previous year. Three hundred and seventy nine students completed the questionnaire, a return rate of 81%. A confirmatory factor analysis (using Amos 7.0) provided evidence for the construct's validity (values in parentheses indicate requirements for validity); $\chi^2 = 2.127$, $p = 0.831$ ($p > 0.05$). Main fit indices also showed a very good model fit (Kline, 2005): RMSEA = 0.000 (< 0.05) with a 90% confidence interval of [0.000, 0.042]; RMR = 0.009 (< 0.05); GFI = 0.998 (> 0.95); NFI = 0.994 (> 0.95); and CFI = 1.000 (> 0.95).

e. Invariance and stability

Final checks on the questionnaire were for invariance and stability. A questionnaire is said to be invariant if the factor structure for data from different samples from the population is consistent. Furthermore, if the factor structure is consistent when the questionnaire is administered at different times, it is said to be stable. Males and females in two different classes were sampled four times in the year (more detail is provided in the next section). No anomalies in the factor structure were found between either gender or times of administration. As the questionnaire was robust the data used for the invariance and stability checks were further examined for trends in student self-reports of self-efficacy.

Part II: Self-efficacy for females and males with and without senior high school physics instruction

The Sample

Two different first year physics classes at the University of Sydney were sampled in first semester, 2007. The Fundamentals (FND) and Regular (REG) classes are designed to cater for students' prior formal instruction in physics. The Fundamentals class (N = 234) is designed for students with no prior formal senior high school instruction in physics. It covers skills and methods while teaching mechanics and waves. The Regular class (N = 351) assumes two years of senior high school physics background and covers mechanics, waves, and thermal physics, where the first two topics are covered more deeply than in the Fundamentals class.

In second semester the students from the Fundamentals and Regular classes have a choice of either enrolling in classes based on their interest or not continuing with physics. Approximately equal numbers of Fundamentals and Regular students enrolled in the larger second semester class, called the Environmental class (N = 246). Hence, the Environmental class was sampled in second semester 2007. The Environmental class focuses on aspects that are relevant to environmental and life sciences, covering properties of matter, electromagnetism, and modern physics, and has been specifically designed for this merging of students (Au & Sharma, 2007). Note that in this paper students carry the labels of FND and REG for *the whole year*, even though they are merged in second semester. This is done because the labels reflect the students' high school physics background, which is a variable of interest in this study.

The structure and assessment of all the classes are similar. Each semester has 13 teaching weeks and one examination study week followed by two examination weeks. During each teaching week students attend three one-hour lectures, one one-hour tutorial, and one three-hour laboratory. The summative assessment is through assignments, collaborative laboratory work, and participation in collaborative tutorials, together with a final examination held during the examination weeks. Even though the structure and assessments are fairly standard, interactive practices are embedded into the curricula (Lindstrøm & Sharma, 2009).

Data Collection

The self-efficacy questionnaire was administered in weeks 3 and 13 of first semester and again in weeks 3 and 13 of second semester. For logistic reasons, the week 3 administrations were in lectures while those in week 13 were in tutorials. The self-efficacy items were the first five questions of a two-page questionnaire addressing various aspects of the classes (students' attitudes towards physics, achievement goal orientations in physics, and opinions of the tutorials offered). For each questionnaire the students were given a short (three-minute) talk by one researcher informing them of the purpose of the research and the privacy protocols. The return rates were (78-91)% for the students attending the lecture or tutorial, which corresponded to (53-61)% for all enrolled students.

Informed consent was obtained from students, and data on year 12 course selections and gender were obtained from records. Only those students in the Fundamentals class *with known* non-physics background and those in the Regular class *with known* physics background are included in the ensuing analysis. These explain why the number of students in any analysis is smaller than the total number of students who completed the course requirements.

Analysis and results

To use the newly created questionnaire we decided to apply unit weighting to each item to produce factor scores for each individual student. In cases where questionnaires are administered and analysed beyond the original sample, this is indeed the recommended method (Gorsuch, 1983). In other words, students' responses to each item were coded as follows: strongly disagree was coded as 1, disagree as 2, neutral as 3, agree as 4, and strongly agree as 5. Each student's response was summed over all 5 items. Physics self-efficacy scores thus ranged between 5 (lowest) and 25 (highest).

To investigate whether students were interpreting the questionnaire in a similar manner across different administrations, we decided to look for correlations of self-efficacy scores between administrations. As none of the self-efficacy distributions was normal, Spearman's ρ (which is a non-parametric correlation analysis) was carried out. The statistic Spearman's ρ is interpreted in the same manner as Pearson's r.

There were large correlations between the pair of questionnaires (pre and post, not separating by class) administered in the first semester ($\rho = 0.61$, $N = 193$, $p < 0.001$) and those in the second semester ($\rho = 0.68$, $N = 92$, $p < 0.001$). The correlations between any two questionnaires administered in different semesters were smaller ($\rho = 0.40$ to 0.51, $N = 88$ to 108, $p < 0.001$ for all). There was more internal consistency between two questionnaires from the same semester, as opposed to two questionnaires completed in different semesters, as expected.

Table 3 shows the means and standard deviations of the self-efficacy scores for females and males in each class at each administration. We see four interesting features. First, females consistently self-report lower self-efficacies than males; second, Fundamentals students report lower self-efficacies than Regular students in second semester, but not so in first semester; third, there is a marked pre-examination drop in self-efficacy in first semester, but not so in second semester; and fourth, upon entry the males in the Fundamentals class report the highest self-efficacy of any group at any administration. The first two of these findings were investigated by conducting four two-way between-groups ANOVAs (one for each time self-efficacy was measured), including both gender and prior formal instruction in physics. The last two findings were based on t-tests and perusing trends in Table 3.

Females consistently self-report lower self-efficacies. Looking at the impact of gender in the four two-way between-groups ANOVAs, three of the four analyses revealed a statistically significant main effect (the measurement early in second semester bordered significance). The effect sizes measured by η^2 indicated a medium effect (according to Cohen (1988), a small effect size is around 0.01, medium is 0.06 and large is 0.14). [$F(1,277) = 17.396$, $p < 0.001$, $\eta^2 = 0.059$; $F(1,267) = 32.123$, $p < 0.001$, $\eta^2 = 0.107$; $F(1,118) = 3.219$, $p = 0.075$, $\eta^2 = 0.027$; $F(1,141) = 9.867$, $p = 0.002$, $\eta^2 = 0.065$].

TABLE 3: The means and standard deviations (SD) of the self-efficacy scores for males and females in each class at each administration. Only students with no formal instruction in senior high school physics are included in the Fundamentals (FND) class, while only those with formal instruction in senior high school physics are included in the Regular (REG) class. The two classes were taught separately in first semester but merged into the Environmental class in second semester. The Fundamentals and Regular labels are retained to reflect students' senior high school physics experience.

		Early semester 1			End semester 1			Early semester 2			End semester 2		
Gender	Class	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Female	FND	51	17.76	3.12	66	16.88	2.52	38	17.18	2.64	43	16.93	3.13
	REG	66	18.06	3.54	59	16.83	3.04	30	18.17	2.67	36	18.47	2.41
Male	FND	39	19.90	2.91	52	18.87	2.80	26	17.88	2.96	31	18.58	3.01
	REG	125	19.20	2.65	94	18.60	2.41	28	19.14	1.80	35	19.74	2.45

Fundamentals students report lower self-efficacies than Regular students in second semester, but not so in first semester. In the four two-way between-groups ANOVAs the main effect for prior formal instruction in physics did not reach statistical significance in first semester [$F(1,277) = 0.262$, $p = 0.609$; $F(1,267) = 0.231$, $p = 0.631$] when the students were separated in the Fundamentals and Regular classes, but it was statistically significant in second semester when all students were enrolled in the same class [$F(1,118) = 5.750$, $p = 0.018$, $\eta^2 = 0.046$; $F(1,141) = 8.456$, $p = 0.004$, $\eta^2 = 0.057$]. Medium effect sizes were observed for prior formal instruction in physics in second semester.

There is a pre-examination drop in self-efficacy. Paired-samples t-tests showed that the pre-examination drops seen in first semester were all significant, except for that of Regular males. Note that the sample sizes were somewhat smaller than those in Table 3 since only those students who responded to *both* first semester questionnaires were included [FND females: mean (early) = 17.92, mean (end) = 17.18, N = 39, t = 2.048, $p = 0.048$; FND males: mean (early) = 19.91, mean (end) = 18.64, N = 33, t = 2.469, $p = 0.019$; REG females: mean (early) = 18.23, mean (end) = 16.63, N = 48, t = 4.418, $p < 0.001$; REG males: mean (early) = 18.81, mean (end) = 18.74, N = 73, t = 0.242, $p = 0.809$]. In second semester there were no pre-examination drops, confirmed by t-tests, which revealed no statistically significant differences.

Males in the Fundamentals class report very high self-efficacies. When comparing all the mean self-efficacy values in Table 3, Fundamentals males at the very beginning of the year reported the highest values, while Regular males at the end of the year reported the second highest values.

Finally, the interaction effect between gender and prior formal instruction in physics did not reach statistical significance at any time [$F(1,277) = 1.603$, $p = 0.207$; $F(1,267) = 0.112$, $p = 0.738$; $F(1,118) = 0.087$, $p = 0.768$; $F(1,141) = 0.167$, $p = 0.684$].

Exploring selection effects: Two questions arise from the results shown above. Is it possible that female students with low self-efficacy in physics are leaving the courses *during* the semester? Is it possible that female students with low self-efficacy in physics are self-selecting out *after* the first semester? In-depth analyses demonstrate that this is not so. In both cases, the mean self-efficacies for females continuing are not statistically significantly higher than for those not continuing. The same applies to males.

Part III: Study of physics self-efficacy and academic achievement across one year

Additional data collected

A subset of the sample described in Part II was used in Part III. Academic achievement was measured using the end of semester examinations. Assignment and laboratory marks were not used because they are group work efforts, which do not reflect individual achievements. The examinations are three hours long with 12 questions. The first six questions (30 marks) are conceptual questions while the remaining six (60 marks) are more traditional questions requiring both calculations and interpretation of answers.

Checks on academic achievement and self-efficacy data

Three checks were carried out on the data: correlations between the two academic achievement measures (the first and second semester examination scores), tests for statistical difference between examination scores for males and females, and correlations between self-efficacy and examination scores for various subgroups (separating by gender and prior formal instruction in physics).

The first and second semester examination scores were first checked for internal consistency using Pearson's correlation coefficient r, as the data were normal. There were large and statistically significant correlations between the examination scores (FND: $r = 0.75$, $N = 105$, $p < 0.001$; REG: $r = 0.77$, $N = 113$, $p < 0.001$) demonstrating internal consistency.

Next we checked whether the gender difference in self-efficacy seen in Table 3 could be explained by gender difference in academic achievement. When we conducted t-tests, the only statistically significant difference between genders was for the Fundamentals students (males: mean = 46.9, $N = 74$, SD = 14.5; females: mean = 42.5, $N = 92$, SD = 14.1; $t = 1.987$, $p = 0.049$). Note, however, that this difference was not observed in second semester. Hence the gender difference in self-efficacy cannot simply be explained by different levels of knowledge, as measured by the examination results.

Lastly, we needed to see if our results confirm current understandings, as the literature generally suggests a robust correlation between self-efficacy and academic achievement. The effect of prior formal instruction in physics was expected to be the strongest in the correlations between self-efficacy and academic achievement in first semester. For the Fundamentals students, there was no correlation at any time during the first semester, neither when analysed for the whole class, nor when split by gender. The Regular class and the second semester correlations were more complicated and at the same time interesting. Consequently only those students who completed both semesters were studied as reported below.

Results and analysis

The group of students who sat either the Fundamentals or Regular examination in first semester *and* sat the Environmental examination in second semester were identified to allow us to track one particular group for the whole year. This is the sample for the study of self-efficacy and academic achievement. Note that the students did not need to have responded to all four questionnaires to be included in the analysis. The mean self-efficacy scores for males and females are presented in Figure 1. Note that the standard error of the mean is about ±0.5 for all means. Since the trends mirror those in Table 3, the smaller samples are adequate for interpretation using correlations with examination scores. Two sets of correlations between self-efficacy and examination scores were conducted as described below.

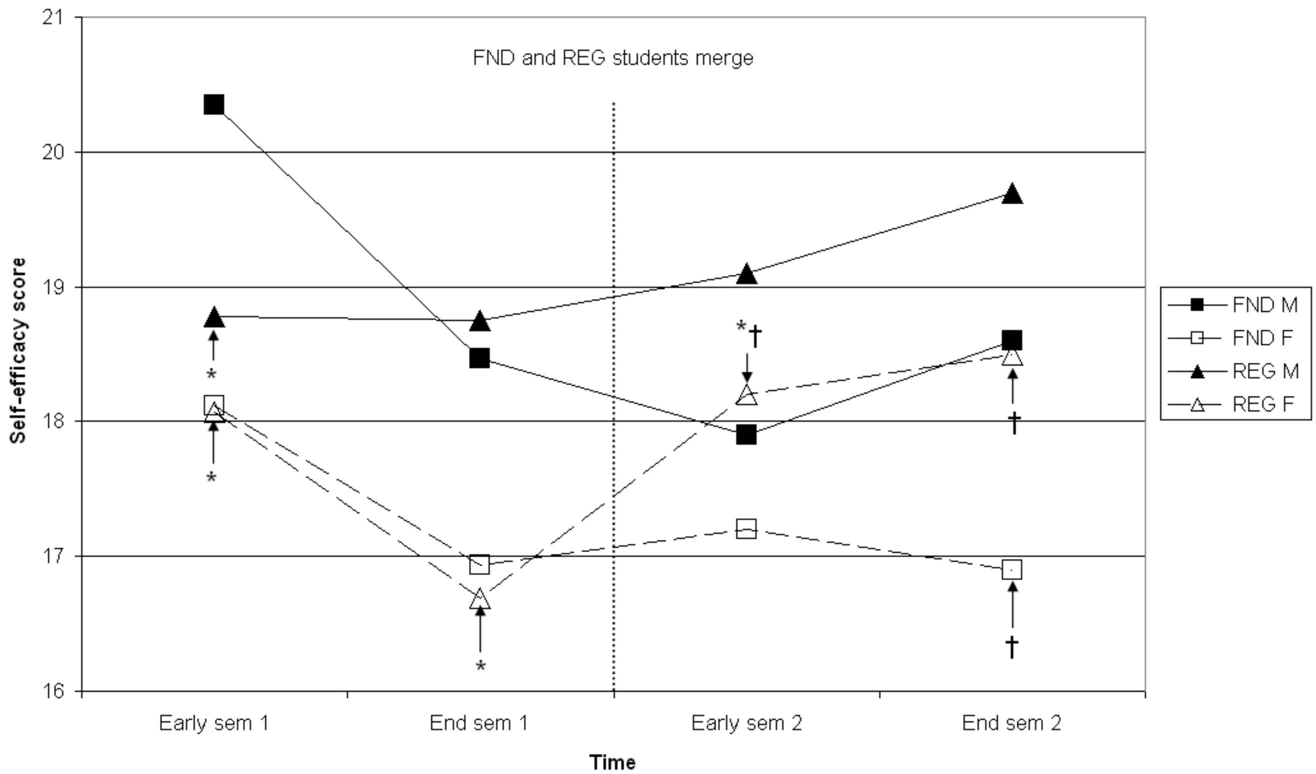

Figure 1: The mean self-efficacy scores for females and males who had completed both semesters. Each mean has a standard error of the mean of about ±0.5. A star indicates a statistically significant correlation between self-efficacy scores and first semester examination scores. A dagger indicates a statistically significant correlation between self-efficacy scores and second semester examination scores.

Correlations with first semester examination scores. We calculated correlation coefficients for self-efficacy measures from the first three questionnaires with the first semester examination scores. Each point that showed a statistically significant correlation is marked by a star on the graph (Figure 1). Neither the Fundamentals males nor females showed any correlations, but there was a significant correlation between self-efficacy and examination score among Regular females in all three cases, but only at the beginning of the year for Regular males.

Correlations with second semester examination scores. Early and end second semester self-efficacy scores were correlated with second semester examination scores. Each data set that showed a statistically significant correlation is marked by a dagger on the graph. In this case there was *one* occurrence of a correlation for the Fundamentals class, namely for females at the end of the semester. Correlations early in the second semester were only seen among the Regular females, while females from both classes showed correlations at the end of the semester.

In summary, Regular females consistently show correlations between self-efficacy and examination scores throughout the year. The Regular males, on the other hand, exhibit correlations only occasionally, whereas only one statistically significant correlation was seen in the Fundamentals class. A further check was carried out to see whether the same trends were found when analysing the data for only those students who responded to all four questionnaires and had sat the two examinations, emulating a longitudinal study. In this case the sample sizes for the various groups were: N(FND females) = 16, N(FND males) = 10, N(REG females) = 17, and N(REG males) = 14. The same trends did indeed emerge, supporting the above findings.

Discussion

The aims of this study were to develop and evaluate a Physics Self-Efficacy Questionnaire that was easy to use, and to analyse self-efficacy and academic achievement over a year of physics study. The data were analysed both with respect to gender and prior formal instruction in physics.

Four items on the Physics Self-Efficacy Questionnaire were adapted to a physics context from Jerusalem and Schwarzer's (Schwarzer, 1993) General Self-Efficacy Scale, and one item was generated by the authors. Exploratory factor analysis was conducted on the first administration of the questionnaire in 2006, and a confirmatory factor analysis was conducted on the second administration in 2007, providing evidence for the questionnaire's construct validity.

In agreement with the literature (Cavallo et al., 2004), it was found that females consistently reported lower self-efficacies than males (regardless of prior formal instruction in physics) even though their academic achievements were generally not statistically significantly different from males'. This supports the different 'metric' theory discussed by Pajares (2002). The data also suggested that females experienced lower self-efficacies prior to the examination in first semester, but not in second semester, which is in agreement with the findings of Spielberger et al. (1983) that test anxiety reduces as students experience more examinations. For males, the major result was that the Fundamentals students at the beginning of first semester exhibited the highest mean physics self-efficacy seen in this study. The finding possibly indicates the existence of a 'male overconfidence syndrome' in males before they have experienced any significant formal instruction in physics, which is the case in week 3 of first semester. This is not unknown in the literature (Cervone & Palmer, 1990). Specifically studying this aspect is worthy of further investigation. No equivalent extreme self-efficacy score was observed for females.

Correlations between physics self-efficacy and students' academic achievements showed the greatest departure from the general literature, although there have been similar findings in tertiary physics (Shaw, 2004). Students without senior high school physics exhibited no correlation between the two measures, with the exception of females by the end of second semester. Our findings suggest that although self-efficacy in novices should be tentative only and easily adjusted when receiving feedback (Cervone & Palmer, 1990), this change in self-efficacy appears to be much slower than expected due to a surprisingly large resistance to change in the face of clear counter-evidence (Lepper et al., 1986). While many of these students are novices in a formal physics classroom setting, all of our students have ideas about how the world works, and they all have some experiences that allow them to think about whether they are 'good' at science or not. These, in addition to the social persuasion of teachers and others, and the vicarious learning of peers, tend to inform self-efficacy in a very powerful way. This may or may not account for the surprisingly large resistance to change.

Drawing all these findings together, the results seem to suggest that when analysing by gender and senior high school physics instruction, both have similar effects on students' self-efficacy. However, no interaction effect between these two variables using a two-way ANOVA was detected. In terms of correlations between self-efficacy and academic achievement, females with prior formal physics instruction exhibit statistically significant correlations, but males (regardless of experience) for the most part do not. This may indicate that females are more receptive to feedback and adjust their self-efficacy accordingly. This idea is supported by the findings for the Fundamentals students in which the females change from performing substantially more poorly than the males in first semester, to being academically on par with the males in second semester. Seemingly, they responded to feedback and changed their self-efficacy. The improved relative performance in semester two also suggests that certain measures were taken with respect to study strategies and/or efforts. The intention of the Fundamentals class to get novices up to speed in one semester appears to have been successful, although it worked better for the women. If the findings of Zeldin et al. (Zeldin et al., 2008; Zeldin & Pajares, 2000) generalise to our study in that self-efficacy tends to be informed in males and females by different forms of feedback, this may be a possible explanation for the gender differences observed. Alternatively, it may be that males and females respond to the questionnaire via different 'metrics', as discussed by Pajares (2002), or that the instrument somehow better addresses the self-efficacy of females than males. These speculations open up directions for further investigation, which can provide interesting insights. We found no other studies with which to compare these interpretations and speculations as there are only a few papers on resistance to change, none of which discuss gender differences. However, results using CLASS (Adams et al., 2006) have found that the level of prior instruction affects students' personal beliefs about physics and about learning physics (Gray et al., 2008). The CLASS data also suggest that 'while women are better at identifying what ideas physicists believe, they are less inclined to feel that these ideas are valid or relevant for their experiences' (Gray et al., 2008, p. 8). Hence, the importance of considering prior formal instruction in physics and gender in tertiary physics has been recognised in parallel by other research groups.

Both a strength and a shortcoming of this study is the length of the questionnaire. A short questionnaire is more practical to implement and easier to analyse but does not provide detail on the different facets of self-efficacy. Targeted interviews on the items and various aspects like feedback would better substantiate speculative ideas presented in this paper. Further studies could focus on larger sample sizes striving for higher response rates, and comparing self-efficacy across different groups of students and institutions. Finally, for logistic reasons the questionnaire was administered in week 3 measuring students' self-efficacy early, but not right at the beginning of the semester. Often when surveying students' incoming beliefs or motivations, timing is crucial and it would be interesting to study students' self-efficacy prior to any exposure to tertiary physics.

Implications for the teaching and learning context

If our interpretations of the findings are correct, the implications for teaching and learning bring us into an educational minefield. The gender issue has been a long-standing hot topic with strong opinions in both camps. However, if we listen to the findings of this study, rather than the politics of the debate, we see two issues that need to be addressed.

First, the findings suggest that the type of feedback required by males and females is different. Since self-efficacy is a measure of a person's belief in his or her ability to perform a certain task, once a person has passed through several rounds of examinations and subsequent feedback, one would expect to observe a correlation between achievement and self-efficacy. We find that this is truer for females than for males. The first year students described in this paper experience many different types of feedback. There is group work in lectures, laboratories and tutorials (Sabella, 1999; Sharma, Khachan, Chan & O'Byrne, 2005; Sharma, Mendez & O'Byrne, 2005; Sharma, Millar & Seth, 1999) where feedback is via social persuasion, watching peers work, and collaborative problem solving. Such interactive learning has been found by some researchers to benefit the academic achievement of females more than males (Lorenzo, Crouch & Mazur, 2006), and is in agreement with Zeldin et al.'s finding that social persuasion and vicarious learning more strongly affects the self-efficacy of females than males (Zeldin et al., 2008; Zeldin & Pajares, 2000). We also provide feedback in terms of mastery experiences with a range of problem-solving tasks in assignments (and examinations). Our findings then suggest the question: are we providing feedback in a way that better addresses females, or are males just more resistant to change? This phenomenon is not understood well enough that we would suggest any changes in educational policy. However, it is worthy of further investigation, and educators may do their students a service by being aware and observant of any gender differences in their students' responses to varying forms of feedback.

Second, the already well documented observation that females report a lower self-efficacy than males when their achievements are comparable deserves some serious attention. *If* self-efficacy affects students' choice of further study in the same way for males and females, then we may need to reconsider how we address students when encouraging further physics studies. More research is needed to elaborate on this issue, and further research must explicitly separate males and females in their analyses.

Acknowledgements

The authors would like to thank the Sydney University Physics Education Research Group for valuable discussions and comments on drafts of the paper, and the anonymous referees who provided extensive feedback, which helped to improve the paper.

References

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research, 2*(010101), 1-14.

Anderman, E. M., & Young, A. J. (1994). Motivation and strategy use in science: Individual differences and classroom effects. *Journal of Research in Science Teaching, 31*(8), 811-831.

Andre, T., Whigham, M., Hendrickson, A., & Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *Journal of Research in Science Teaching, 36*(6), 719-747.

Arch, E. C. (1987). Differential responses of remales and males to evaluative stress: Anxiety, self-esteem, efficacy, and willingness to participate. In R. Schwarzer, H. M. van der Ploeg & C. D. Spielberger (Eds.), *Advances in test anxiety research* (Vol. 5, pp. 97-106). Hillsdale, NJ: Erlbaum.

Au, T., & Sharma, M. D. (2007). Student performance in first-year physics: does high school matter?

Bandura, A. (1994). Self-Efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior* (Vol. 4, pp. 71-81). New York: Academic Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Becker, J. B. (1989). Gender and science achievements: A reanalysis of studies from two meta-analyses. *Journal of Research in Science Teaching, 26*(2), 141-169.

Cantor, N., & Kilhstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Cavallo, A. M. L., Rozman, M., & Potter, W. H. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors. *School Science and Mathematics, 104*(6), 288-300.

Cervone, D. (2000). Thinking about self-efficacy. *Behavior Modification, 24*(1), 30-56.

Cervone, D., & Palmer, B. W. (1990). Anchoring biases and the perserverance of self-efficacy beliefs. *Cognitive therapy and research, 14*(4), 401-416.

Choi, N. (2005). Self-Efficacy and Self-Concept as Predictors of College Students' Academic Performance. *Psychology in the Schools, 42*(2), 197-205.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. New York: Erlbaum.

Dalgety, J., & Coll, R. K. (2006). Exploring First-Year Science Students' Chemistry Self-Efficacy. *International Journal of Science and Mathematics Education, 4*(1), 97-116.

Fencl, H., & Scheel, K. (2004). Pedagogical approaches, contextual variables, and the development of student self-efficacy in undergraduate physics courses. In J. Marx, S. Franklin & K. Cummings (Eds.), *2003 Physics Education Research Conference: AIP Conference Proceedings (720)* (pp. 173-176). Melville, NY: AIP.

Fencl, H., & Scheel, K. (2005). Engaging students. *Journal of College Science Teaching, 35*(1), 20-24.

Field, A. (2000). *Discovering Statistics Using SPSS for Windows*. London: Sage Publications Ltd.

Floyd, F. J., & Widaman, K. F. (1995). Factor Analysis in the Development and Refinement of Clinical Assessment Instruments. *Psychological Assessment, 7*(3), 286-299.

Gire, E., & Jones, B. (2009). Characterizing the epistemological development of physics majors. *Physical Review Special Topics - Physics Education Research, 5*(010103), 1-6.

Gorsuch, R. L. (1983). *Factor analysis*. Hillsdale, N.J.: Erlbaum Associates.

Gray, K. E., Adams, W. K., Wieman, C. E., & Perkins, K. K. (2008). Students know what physicists believe, but they don't agree: A study using the CLASS survey. *Physical Review Special Topics - Physics Education Research, 4*(020106), 1-10.

Gungor, A., Eryilmaz, A., & Fakioglu, T. (2007). The relationship of freshmen's physics achievement and their related affective characteristics. *Journal of Research in Science Teaching, 44*(8), 1036-1056.

Hackett, G. (1995). Self-efficacy in career choice and development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 232-258). New York: Cambridge University Press.

Hembree, R. (1988). Correlates, Causes, Effects, and Treatment of Test Anxiety. *Review of Educational Research, 58*(1), 47-77.

Hembree, R. (1990). The Nature, Effects, and Relief of Mathematics Anxiety. *Journal for Research in Mathematics Education, 21*(1), 33-46.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd ed.). New York: The Guilford press.

Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Discriminant and Predictive Validity of Academic Self-Concept, Academic Self-Efficacy, and Mathematics-Specific Self-Efficacy. *Journal of Counseling Psychology, 44*(3), 307-315.

Lepper, M. R., Ross, L., & Lau, R. R. (1986). Persistence of inaccurate beliefs about the self: Perseverance effects in the classroom. *Journal of Personality and Social Psychology, 50*, 482-491.

Lindstrøm, C., & Sharma, M. D. (2009). Link maps and map meetings: Scaffolding student learning. *Physical Review Special Topics - Physics Education Research, 5*(1), 010102.

Lorenzo, S., Crouch, C., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics, 74*(2), 118-122.

McInerney, D. M., & McInerney, V. (2002). *Educational Psychology: Constructing Learning* (3rd ed.). Frenchs Forest, NSW: Pearson Australia.

Otero, V. K., & Gray, K. E. (2008). Attitudinal gains across multiple universities using the Physics and Everyday Thinking curriculum. *Physical Review Special Topics - Physics Education Research, 4*(020104).

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into practice*, *41*(2), 116-125.

Palmer, D. (2006). Durability of Changes in Self-efficacy of Preservice Primary Teachers *International Journal of Science Education*, *28*(6), 655-671.

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

Ruthig, J. C., Perry, R. P., Hall, N. C., & Hladkyj, S. (2004). Optimism and Attributional Retraining: Longitudinal Effects on Academic Achievement, Test Anxiety, and Voluntary Course Withdrawal in College Students. *Journal of Applied Social Psychology, 34*(4), 709-730.

Sabella, M. S. (1999). *Using the context of physics problem solving to evaluate the coherence of student knowledge.* University of Maryland Physics Education Research Group, College Park, Maryland.

Sawtelle, V., Brewe, E., & Kramer, L. (in press). An exploratory qualitative study of the proximal goal setting of two introductory modeling instruction physics students *2009 Physics Education Research Conference: AIP Conference Proceedings*.

Schwarzer, R. (1993). Measurement of perceived self-efficacy. Psychometric scales for cross-cultural research. Berlin, Germany: Freie Universität Berlin.

Schwarzer, R., Bassler, J., Kwiatek, P., Schroder, K., & Zhang, J. X. (1997). The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the General Self-efficacy Scale. *Applied Psychology: An international review, 46*(1), 69-88.

Sharma, M. D., Khachan, J., Chan, B., & O'Byrne, J. W. (2005). An investigation of the effectiveness of electronic classroom communication systems in large lecture classes. *Australasian Journal of Educational Technology, 21*(2), 137-154.

Sharma, M. D., Mendez, A., & O'Byrne, J. W. (2005). The relationship between attendance in student-centred physics tutorials and performance in university examinations. *International Journal of Science Education, 27*(11), 1375-1389.

Sharma, M. D., Millar, R., & Seth, S. (1999). Workshop tutorials: accommodating studentcentred learning in large first year university physics courses. *International Journal of Science Education, 21*(8), 839-853.

Shaw, K. A. (2004). The development of a physics self-efficacy instrument for use in the introductory classroom. In J. Marx, S. Franklin & K. Cummings (Eds.), *2003 Physics Education Research Conference: AIP conference proceedings (720)* (pp. 137-140). Melville, NY: AIP.

Speilberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait anxiety inventory STAI (Form)*. Palo Alto: Consulting Psychology Press.

Streiner, D. L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry, 39*, 135-140.

Tuan, H.-L., Chin, C.-C., & Shieh, S.-H. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education, 27*(6), 639-654.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Wigfield, A., Eccles, J. S., & Pintrich, P. R. (1996). Development between the ages of 11 and 25. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 148-185). New York: Simon & Schuster Macmillan.

Zeldin, A. L., Britner, S. L., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 45*(9), 1036-1058.

Zeldin, A. L., & Pajares, F. (2000). Against the odds: self-efficacy beliefs of women in mathematical, scientific, and technological careers. *American Educational Research Journal, 37*(1), 215-246.

Zoller, U., & Ben-Chaim, D. (1989). Interaction between examination type, anxiety state, and academic achievement in college science; an action-oriented research. *Journal of Research in Science Teaching, 26*(1), 65-77.