Predicting Student Success in Statics

Kate F. Wilson[a], David J. Low[b]
School of Engineering and Information Technology, UNSW Canberra [a], School of Physical, Environmental and Mathematical Sciences, UNSW Canberra [b]
Corresponding Author's Email: k.wilson@adfa.edu.au

Structured Abstract

BACKGROUND
There is increasing pressure in Australian universities to ensure that students successfully complete their degree programs. The early and accurate identification of students who are at significant risk of failure is thus very desirable. Previous studies have found that university entry scores are a good predictor of student success, but the effects of age, gender, social integration and factors such as study skills are unclear, and may be sensitive to degree and institution (see, for example, McKenzie and Schweitzer 2001, Zhang *et al* 2004, Sheard 2009, Litzler and Young, 2012, Hacker *et al* 2013). In this study we look specifically at predictors of student success in Statics, as it is a foundation course for several engineering programs including civil, mechanical and aeronautical engineering.

PURPOSE
Our aim in this study was to establish what simple demographic factors might be predictors of student success in Statics, and to measure the usefulness of various diagnostic tests.

DESIGN/METHOD
Demographic data was collected for the 2013 and 2014 student cohorts, particularly gender, and the state and year in which they finished high school. Age and gender was also available for the 2011 and 2012 cohorts. Students were given a diagnostic test in their first week of statics, as well as different diagnostic tests in their physics and mathematics courses in 2013 and 2014.

We analysed the correlation of final statics marks with the results of the three different diagnostic tests, and the students' final physics and mathematics marks. We also looked at how different demographic groups (males and females, older and younger students) performed on these tests, compared to their final course marks.

RESULTS
The correlation between the final mark in Statics and any of the diagnostic tests was lower than expected, and was particularly low for the physics diagnostic test. The correlation of final subject marks with the Statics final mark was much higher, with a strong correlation between the final Statics and Mathematics marks.

Students who were more than two years out of high school had as good or better final marks, and higher normalised learning gains on the physics test, even though they achieved lower marks on the various diagnostic tests. There were significant differences in the performances of male and female students, with females achieving lower scores in the diagnostic tests, and poorer final marks overall.

CONCLUSIONS
Compared to diagnostic testing, demographic information may be just as useful, or even more so, at predicting future success and in identifying students at risk of failure. Based on these results, it appears that the skills and personal attributes that students bring to their university studies, such as time management and life experience, may be just as important as their prior knowledge of relevant areas such as mathematics and physics. In light of this, we propose that effective early intervention and remedial programs should include tuition in effective study skills together with 'content'.

KEYWORDS
Diagnostic tests, student demographics, early intervention, remedial programs, 'at risk' students

Introduction

There is increasing pressure on students in Australian universities to complete their degree programs quickly, due to the large fees charged for many courses and the significant debts that students accrue for their degrees. At the University of New South Wales at the Australian Defence Force Academy (UNSW Canberra), students undertake military training concurrently with their academic studies. The Department of Defence recruits these students, pays them a salary, and covers their tuition fees; in exchange, the students accumulate a "return of service" obligation based on the length of their degree program. Hence while these students will not leave university with a debt for their degree, there is still significant career pressure on the students to complete their degree in minimum time.

Student progress at UNSW Canberra is more closely monitored than is usual at large research intensive universities: academic staff are required to report student academic issues to military staff. There is thus institutional pressure to identify students "at risk" of failure early, and give them every opportunity for academic improvement. Naturally, this pressure must be balanced against the other commitments of academic staff: ideally, the early identification of "at risk" students should not create a significant assessment burden. Similarly, to be effective, the assistance provided to these students should be targeted to their individual needs, to avoid wasting both staff and student time.

In 2012, UNSW Canberra introduced diagnostic tests for core engineering courses, in an attempt to identify at risk students. Remedial tutorials and/or additional assignment work was provided for students who did not pass these tests. This imposed an additional load on academics, and in many subjects it was unclear whether the remedial work was effective. The usefulness of the diagnostic tests was questionable: while the diagnostics could identify students with significant gaps in their knowledge, or an inability to apply that knowledge, they could not predict how well students would be able to address any identified deficiencies.

Some previous studies (e.g. Power *et al* 1987; Trigwell *et al* 2013) have found that university entry score is a good predictor of student success in various degree programs, including engineering programs (e.g. Zhang *et al*, 2004; Hacker *et al* 2013). Thomas *et al* (2009) found, in a study involving four Australian institutions, that entry score was a poor predictor of grades in mechanics courses, but a risk factor for students with low entry scores. However the minimum entry score for engineering at UNSW Canberra is high, and the range of scores is small, which reduces the usefulness of these scores as a discriminating predictor or a risk factor. In addition, these scores are not easily accessible to academics, and most mature-age students do not enter their degree program via such a score.

Other studies have looked at the effects of age, gender, motivation, social integration and factors such as study skills on student success. The effects of age are unclear, and may be sensitive to degree and institution, with some studies finding age to correlate positively with high grades and student completion (e.g. Hacker *et al* 2013; Sheard 2009) but others finding no correlation (McKenzie and Schweitzer 2001) or a negative correlation (de Koning *et al* 2012). Gender is already known to be a significant issue in engineering and the physical sciences, with females making up only a small fraction of the enrolment in these subjects. In some disciplines, females have been seen to outperform males relative to entry-level measures of ability (Sheard 2009). However, this does not appear to be the case in engineering, where Zhang *et al* (2004) and Litzler and Young (2012) and found that females underperformed compared to males; more recently, Hacker *et al* (2013) found that gender was not significant.

Motivation has been generally shown to correlate positively with student success and retention (e.g. French *et al* 2005). Motivation interacts with other factors such as social integration, and impacts on students' marks via the time and effort put into study. Social integration has also been shown to have complex effects on student success and retention. Social isolation has been shown to be a predictor of attrition, with socially isolated students and students who do not feel a sense of "belonging" at their institution more likely to drop out of their studies (Gerdes and Mallinckrodt 1994; French *et al* 2005).

Hence, in addition to academic factors such as prior performance, demographic factors such as age and gender, and less tangible factors such as motivation and social integration, all play a part in determining how well a student succeeds in their chosen courses.

Our aim in this study was to establish which of these factors are useful predictors of students' later performance in Statics, which is a foundation course for several engineering programs at UNSW Canberra, and attracts a significant fraction of the student intake. The course is a fairly traditional one, and it is not the aim of the current work to investigate the effects of changing the structure of course, although this may be investigated in the future. At present, we hope that by identifying and appropriately supporting students at risk of failing at this early stage, it will help them throughout their degree program, beginning with the Statics course.

The Student Cohort

The Australian Defence Force Academy (ADFA) recruits students from across all states and territories in Australia. Approximately 90% of the students are officer cadets and midshipmen, who undertake military training and education concurrently with their full-time degree programs at UNSW Canberra. Almost all of these students arrive directly from school, or with a single gap-year. The cadets and midshipmen live on-campus, in communal accommodation blocks. Collaboration is generally encouraged both in their military and academic training. Between their living arrangements, military training and study, students spend a great deal of time together. Social isolation, even if desired, would be difficult to achieve in this environment. The remainder of the cohort comprises mature-age serving officers in the ADF, and a few civilians on scholarships from Defence-related organisations.

The fraction of females in the intake is around 30% overall, and typically only around 10% in engineering. These numbers are both significantly lower than would be found at a civilian university. Students from rural and regional areas comprise almost half of the intake (which is relatively very high), and every state and territory is significantly represented in the intake.

A "First-Year Experience" survey based on the James *et al* (2010) survey was administered in 2013. It indicated that, compared to the first year cohort across all Australian universities, UNSW Canberra students feel a stronger sense of belonging to the university community. Their primary sense of "belonging", however, is to the military. The same survey also found that UNSW Canberra students were generally less motivated towards their study. This is perhaps not surprising, given that the cadets and midshipmen are primarily motivated to become military officers: their academic studies are seen as a step on the way to that goal. In contrast, the majority of the older students are already officers: they are more academically motivated, as they aim to further their career by obtaining a degree.

Method

From 2011 to 2013 the statics course was taught by a single lecturer, and largely based on the textbook *Engineering Mechanics: Statics* by Hibbeler (2013). In 2014 the course was taught by two different lecturers (including one of the authors), but with the same curriculum and based on the same lecture notes and using many of the same assessment tasks. The assessment for the course consisted of ten assignments, two quizzes, a laboratory task and a final exam. The final exam was weighted 60% and was very similar in all years 2011 to 2014. The other components had large overlap. Marks and basic demographic data from 2011 to 2014 were used in this study.

In 2013 and 2014, the Statics students (79 students in 2013, 80 in 2014) were given three diagnostic tests during or prior to the first week of semester: a multiple choice basic mathematics test, given prior to the start of semester or in the first week; a written mathematics test in a statics lecture; and a multiple choice physics test in a physics lecture.

The basic mathematics diagnostic test consisted of 20 multiple choice questions addressing techniques typically encountered in year 10 and 11 at high school. This diagnostic was administered in the first week of semester in 2013 to all students taking mathematics as part of their degree program; and in the week prior to semester starting in 2014 to the new intake of officer cadets and midshipmen in all science and engineering degree programs (civilians, officers and repeating students were not available to take the mathematics test at this time).

The statics diagnostic test consisted of 10 questions in 2013, and 7 (of these 10) questions in 2014 (three questions were discarded to eliminate repetition. The test covered solving simultaneous equations, integrating simple polynomials and trigonometric functions, simple geometry including using the sine and cosine rules, and finding the components of a vector in Cartesian coordinates in 2 and 3 dimensions. All but one of the questions were taken directly from tutorial problems which the students met later in the course, with the engineering context removed.

The physics test was the Force Concept Inventory (FCI), a standard diagnostic test to measure students' understanding of Newtonian mechanics (Hestenes *et al* 1992; Halloun *et al* 1995). The FCI is a 30 question multiple choice test with questions requiring students to identify forces acting in various situations and apply Newton's laws in both equilibrium and non-equilibrium situations. Students were given 30 minutes to complete the test in their first physics lecture in 2013. Unfortunately in 2014, due to technical problems, students only had 20 minutes to complete the FCI and many were not able to complete the test in this time. At the beginning of second semester, in 2013 and 2014, the students were given the full FCI.

Demographic data was collected from students in the 2013 and 2014 statics cohorts including the year and state in which they had undertaken year 12 physics. Gender and age information was also collected from student records for the 2011 to 2014 statics cohorts. University entry/ranking scores were available for some students but, due to the Australia-wide recruitment, these scores are not easily comparable (different states use different scales). Given the high entry requirement and small spread of scores for the first-year engineering cohort noted earlier, these data were not used in this study.

Results

We have calculated the correlation coefficients between the cohort's final course grade in Statics, with the various diagnostic tests, and the final course grades in both Physics and Mathematics. The various correlations are summarised in Table 1.

Table 1: Correlation of statics final course grade with diagnostic test grades and mathematics and physics final course grades.

	Diagnostics			Finals		
Cohort	**Maths**	**Statics**	**FCI-pre**	**Physics**	**Maths**	**FCI-post**
2013	0.41	0.38	0.25	0.61	0.81	0.38
2014	0.46	0.42	0.42*	0.77	**	0.41

* In 2014 only the first 20 questions were used. **Mathematics marks are only available at the end of the academic year.

The correlation between the final statics mark and the statics diagnostic test mark are positive and significant at the $\alpha = 0.01$ level, however neither is a large correlation. If we simply divide the class into those that passed the diagnostic test (about 60% of the class) and those that did not (about 40%), the average final mark for those that passed was 73%, while the average for those that failed the diagnostic was 61%. Of the 64 students who failed the diagnostic in 2013 and 2014, 12 went on to fail the course, while 52 passed. An additional 4 students clearly passed the diagnostic but went on to fail the course. Of the 52 students who failed the diagnostic but passed the course, many attained good course marks, including distinctions and high distinctions. Hence while the diagnostic test gives some indication of how students are likely to perform in the course, it is a fairly blunt instrument.

Students who failed the diagnostic test were directed to attend remedial tutorials, and most did attend at least one remedial tutorial in 2013. However due to timetabling problems in 2014, many of the students did not attend. There is no clear relationship between attending the remedial tutorials and final course marks in either 2013 or 2014. The fail rate did not significantly decrease following the introduction of diagnostic tests and remedial tutorials, nor did those who attended the remedial tutorials do noticeably better than those who did not.

The correlation between FCI marks and final statics mark was 0.25 in 2013 which is significant at the $\alpha = 0.05$ but not the 0.01 level. This is a surprisingly small correlation, given that the conceptual basis of statics is Newtonian mechanics. The correlation with the second FCI mark, from start of second semester, was higher but still not large. This indicates that not only was the students' prior understanding of Newtonian mechanics a poor predictor of success in statics, but that their ability to succeed in statics correlated only weakly with their conceptual understanding of Newtonian mechanics. While correlation does not imply causality, it seems that, for the 2013 cohort, a good conceptual understanding of Newtonian mechanics was not a prerequisite for success in statics, nor did studying statics necessarily improve the cohort's conceptual understanding of Newtonian mechanics.

In 2014 the correlation with the FCI at the start of semester was substantially higher at 0.42. This number is not directly comparable to the 2013 result as it is calculated using only the first 20 (of 30) FCI questions. However we note that the change of lecturers in 2014 may have been significant here. In 2014, there was an increased emphasis on setting up problems in more detail and showing free body diagrams, particularly in the first half of the course. While the course content and most of the assessment tasks were unchanged, there may also have been a small difference in the way marks were assigned, with more marks for demonstrating conceptual understanding than would have been awarded in previous years.

The mathematics diagnostic test in both 2013 and 2014 had similar or slightly higher correlation than the statics diagnostic test with the final statics mark. This was an interesting result because the statics diagnostic test was particularly designed to test mathematics that would be required during the course. It was also a written test, so students needed to demonstrate that they could use the mathematical ideas. The mathematics test was a multiple choice test requiring more fundamental mathematical knowledge.

When students final marks in statics were compared with their marks in mathematics and physics for 2013, it was found that there was a strong correlation with physics final marks (0.61) and even stronger with mathematics final mark (0.81). Note that the mathematics mark is for the entire year, as all students do two semesters of mathematics and there is no separate first semester mark. For 2014 the correlation with physics marks was 0.77, which may again indicate that the skills and abilities needed for student success in statics in 2014 were more similar to those required for physics, than was perhaps the case in 2013.

Our data includes two significant minorities: females, and older students. In the four years for which data was available (2011 – 2014), there were a total of 50 female students out of a total 357 students in the four cohorts, and 58 students (mostly mature-age officers) who were more than one year from leaving high school, with an overlap of 2 students in both minority groups. To allow for comparison across the years, we calculated the average normalised difference *d* between the results for each minority and their year-cohort average:

$$d = \left\langle \frac{\langle \text{minority} \rangle - \langle \text{cohort} \rangle}{\langle \text{cohort} \rangle} \right\rangle \tag{1}$$

To gauge the significance of each difference, we calculated the Behrens statistic t^* (a modified t-statistic for the comparison of non-Normal, different-variance distributions):

$$t^* = \frac{\langle \text{sub-cohort} \rangle - \langle \text{others} \rangle}{\sqrt{\frac{\sigma^2_{\text{sub-cohort}}}{n_{\text{sub-cohort}}} + \frac{\sigma^2_{\text{others}}}{n_{\text{others}}}}} \quad (2)$$

Table 2 shows the values of these measures as well as number of students and mean marks over the four years.

Table 2: Cohort subset average performances, for 2011 to 2014.

	All students	Females	Older students
Number	357	50	58
% of cohort	100	14	16
average grade (%)	72.9	67.3	74.4
d	-	-0.07	+0.04
*t**	-	-3.0	+0.75

The large negative value of t^* strongly supports the hypothesis that the female average is lower than the male average. The older students have a higher average final score than the younger students, however this is not significant at the 0.05 level.

Differences by state in final statics grades were not statistically significant, which may be due to the small sample sizes once the cohort is split by states. Ongoing data collection and analysis may reveal significant differences, although the imminent implementation of the Australian curriculum at high school level may well have an impact over time.

In addition to looking at final marks, we also looked at how the cohort and the two minority groups improved over the semester, relative to their initial performance in the diagnostic tests. Table 3 shows the performance of the cohort and subgroups in each diagnostic test and the final average grades for 2013 and 2014.

Table 3: Cohort subset performance on diagnostics and final grades, for 2013 and 2014.

		Diagnostics (%)			Finals (%)	
Cohort	***n***	**Maths**	**Statics**	**FCI-pre***	**Statics**	**FCI-post**
All	159	65	53	57	67	71
Female	23	61	48	42	60	61
Older	26	**63	48	54	72	76

* for first 20 questions only in 2014. **2013 only

Looking at Table 3, we can see that female students and older students performed poorly compared to the overall cohort on all three diagnostic tests. However unlike the female students, the older students were (on average) able to achieve grades just as high – or even higher – than the younger students. This was particularly noticeable on the FCI. The FCI data for the entire physics cohort showed this same pattern, with older students having significantly higher normalised learning gains than their younger colleagues. This may be evidence that the older students, while starting with less content knowledge, are more efficient and/or effective learners than the younger students (Low and Wilson, 2014).

The relatively poor performance of females in the diagnostic tests may be due to poorer knowledge and skills in the relevant areas, but it may also be at least partly attributed to the design of the tests. It is well known that males outperform females on the FCI (for example, McCullough 2004) and other multiple choice tests (for example, Madsen *et al* 2013, Bolger and Kellaghan 1990), and this has also been observed on written tests (Wilson *et al* 2007). It may be that even if the females had the same level of knowledge and skills as the males in the cohort, they would still obtain lower grades because of the nature of the assessment.

Implications

The correlation of final statics grade with each of the three diagnostic tests is significant, but given that these tests are meant to identify students at risk, they are not as high as we might hope. This is particularly disappointing because the diagnostics require significant time to administer and to mark. As the mathematics multiple choice test is a better predictor of success in statics than the other diagnostics, time could be saved by just using the data gathered in mathematics.

In contrast to the diagnostic tests, the correlations with final marks in other subjects are much higher, indicating that students who do well in one subject are likely to do well in their other subjects. We have already noted that the older students perform as well as or better than the younger students, even though they enter with less prior knowledge. This may be due to better learning or study skills, or to higher levels of motivation.

Based on these results, we suggest that the willingness and ability to learn is at least as important as prior content knowledge. Students who enter with low levels of prior knowledge, but who are able to quickly learn and apply new skills and content, are likely to out-perform those with more prior knowledge but poorer learning skills.

Hence we suggest employing a targeted approach to diagnosis and remediation. Firstly, identify students who are at risk of struggling early due to a lack of prior knowledge, and offer them “content” support to bring them up to speed. Identifying these students is relatively simple, and can be done with a range of diagnostic tests, including multiple choice tests that are quick to mark and analyse. Secondly – in parallel – identify students who are at risk of not keeping up with new material due to a lack of learning skills, and provide them with ongoing support (such as study groups) to develop their ability in this area.

Offering support simply by the provision of additional content is likely to be counter-productive for students identified as needing assistance in learning. These students primarily need assistance in learning how to learn, beginning with techniques for recognising *how* they learn. Study skills sessions, which use relevant examples from their courses, and explicit teaching of problem solving techniques and ways of organising information such as mind-mapping, are also likely to be helpful.

Identifying students who suffer from a lack of learning skills is more problematic. The literature on testing learning ability is largely focussed on learning disabilities, and hence the tests are not appropriate in this context, given the rigorous testing the students have already undergone during their recruitment. The closest parallels we have found appear in the recruitment of staff to cutting-edge technology implementation and consultancy firms, who want to hire people who can solve “new” problems on a regular basis: they do not want to employ staff who are simply good at content-recall. The recruitment activities used by such firms to identify “good learners” can also be used to identify those who are likely to struggle with synthesising new ideas.

For example, students could be set a task that required them to explain a phenomenon which they could legitimately be expected to understand given their incoming knowledge, but that they would not have encountered previously. Then, after a suitable interval (perhaps a few days), students are asked (a) to explain the phenomenon that they were tasked to explore; but also (b) to explain a similar phenomenon that is based on the same fundamental principle. Students who are “good learners” will be able to do both. Students who have at least applied themselves to the task should be able to do (a) even if they can’t do (b).

Students who lack either the ability or motivation to learn are unlikely to be able to succeed at either task, and should be provided with assistance in “learning how to learn”.

The two groups of particular interest that we have identified are likely to be treated differently under this approach. The older students, who appear to be more likely to lack content than learning skills, would be given extra content-only tutorials, possibly online. These could easily be made available to all students. The younger students who are at risk of not succeeding, including the females, are more likely to be identified as needing help with both content and learning skills. Hence they would be offered a program of content and study skill support.

Whether it is desirable to split the females into separate groups for remedial programs is highly debatable, particularly in the UNSW Canberra / ADFA context where both genders train and work together as equals. However there is good evidence that girls in the physical sciences benefit from being in classes where there is a “critical mass” of females (Feteris 2005). It is relatively simple to ensure that laboratory and tutorial groups include more than this minimum amount (with the trade-off being that some groups may be all-male).

Given that the relatively poor performance of females may also be due to the way in which tests (both diagnostic and summative) are written and administered, it is important to consider whether the range of assessment tasks and modes of assessment could be expanded, enabling female students to better demonstrate their knowledge and skills. The assessment in Statics at UNSW Canberra is largely exam based, with 80% of the final mark coming from exams. A broader range of assessment tasks, such as laboratory logbooks, oral presentations and open assignments might allow female students to better demonstrate their knowledge and skills, and thus perform better, without compromising the academic standards of the course. In addition, the lower correlation of Statics assessment marks with both the FCI and physics final marks, compared with mathematics, may indicate that the current assessment is too strongly biased towards testing procedural knowledge at the expense of conceptual understanding.

Conclusions

Simple diagnostic tests, such as basic mathematics multiple choice tests, may be just as effective at identifying students with inadequate preparation to succeed in a traditional statics course as more complex and time consuming tests. However, diagnostic tests are unlikely to identify students at risk of failure because of poor learning skills. A different type of test which identifies learning ability is needed in parallel with traditional diagnostic tests.

Once “at risk” students have been identified by the appropriate tests, the appropriate support can be offered. This may be content focussed, learning skill focussed or both. Such an approach will benefit all “at risk” students. It will allow resources to be better targeted to meet students’ needs. It may also reduce the burden on academics to provide remedial content based tutorials for students who do not need them, such as older students with better developed learning skills. Attention should be paid to the classroom environment and assessment methods used for female students.

References

Bolger, N., and Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*,*27*(2), 165-174.

de Koning, B. B., Loyens, S. M., Rikers, R. M., Smeets, G., and van der Molen, H. T. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual differences*, *22*(3), 313-323.

Feteris, S. (2005). Seating in laboratory classes-achieving critical mass. In *Proceedings of the 16th National Australian Institute of Physics Congress [http://aipcongress2005.anu.edu.au/pdf/AIPC_Handbook_V2.pdf]*.

French, B. F., Immekus, J. C., and Oakes, W. C. (2005). An examination of indicators of engineering students' success and persistence. *Journal of Engineering Education*, *94*(4), 419-425.

Gerdes, H., and Mallinckrodt, B. (1994). Emotional, social, and academic adjustment of college students: A longitudinal study of retention. *Journal of Counseling & Development*, *72*(3), 281-288.
Goldfinch, T. L., Carew, A. L. and McCarthy, T. J. (2008). Improving Learning in Engineering Mechanics: The Significance of Understanding. In L. Mann, A. Thompson P. Howard (Eds.), *Proceedings of the 19th Annual Conference of the Australasian Association for Engineering Education AAEE2008* (pp. 1-6).
Hacker, C., Seagar, A., and Cassidy, P. (2013) Can 'at risk students' be identified based on demographics? In *Proceedings of the 24th Annual Conference of the Australasian Association for Engineering Education AAEE2013*.
Halloun, I., Hake, R., Mosca, E., and Hestenes, D. (1995). Force Concept Inventory (revised 1995). Retrieved from http://modeling.asu.edu/R&E/Research.html (password protected).
Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30, 141158. doi: 10.1119/1.2343497
Hibbeler, R. C. (2013). Engineering Mechanics: Statics and Dynamics 13/e. Pearson Education, Singapore
James, R., Krause, K. L., and Jennings, C. (2010). *The first year experience in Australian universities*. Canberra: DEEWR.
Litzler, E. and Young, J. (2012), Understanding the Risk of Attrition in Undergraduate Engineering: Results from the Project to Assess Climate in Engineering. *Journal of Engineering Education*, 101: 319–345. doi: 10.1002/j.2168-9830.2012.tb00052.x
Low, D. and Wilson, K.F. (2014), Teachability: it's harder to teach a young dog new tricks. *Proceedings of the 2014 Australian Conference on Science and Mathematics Education*, Sydney, Australia (in press).
Madsen, A., McKagan, S. B., and Sayre, E. C. (2013). Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? *Physical Review Special Topics-Physics Education Research*, *9*(2), 020121.
McCullough, L (2004). Gender, Context, and Physics Assessment. *Journal of International Women's Studies*, 5(4), 20-30.
McKenzie, K and Schweitzer, R (2001) Who Succeeds at University? Factors predicting academic performance in first year Australian university students, *Higher Education Research & Development*, 20:1, 21-33, doi: 10.1080/07924360120043621
Power, C., Robertson, F., and Baker, M. (1987). *Success in higher education.* Canberra: Australian Government Publishing Service.
Sheard, M. (2009), Hardiness commitment, gender, and age differentiate university academic performance. *British Journal of Educational Psychology*, 79: 189–204. doi: 10.1348/000709908X304406
Thomas, G., Henderson, A., and Goldfinch, T. (2009). The influence of University entry scores on performance in engineering mechanics. In *20th Annual Conference for the Australasian Association for Engineering Education, 6-9 December 2009* (p. 980). Engineers Australia.
Trigwell, K., Ashwin, P. and Millan, E. S. (2013), Evoked prior learning experience and approach to learning as predictors of academic achievement. British Journal of Educational Psychology, 83: 363–378. doi: 10.1111/j.2044-8279.2012.02066.x
Wilson, K., Kueter, N., Dennis, G., Nulsen, A., and Verdon, M. (2007). Addressing gender disparity in the Physics National Qualifying Exam for the Australian Science Olympiads. *Teaching Science: The Journal of the Australian Science Teachers Association*, *53*(1).
Zhang, G., Anderson, T. J., Ohland, M. W. and Thorndyke, B. R. (2004), Identifying Factors Influencing Engineering Student Graduation: A Longitudinal and Cross-Institutional Study. *Journal of Engineering Education*, 93: 313–320. doi: 10.1002/j.2168-9830.2004.tb00820.x

Acknowledgements

The authors would like to acknowledge their colleagues in mathematics and the Teaching and Learning Group at UNSW Canberra who kindly shared their data.

Copyright statement