2006

Identifying at-risk students: is it possible in a tertiary preparation course for adults?

Colin S. Carmichael
University of Southern Queensland

Peter K. Dunn
University of Southern Queensland

Janet A. Taylor
University of Southern Queensland

Publication details

Carmichael, CS, Dunn, PK & Taylor, JA 2006, 'Identifying at-risk students: is it possible in a tertiary preparation course for adults?', in P Grootenboer, R Zevenbergen, & M Chinnappan (eds), *Proceedings of Identities, cultures and learning spaces: 29th Annual Conference of the Mathematics Education Research Group of Australasia (MERGA29)*, MERGA, Canberra, ACT, 1-5 July.

Identifying at-risk students: Is it possible in a tertiary preparation course for adults?

Colin Carmichael
University of Southern Queensland
<carmich@usq.edu.au>

Peter Dunn
University of Southern Queensland
<dunn@usq.edu.au>

Janet Taylor
University of Southern Queensland
<taylorja@usq.edu.au>

In the current educational climate in Australia, there is an imperative on university administrators to maintain student enrolments since funding is explicitly linked to student numbers. The most effective way of achieving this is to retain those students who have already enrolled. Consequently many universities currently seek to identify those students at risk of either failing a course or withdrawing from the university. In this paper we report on an initial study into the use of pre-entry measures to identify at-risk students in the context of a tertiary preparatory course offered entirely in the distance mode. We conclude that such measures are at best rough indicators of at-risk students and that results from such measures should be used in a non-directive manner only.

Effective teachers in the school classroom context continually monitor their students using informal tests, discussions and observations to detect students at risk of failing their subject. They are then in a position to assist the student, perhaps directing them to additional resources, encouraging them or even seeking assistance from learning support specialists. Teachers in university settings, where classes may be in the order of several hundred students are often unable to establish the type of student/teacher relationship that exists in a school setting. Identifying those students in the class who lack confidence, motivation or knowledge becomes difficult and is exacerbated when students study in a distance mode. Rather than use observation of students, teachers in this context need to rely almost entirely on student responses to pre-course measures and perhaps early assessment items. In many courses, students who perform poorly in pre-course measures (for example pre-tests) are targeted for intervention programs. For example, in a 2001 study of mathematics, engineering and physical sciences departments in the United Kingdom (Learning & Teaching Support Network, n.d.), 68% of respondents indicated that their department used a diagnostic mathematics test in the first few weeks of semester. In the majority of instances (70%) these tests were used to determine the level of mathematics help required by students. If such intervention relies on diagnostic tests and other pre-course measures, then it is important that such tests are able to identify at-risk students. In this paper we investigate, using statistical modelling, the ability of a pre-entry mathematics test and other measures to identify at-risk students within the context of a Tertiary Preparatory Program (TPP) mathematics course.

Background

Experienced teachers claim to ‘know’ from the results of pre-entry test whether a student is at risk of failing. This knowledge, however, is rarely tested. A pre-course measure that purports to identify at-risk students should be based upon a model that seeks to explain the variation in final student educational performance, with such performance including measures of student progression through the course and their final achievement in the course. Variation in student performance can be explained by factors that relate to the individual, their environment and the interaction between the two.

Traditionally models of student performance have concentrated on cognitive explanatory factors such as prior achievement. In the tertiary environment, Power, Robertson and Baker (as cited in Zeegers, 2004) found in a cross-institutional study of 5000 students that tertiary entrance score was the best predictor of academic performance. Similarly, in a meta-analysis of 109 studies; Robbins et al. (2004) found that a combination of high school grades and standardised test scores accounted for the majority of explained variance in statistical models of performance.

However, other factors from the affective and conative domains have also been shown to be important in modelling performance. Pintrich and De Groot, (1990) indicate that learners who can self-regulate (have strong conative and cognitive skills) will out-perform other learners, while Bower, (1981) has shown that a student's mood during learning will influence the quality of their learning and their subsequent recall. In the conative domain, "…contemporary motivational theories are emerging as strong models of academic achievement" (Robbins et al., 2004, p.263) with mathematical self-efficacy becoming an accepted predictor of achievement in mathematics (Pajeres & Graham, 1999; Pajeres & Miller, 1994; Stevens, Olivarez, Lan, & Tallent-Runnels, 2004).

Environmental factors that are known to influence student performance are many and varied and can include: an ability to access financial resources (Considine & Zappala, 2001); social support (DeBerard, Spielmans, & Julka, 2004), and the mode and quality of teaching.

While pre-course measures that seek to identify at-risk students ideally should reflect models of performance that encompass environmental influences and factors that collectively span all three domains, in practice this rarely occurs. Not only would such pre-course instruments be so large and create respondent burden, they may not add to the predictive strength of the measures. In a meta-analysis reviewing influences of psycho-social variables on student performance Robbins et al. (2004) found that overall prior knowledge outweighed other factors in their predictive strength.

The current study focuses on the performance of students within a Tertiary Preparation Program (TPP). This program has been designed for students returning to study after an absence and is often composed of students who are mature aged (median 30 years) and who have not had success in studying mathematics. In this study we employ an existing mathematics pre-test (called the M-test) to assess student's prior knowledge in mathematics along with measures of student's self-efficacy and some demographic variables. The latter measure is included as it is known to predict performance in mathematics for older and lower achieving students (Multon, Brown, & Lent, 1991).

Methodology

In the first semester of 2005, all students enrolled in a TPP Mathematics Course (TPP7181) were required to complete the M-test. Part A of this test contained 16 items that ranged from basic calculations of the type '$102-36$' to the interpretation of trend graphs and was scored out of a total of 16. Part B of the test contained 21 items that ranged from concepts of ratio such as $\frac{3}{4}=\frac{15}{?}$ to the drawing of a trend graph, and was scored out of a total of 22.

All enrolled students were also invited to participate in a survey designed to assess their self-efficacy (reported in Carmichael & Taylor, 2005). This survey utilised existing scales that were able to provide an assessment of student self-efficacy at three levels:

- at the course level (confidence in succeeding in the course);
- at the topic level (confidence in succeeding in, for example, algebra); and,
- at the task level (confidence in succeeding in a given and specific task).

Demographic information collected included; the student's age, gender, the number of years since they last studied formal mathematics and their highest year level reached at school.

Performance data for the students included whether they remained for the duration of the course (and completed the final examination) and their total course score. The latter is an aggregate of scores obtained from five assignments and one examination.

As the purpose of this study was to establish the predictive validity of the pre-course measure, regression-type models were employed. In particular a standard linear regression model was applied to the total course score and a logistic regression model was applied to completion data. The models were then externally validated against the data obtained from 202 students enrolled in the course during the second semester of 2005.

Results

One hundred and twenty five students participated in the study out of a total of 300 enrolled students. Of these respondents, 47 were male and 78 were female, with ages ranging from 18 to 54 years (median: 30 years). The number of years since the student last studied mathematics formally ranged from 0 to 40 (median: 12). Forty six percent of respondents had completed formal study to a year 12 (or equivalent) level while eighteen percent had not even completed formal study to a year 10 level.

Of the 125 respondents, 71 completed the course, 32 completed some components of the course and 22 withdrew without submitting any assessment tasks. The latter group were awarded a total course score of zero. The distribution of total course score for all students is shown in Figure 1. This distribution is obviously not Gaussian and therefore it was inappropriate to apply standard linear regression models to the entire total course score distribution. The distribution of scores for students who completed all assessment items in the course, however, was close to symmetric, and we were able to apply a linear regression model to this subset of the data.

The M-test results for all respondents ranged from 15 to 38 (out of a total score of 38) with a mean of 30. Only three students scored less than 19 in the M-test and all three subsequently withdrew from the course. Nineteen students scored less than 28 in the M-test (approximately 75% of the total score) and of these, 14 subsequently withdrew from the course. These results indicate that students who achieved a poor score in the M-test were more likely to be at risk of withdrawing from the course; however many students who performed well in this test also withdrew or failed the course.

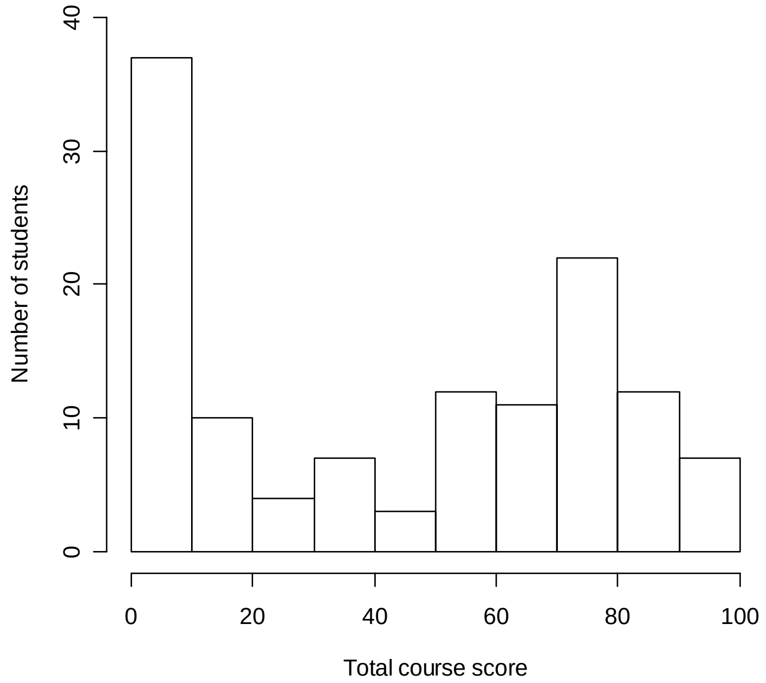

Figure 1: Distribution of total course score (n = 125).

The application of a linear model to the total course score of the 71 students who completed the course resulted in two significant explanatory factors, their score in the M-test ($p = 0.001$) and their gender ($p = 0.003$). The model (see Equation 1.1) was able to explain 20.6% of the variation in total course scores.

$$\text{Score} = 17.5 + 1.5\text{Mtest} + 11.5\text{Gender} \tag{1.1}$$

(where Gender is an indicator variable equal to 1 when students are female and 0 otherwise).

From this model, we see that with all other factors constant, an increase in the M-test result of 1 mark should lead to an increase in total score of 1.5%. Similarly, with other factors constant, a female student should score 11.5% more marks than her male counterpart. In other words, there was some evidence to suggest that poor performance in the M-test may identify at-risk students and that males were more at risk than females. However, these conclusions were restricted to those students who completed the course.

A dichotomous variable was created to model whether the students completed all assessment tasks or did not. A generalised linear model with a logit link was applied to these data resulting in two significant explanatory variables, the student's score in the M-test ($p = 0.001$) and their age ($p = 0.01$). As this particular model was non-linear, a pseudo R^2 value based on McFadden's likelihood ratio statistic was used (Hardin & Hilbe, 2001). The model (see Equation 1.2) had a pseudo R^2 value of 12.2%.

$$\text{logit(complete)} = -2.63 + 0.15\text{Mtest} - 0.05\text{Age} \tag{1.2}$$

(where complete is equal to 1 if students complete the course, and 0 otherwise).

We see that with all other factors constant, an additional one mark in the M-Test will increase the natural log odds (logit) of a person completing the course by 0.15. Similarly, an increase in age of one year will decrease the natural log odds of the person completing the course by 0.05. This model provides some evidence to suggest that poor performance in the M-test may indicate students are at risk of dropping the course and older students are more at risk than younger students.

Attenuated measures of self-efficacy were not predictive of total course score or course completion, although there was a weak significant correlation between question efficacy and the first assignment result ($r = 0.2,\ p = 0.04$) and a moderate correlation between the first assignment result and the final examination ($r = 0.4,\ p = 0.01$).

In order to externally validate the above models we used pre-course measures obtained in the second semester of 2005. Of the 300 students initially enrolled in the course for semester 2, M-test results were available for 202 students. In this subgroup, 87 students completed the course (that is sat for the examination) and 115 did not. As the first regression model obtained above was based only on data for students who completed the course, it was felt that only the second model could adequately identify at-risk students. Accordingly this model was applied to the ages and M-test results for the 202 students and the probability of the student completing the course (*C*) was then estimated. This probability was compared with benchmarks of 0.1 to 0.5 (to arbitrarily gauge when a student was at risk of not completing) and then at-risk students were identified. Numbers of students correctly identified as being at risk of not completing the course and numbers of students incorrectly identified are shown in Table 1. If we assume that those students whose probability of completion is less than 0.1 are at risk of not-completing, then the model is able to correctly identify 4 of the 115 students who did not complete, further it incorrectly identifies 4 students (of the 87 who did complete). In other words the model is almost useless, although for larger benchmarks it seems to be able to correctly identify a larger proportion of students.

Table 1
Number of students correctly and incorrectly identified as being at-risk

	$C = 0.1$*	$C = 0.2$	$C = 0.3$	$C = 0.4$	$C = 0.5$
Number of students correctly identified as being at-risk	4	7	17	36	60
Number of students incorrectly identified as being at-risk	4	8	13	21	37

*where *C* is the probability of a student completing the course

Discussion

There are several issues that arise from these almost inconclusive results. These relate to the difficulties in obtaining suitable predictors of performance and the statistical difficulties associated with modelling the data in this particular context.

Although both models reported above produced significant predictor variables, the percentage of explained variance was low, although within the range reported in similar studies (see Robbins et al., 2004). So despite being able to explain some of the variation in final performance we did not have sufficient measures to account for the vast majority of this variation. Arguably many of the predictors in such a model of performance will relate to measures associated with the mind. Such measures are difficult to measure and are often unstable over time. For example a strong correlation exists between self-efficacy and performance when both measures are proximal (Pajeres & Graham, 1999), however a much weaker correlation exists when the two measures are not proximal (Lane & Lane, 2001). If we intend to identify students at risk of failing a course through the use of such pre-course measures, then it is doubtful that we will ever explain a large proportion of the variation in final course performance. This is because measures of those attributes that explain performance will invariably change through the duration of a semester course. In fact many preparatory courses are designed to do just that (Taylor & Mohr, 2001).

There were statistical problems in this particular context caused by the large proportion of students who withdrew from the course. In order to use regression techniques to validate our pre-course measures, a large amount of data was either discarded or not fully utilised. In the linear model of student achievement, results for the 54 students who made partial or no progress were ignored. Similarly in the model of student completion, the achievement results of the 32 students who made partial progress were not fully utilised. This treatment of the data reduced the power of the model and possibly produced biased model estimates. Two possible methods for addressing this problem are being investigated in a related study. One makes use of a generalised linear model that is based upon a class of power variance dispersion models that include exact zeros (Jorgensen, 1997). Unlike standard linear regression that assumes a constant variance, these models assume that the variance will vary according to some power of the mean, that is $\mathrm{var}(Y) = \mu^p$. Methods have been developed for fitting generalised linear models that are based on these dispersion models (Dunn, 2001) and initial results suggest that they can be applied to educational performance data.

The second possible method for dealing with these data is to apply a Tobit model (Long, 1997). Such a model assumes the existence of an underlying latent variable that is only observable when it reaches and exceeds some threshold. So we could hypothesise the existence of a latent variable called, for example, 'propensity to learn mathematics'. This variable is then only observable in students when it exceeds, for example, the value of zero. Then it can be measured using their total course score. Again, initial results suggest that a Tobit model can be applied to this situation.

It may be possible, that despite the problems associated with actually locating suitable predictors of at-risk students, we may be able to at least use all of the data available to us and do so in a more effective way.

Conclusion

In this paper we have reported on an initial study that sought to assess the predictive validity of pre-course measures aimed to identify at-risk students in a distance education course. We used regression-type models for this purpose and were able to report on models of both student achievement and completion. These models, however, were of limited use, as they both failed to explain very much of the variation in the final performance of students in our sample. It is possible that with the inclusion of more measures aimed to assess other areas in the affective and conative domains, we may be able to develop a model that can explain more of this variation. For example, Petry & Craft (1976) were able to explain 55% of the variation in performance for high-risk college students using several pre-course measures; however their questionnaire included hundreds of individual items.

Such a large number of items is counterproductive and in any case, with such a model we are only just over 50% certain that we can identify at-risk students.

The use of pre-course measures to identify at-risk students is akin to measuring height with a yardstick; they are of limited use. In this study we found some evidence to suggest that older students or those who perform poorly in a pre-course mathematics test are at risk of withdrawing from the course. Certainly many of the students who performed poorly in the M-test did not succeed in the course, however many students who did perform satisfactorily in the M-test also did not succeed. Despite the limitations of such tests, academics continue to use them, and direct students who perform poorly in these tests towards learning support programs. Lane, Hall, & Lane, (2004), for example, argue that students scoring less than 1 standard deviation below the mean result in a self-efficacy scale could be at risk of failing. This is despite the measure not correlating with final performance when it is implemented early in the course and then moderately correlating ($r = 0.41$) with final performance when it is implemented mid-way through the course.

We have discussed at length the poor predictive validity of pre-course measures in general. Another issue that needs to be addressed is the reliability of such measures. In this study as presumably in many other instances, the pre-course test was implemented in the first few weeks of semester and it did not contribute towards students' final assessment. It is conceivable that if the same test were a 'high stakes' assessment item the same students may achieve markedly different results. To overcome this, some institutions allow students three attempts to complete 'parallel' versions of the test (Golden, n.d.) while at the same time providing revision material. Others have developed the test as a self-assessment instrument linked with developmental resources (Taylor, 1998). In this latter case it was noted that the test itself acted as a tutoring tool. In the current study the M-test required only short answer responses, so there was no way to identify students who had partial knowledge of some questions. This problem is exacerbated if the test is entirely multiple-choice.

We argue that there is a place for pre-course measures and the use of associated statistical models designed to explain student performance. Through their use we gain knowledge of the factors that influence student performance and also a better understanding of the profile of our enrolled students. We advocate that pre-course tests should not be used in a prescriptive manner. Instead these tests would be more beneficial if used as a self-diagnostic for students. In other words students need to be encouraged to use their performance in a pre-test, to reflect on their readiness for the course ahead. Students who feel under prepared (and this is probably as good a predictor of final performance than any we can measure) should then be guided to learning support resources. The results of this study do not support the practice of excluding or targeting individual students who do not perform well in a one-off pre-course measure of readiness.

References

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*(2), 129-148.

Carmichael, C. S., & Taylor, J. A. (2005). The analysis of student beliefs in a tertiary preparatory mathematics course. *International Journal of Mathematical Education in Science and Technology, 36*(7), 713-179.

Considine, G., & Zappala, G. (2001). *Factors influencing the educational performance of students from disadvantaged backgrounds.* Paper presented at the National Social Policy Conference, Sydney, N.S.W.

DeBerard, M. S., Spielmans, G. I., & Julka, D. C. (2004). Predictors of academic achievement and retention among College Freshmen: A longitudinal study. *College Student Journal, 38*(1), 66-81.

Dunn, P. K. (2001). *Likelihood-based inference for Tweedie exponential dispersion models.* Unpublished Ph.D., University of Queensland, Brisbane.

Golden, K. (n.d.) Diagnostic testing and student support. In *Diagnostic Testing for Mathematics*. Retrieved March 27, 2006, from: www.engsc.uc.uk/downloads/mathsteam/diagnostic_test.pdf

Hardin, J., & Hilbe, J. (2001). *Generalised Linear Models and Extensions*. Texas: Stata Press.

Jorgensen, B. (1997). *Theory of Dispersion Models*. London: Chapman and Hall.

Lane, A., Hall, R., & Lane, J. (2004). Self-efficacy and statistics performance among sport studies students. *Teaching in Higher Education, 9*(4), 435-448.

Lane, J., & Lane, A. (2001). Self-efficacy and academic performance. *Social Behaviour and Personality, 29*(7), 687-694.

Learning and Teaching Support Network. (n.d). *Diagnostic Testing for Mathematics*. Retrieved March 27, 2006, from: www.engsc.uc.uk/downloads/mathsteam/diagnostic_test.pdf

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks: SAGE Publications.

Multon, J., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 31-38.

Pajeres, F., & Graham, L. (1999). Self-Efficacy, Motivation Constructs, and Mathematics Performance of Entering Middle School Students. *Contemporary Educational Psychology, 24*, 124-139.

Pajeres, F., & Miller, M. D. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

Petry, J.R. & Craft, P.A. (1976). Investigation of instruments to predict academic performance of high-risk college students. *The Journal of Educational Research, 70*(1), 21-25.

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

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288.

Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *Journal of Educational Research, 97*(4), 208-221.

Taylor, J. A. (1998). Self Test: a flexible self assessment package for distance and other learners. *Computers and Education, 31*, 319-328.

Taylor, J. A., & Mohr, J. (2001). Mathematics for math anxious students studying at a distance. *Journal of Developmental Education, 25*(1).

Zeegers, P. (2004). Student learning in higher education: a path analysis of academic achievement. *Higher Education Research and Development, 23*(1), 35-56.