# CORRELATION BETWEEN STUDENT PERFORMANCE IN LINEAR ALGEBRA AND CATEGORIES OF A TAXONOMY


**Leigh N. WOOD**
University of Technology, Sydney
PO Box 123, Broadway, Australia, 2007
leigh.wood@uts.edu.au

**Geoffrey H. SMITH**
University of Technology, Sydney
PO Box 123, Broadway, Australia, 2007
gsmith@maths.uts.edu.au

**Peter PETOCZ**
University of Technology, Sydney
PO Box 123, Broadway, Australia, 2007
Peter.petocz@uts.edu.au

**Anna REID**
Macquarie University
North Ryde, Australia, 2114
anna.reid@mq.edu.au



## ABSTRACT

This paper concerns a study of the performance of students in a recent linear algebra examination. We investigated differences in performance in tasks requiring understanding of the concepts with those that required only the use of routine procedures and factual recall. Central to this study was the use of a taxonomy, based on Bloom's Taxonomy, for characterising assessment tasks, which we have described in previous publications. The full taxonomy has 8 categories, which fall into 3 broad groups. The first group (A) encompasses tasks which could be successfully done using a surface learning approach, while the other two (B and C) require a deeper learning approach for their successful completion. Tasks on the examination paper were put into one of the three groups and comparisons were made concerning the performance of individual students in each of these areas.

There are several interesting areas to investigate. The first is to identify those students whose performance in group A was markedly different to their performance on groups B and C. There is considerable disquiet amongst mathematics lecturers at tertiary level as to the routine algebraic skills of incoming students and of students studying mathematics at university (see for example the *ICMI Study into the Teaching and Learning of Mathematics at University Level*, 2001). There is a conjecture that students who have poor technical skills are not able to succeed in university mathematics. The contrapositive conjecture that good technical skills (such as algebraic dexterity) are necessary for success in university mathematics is often taken for granted. The taxonomy allows us to test this hypothesis as we can compare performance in group A tasks (routine) with performance in higher level B and C tasks.

We have also investigated whether or not the data supports any systematic effect of differences in sex or language background in the performance on the three groups.

The sample contained a large cohort of students with who had a home language other than English. We tested the hypothesis that such students would have difficulty with the conceptual aspects of the course, since these normally require greater language facility. This proved not to be the case.

# 1. Introduction

This paper investigates students' performance on an examination—and by extension their learning in the subject—from the point of view of a taxonomy of mathematical tasks. It examines various hypotheses about factors that may affect the nature and success of students learning.

Assessment is a central feature of teaching in formal institutions and can take a multitude of forms, fulfilling many functions, both intended and unintended. Ideally assessment should be linked closely with student learning. We look at a taxonomy for learning in mathematics (Smith *et al* 1996) that is related to that of Bloom (1956). It transforms the notion that learning is related to what we as educators do to students, to how students understand a specific learning domain, how they perceive their learning situation and how they respond to this perception within exam conditions.

We will particularly look at examinations because we believe that a major component of the final grade will continue to be contributed by examination of individual students. As Krantz (1999:57) says '*The principle device for determining grades is the examination*'. There are many reasons for this. Firstly, it is a practical, cost-effective way to assess large numbers of students. Secondly, examinations are seen by many as objective with no favouritism and providing equity, as all students are treated under the same conditions. Thirdly, examinations provide quality assurance and accountability, especially for administrators. Fourthly, examinations have a long historical precedent in mathematics and in educational areas where certification is involved. All of these reasons for maintaining exams focus on their format and administration.

Whether we focus on examinations or other forms of assessment, we can use a range of techniques to assess the nature and extent of student learning. Our decisions about just which forms of assessment we choose are likely to be affected by the particular learning context and by the type of learning outcome we wish to achieve. Essentially, good assessment processes:

- **Encourage meaningful learning** when tasks encourage understanding, integration and application.
- **Are valid** when tasks and criteria are clearly related to the learning objectives and when marks or grades genuinely reflect students' levels of achievement.
- **Are reliable** when markers have a shared understanding of what the criteria are and what they mean.
- **Are fair** if students know when and how they are going to be assessed, what is important and what standards are expected.
- **Are equitable** when they ensure that students are assessed on their learning in relation to the objectives.
- **Inform teachers about their students' learning** (see Brown *et al*, 1997, Brockbank & McGill, 1998 or Biggs, 1999 for greater discussion on the relations between assessment and learning)**.**

With regard to the importance of assessment, Ramsden (1992) says that '*From our students' point of view, assessment always defines the actual curriculum. In the last analysis, that is where the curriculum resides for them, not in the lists of topics or objectives. Assessment sends messages about the standard and amount of work required, and what aspects of the syllabus are most important. Too much assessed work leads to superficial approaches; clear indications of priorities in what has to be learned, and why it has to be learned, provide fertile ground for deep approaches'* (p187).

It follows that students will look carefully at the range of assessment tasks—including examinations—that are involved in any course of study. In mathematics courses, students usually have access to previous examination papers and these very papers give a clear indication of the nature and extent of their course, and the sorts of things that they need to concentrate on in order to achieve high marks or grades in their courses.

# 2. Examinations

The nature of examinations themselves will change in both content and format. Online delivery with individualised questions will supplement paper-based and oral examinations providing a range of flexibility. The content will change with access to technology, which makes many routine skills less important. Employers are looking for cognitive and communication skills in graduates and this will be reflected in the questions asked in examinations. Students will be able to use a variety of tools in the examination, open-book and/or computer if appropriate. These changes take place in the context of changing classroom environments where higher order conceptions of learning are encouraged through the use of supporting student focused activities (Reid & Petocz, 2001). Assessment is a tool that can be used by students to develop the depth of their understanding of a topic, and also to demonstrate this depth to their teachers. Examinations have the same potential but often send a contrary message. This contrary message is generated by the weighting given to certain questions and thus to the relative importance given to them by students. Hence academics setting examinations need to consider the examination as part of the students' overall learning experience and accordingly need to focus the exam on issues and contexts that encourage a continuation of higher order conceptual thinking. It is important to remember that one quality of higher order conceptions of learning is that they are inclusive and integrated. This means that by encouraging higher order conceptions through class activities and assessments, we are also encouraging the use of routine activities within that context. Crawford *et al* (1994) show this clearly in their categories that describe student learning of mathematics. In their work on innovative examination questions, Smith *et al* (1996) and Ball *et al* (1998) show how the nature of the examination questions directs students toward demonstrating either their understanding of ideas or simply their ability to perform routine functions.

Our categories of mathematics learning, developed from Blooms' taxonomy, provide a schema through which we can evaluate the nature of examination questions in mathematics to ensure that there is a mix of questions that will enable students to show the quality of their learning at several levels.

# 3. Use of a taxonomy

We have been using a taxonomy (Table 1) to ensure that examinations contain a mix of questions to test skills and concepts. The taxonomy was developed due to our desire to encourage a deep approach to learning. Previous studies have shown that many students arrive at university with a surface approach to learning mathematics (Crawford *et al*, 1996) and that this affects their results at university. There are many ways to encourage a shift to deep learning, including assessment, learning experiences, teaching methods, and attitudinal changes. The taxonomy addresses the issue of assessment. It can be applied to all assessment tasks but in this paper it is specifically applied to examinations. The taxonomy has eight categories, falling into three main groups (Smith *et al*, 1996). Group A consists of tasks which students will have been given in lectures or will have practised extensively in tutorials. In group B tasks, students are required to apply their learning to new situations, or to present information in a new or different way. Group C encompasses the skills of justification, interpretation and evaluation.

| Group A | Group B | Group C |
|---|---|---|
| Factual knowledge | Information transfer | Justifying and interpreting |
| Comprehension | Applications in new situations | Implication, conjectures and comparisons |
| Routine use of procedures | | Evaluation |

**Table 1**. MATH Taxonomy (after Bloom). Smith *et al*, 1996

In a previous study (Smith & Wood, 1998), when we looked at the contribution of group A to the total mark gained by the student, we found a significant difference between the performance of males and females. The contribution of group A to the total mark was greater for females, even though there was no significant difference between males and females on the total score. This finding was also investigated with the present data.

The categories of the taxonomy are context specific—proving a theorem when the proof has been emphasized in class is a group A task, while proving the same theorem *ab initio* is a group C task. The taxonomy encourages us to think more about our first attempts at constructing exercises. Whether we act consciously on this influence or simply make changes instinctively, it provides a useful check on whether we have "tested" all the skills, knowledge and abilities that we wish our students to demonstrate.

# 4. Construction of the examination

We have taken a typical examination of the subject Linear Algebra. This subject was neither taught nor assessed by any of the authors of this paper. The examination was a formal 3-hour university examination in June 2001 with students being able to use scientific calculators and no other aids. Eighty-five students completed the paper and we have data on their marks in all subsections. We also have data on their sex, language background and the number of years in Australia.

The examination consisted of 88 marks of group A tasks, 15 group B and 27 marks in group C, for a total of 130 marks. It is obvious from the weighting of the group A tasks that the lecturer considered that routine tasks were the most important aspects of the subject, or perhaps was setting the exam in a "traditional" way, without using a broad range of question types.

- An example of a group A task (routine procedure) on the paper is
  - (i) Find the eigenvalues and corresponding eigenvectors of the matrix
$$A = \begin{pmatrix} 1 & 5 \\ 5 & 1 \end{pmatrix}.$$
  - (ii) Hence or otherwise, find the diagonal matrix $D$ and an invertible matrix $P$ such that $A = PDP^{-1}$
  - (iii) Calculate the spectral decomposition of $A$
  - (iv) Use the spectral decomposition to calculate the inverse of $A$

  In this task, the main requirement was for the student to reproduce work done in class.
- An example of a group B task on the paper is

Explain how the LU decomposition of a matrix $A$ is used to solve the system of linear equations $A\,\mathbf{x} = B$.

In this task, the student is required to transform their knowledge of a routine skill to the meta-knowledge of explaining the skill.

- An example of a group C task (justifying) on the paper is

  Let $T = \{v_1, v_2, \ldots, v_r\}$ be a linearly independent set of vectors and let $A$ be an $n \times n$ matrix. Show that the set $T = \{Av_1, Av_2, \ldots, Av_r\}$ is also linearly independent.

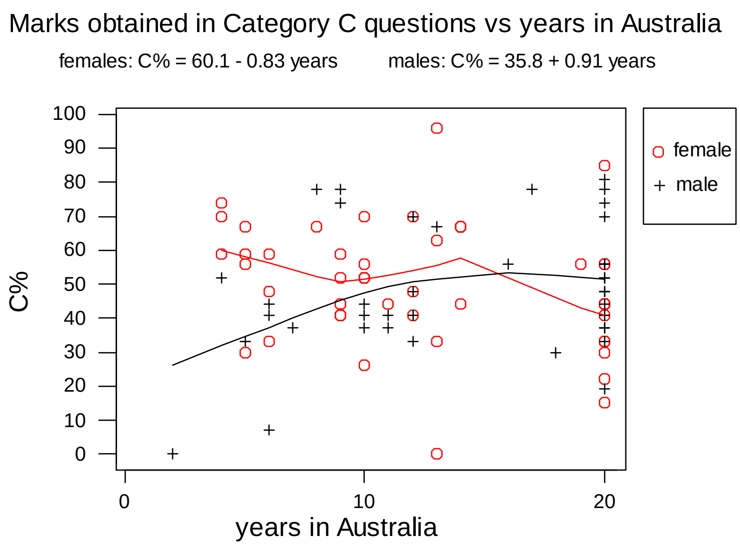
The examination was long, so none of the students completed the whole paper. So although students could have answered all sections, the length of the paper meant that in fact they could choose which sections to attempt. The majority of students started from the beginning and did not make full attempts at the later questions. This did not influence their results on the A, B and C tasks because they were distributed throughout the paper. It did influence the average mark for the examination.

## 5. Results

The correlations between the scores on group A, B and C tasks were significant and high (the correlations were 0.83, 0.67, 0.65) indicating that all components were measuring the same skill or that students were able to work equally across all groups. On average, students obtained 46% of the available marks for group A, 40% for group B and 49% for group C: the differences probably reflect the marking scheme rather than the difficulty of the questions.

We used a general linear model to investigate the differences between various groups of students in the marks they obtained for questions in group A, B and C. The models used sex, non-English speaking background, length of time in Australia and the particular course of enrolment as explanatory variables.

**Figure 1.** Marks obtained in group C questions vs years in Australia




The only statistically significant differences were due to interaction between sex and recent arrival in Australia on the marks achieved in “group C” questions. We investigated these differences first by categorising students into those who had arrived in Australia since 1989 (ie those who were likely not to have done the *whole* of their schooling in Australia). We then categorised students into those who had arrived in Australia since 1994 (ie those who were not likely to have done their *secondary* schooling in Australia). Finally, we used years in Australia as a covariate (making the assumption that the students born in Australia were 20 years old).

Looking at the students who had arrived since 1994, the males obtained significantly lower marks (mean C% = 27, $p = 0.001$) in group C questions than all the other groups—female recent arrivals and males and females who had been done their secondary schooling in Australia (mean C% = 55, 53 and 47 respectively). Looking at students who arrived since 1989, the pattern of results was similar although the differences did not quite reach statistical significance ($p = 0.067$). Using years in Australia as a (continuous) covariate, the sex-by-years interaction was significant ($p = 0.014$) and showed the same general picture: males who had not been long in Australia performed lower than other groups on group C questions.

With the exception of this one finding, no other variables or (two-way) interactions showed any significant effects on performance.

# 6. Conclusion

People who did well overall scored evenly on all groups. This need not have been the case, since the high proportion of group A tasks made it possible to reach high scores without doing particularly well on groups B and C. On the other hand, students who did badly had a mixed performance on the various groups. Two students performed very well in group B and C tasks but not in A. One of these students had a sick wife and, whilst he understood the work well, did not have time to practice the routine procedures. This unusual case shows that it is possible for students who do not perform well at routine procedures to demonstrate deep learning. In general, though, we find that the correlation between A% and the average of B and C% is a very high 0.83. Investigation of outliers may give interesting insights to learning.

There is considerable disquiet amongst mathematics lecturers at tertiary level as to the routine algebraic skills of incoming students and of students studying mathematics at university (see for example the *ICMI Study into the Teaching and Learning of Mathematics at University Level*, 2001). There is a conjecture that students who have poor technical skills are not able to succeed in university mathematics. The contrapositive conjecture that good technical skills (such as algebraic dexterity) are necessary for success in university mathematics is often taken for granted. The taxonomy allows us to test this hypothesis as we can compare performance in group A tasks (routine) with performance in higher level B and C tasks. We have shown in isolated cases that it is possible for students to do well in groups B and C and not in group A. It would be interesting to investigate this further. Clearly a base level of algebraic dexterity is necessary but what is that base?

In retrospect, the examination that was analysed was not ideal in that the questions contained a strong emphasis on routine skills. We suspect that the length of the examination benefited those students who had memorised material and who had practiced techniques. The finding in our previous study (Smith and Wood, 1998) that females scored a higher percentage of their total mark on group A tasks was not replicated. In the present study the same pattern was evident, but was not statistically significant. More work along these lines would be interesting.

Without setting out to test this particular idea, we found that male students who had recently arrived in Australia (but not female recent arrivals) scored significantly lower on group C questions. We are not sure what this suggests. It can’t be simply be due to language, or the females would show the same pattern. We need to investigate this further by interviews with these students and consider teaching interventions to improve their performance.

The hypothesis that non-English speaking background students had "difficulty with the conceptual aspects of the course" was investigated. The variable showing language background was not significant in any model, singly or in interaction with any other variable. In fact, both groups scored an average of 49% on the C questions.

## REFERENCES


-Ball, G., Stephenson, B. Smith, G.H., Wood, L.N., Coupland, M. & Crawford, K., 1998, Creating a diversity of experiences for tertiary students, *Int. J. Math. Educ. Sci. Technol. 29*, 6, 827—841.

-Biggs, J.B. & Collis, K.F., 1982, *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York: Academic Press

-Biggs, J., 1999, *Teaching for Quality Learning at University*. SRHE & OUP, UK.

-Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H. &Krathwohl, D.R., 1956, *taxonomy of Educational Objectives: Cognitive Domain*. New York:McKay.

-Brockbank, A. & McGill, I., 1998, *Facilitating Reflective Learning in Higher Education.* SRHE &OUP UK.

-Brown, G., Bull, J. & Pendlebury, M., 1997, *Assessing Student Learning in Higher Education.* Routledge, London.

-Crawford, K., Gordon, S., Nicholas, J. and Prosser, M., 1994, Conceptions of mathematics and how it is learned: perspectives of students entering university. *Learning and Instruction*, 4, 331–345.

-Holton, D., 2001, *ICMI Study on the Teaching and Learning of Mathematics at University Level.* Boston:Kluwer.

-Krantz, S., 1999, *How to Teach Mathematics.* American Mathematical Society: Rhode Island

-Ramsden, P., 1992, *Learning to Teach in Higher Education*. NY:Routeledge.

-Reid A. and Petocz P., 2001, Using Professional Development to Improve the Quality of Assessment Tasks and Student Learning Environments, in *Improving Student Learning Strategically* 8, 161-168, Oxford Brookes.

-Smith, G.H. & Wood, L.N., 1998, Examination responses to changes in undergraduate mathematics assessment. *International Conference on the Teaching of Mathematics.* Samos, 269–272.

-Smith, G.H., Wood, L.N., Coupland, M., Stephenson, B., Crawford, K. & Ball, G.,(1996). Constructing mathematical examinations to assess a range of knowledge and skills*; Int. J. Math. Educ. Sci. Technol., 27*, 1, 65-77.

-Wood, L.N. & Perrett, G., 1997, *Advanced Mathematical Discourse*, Sydney:UTS.

-Wood, L.N. & Smith, G.H., 1999, Flexible assessment. In *The Challenge of Diversity* (Eds. Spunde, W., Cretchley, P. & Hubbard, R.). Laguna Quays: University of Southern Queensland Press, 229-236.