Students' perceptions of their understanding in *Chemistry 1* for Veterinary Science

Justin Read, **Adrian George** and **Anthony Masters**, School of Chemistry, University of Sydney
Mike King, Faculty of Education and Social Work, University of Sydney
j.read@chem.usyd.edu.au george@chem.usyd.edu.au a.masters@chem.usyd.edu.au
m.king@edfac.usyd.edu.au

Introduction

The aim of this study was to investigate the relationship between students' perceptions of their understanding of chemistry, and their performance as measured by the end of semester examinations. Prior to commencing the study, it was hypothesised that there should be some correlation between students' perceived understanding and exam performance. Furthermore, experience suggested that high achieving students are generally better able to identify their strengths and weaknesses than are weaker students. It seemed logical, therefore, that the strength of any correlation should vary with exam performance. This study was designed to test this hypothesis, and this paper is the first refereed report of results from this on-going investigation.

A search of the literature found no previous studies of direct relevance to this work. However, the literature does offer some background. A number of studies have examined students' perception of their exam performance after completing an exam (e.g., Beyer, Riesselmann and Warren 2002), and students' overall expectations of academic performance has also been examined (e.g., de Campos, Grinberg, Garcia, Parise, da Silveira and Dumont 1998). Both are poor predictors of academic performance. Student self-marking has been shown to correlate well with the marks given by their professors for lower-order cognitive skills questions, but not for questions requiring high-order cognitive skills (Zoller, Fastow, Lubezsky and Tsaparlis 1999). Academic self-efficacy (confidence in one's ability to complete academic tasks) has been shown to be positively correlated with academic performance (Chemers, Hu, and Garcia 2001; Vrugt, Langereis and Hoogstraten 1997). However, the Chemers et al. (2001) study examined generic skills and overall performance in a degree program, and was not linked to a domain. The Vrugt et al. (1997) study examined psychology freshmen, and whilst subject matter understanding was included in their model, they found that 'self-efficacy and goals accounted for 5% of the variance in exam performance' (p. 67), and thus their model has a poor predicting power for student achievement. House (2000, 2003) examined self-beliefs (measuring agreement/disagreement with statements such as 'Science is boring', 'I enjoy learning Science' and 'Science is important to everyone's life') amongst 13-year-olds. These studies found a correlation between self-beliefs and science achievement test scores, but these beliefs were also found to be poor predictors of performance, explaining 6.29% of the variance in test scores in Hong Kong (House 2003, p. 201) and 6.8% in Ireland (House 2000, p. 110).

Method

CHEM1405 unit of study

The CHEM1405 (Chemistry 1 for Veterinary Science) unit of study is a compulsory, one-semester-long subject, taken in the first semester of the Bachelor of Veterinary Science degree program. It consists of 24 lectures of general/inorganic chemistry (equilibrium, thermodynamics, osmosis, acids and bases, redox/electrochemistry and kinetics, etc.) and 28 lectures of organic chemistry (simple organic transformations, spectroscopy, amino acids and proteins, carbohydrates and DNA). Students also undertake 27 hours of laboratory work during the semester. Some prior knowledge of chemistry is assumed (to around HSC level), and those students who have not previously studied chemistry are strongly encouraged to complete a bridging course prior to commencing the unit of study.

The unit of study is assessed primarily by a 3 hour long end-of-semester exam, which constitutes 75% of the assessment program. This exam is divided approximately equally between multiple choice and short answer questions, and between general/inorganic and organic questions. Analysis of the exam scripts for this study has involved both quantitative (statistical) and qualitative methods, with the qualitative analysis looking for evidence of misconceptions and examining commonalities in student approaches (both correct and incorrect). Some of the results from this analysis are summarised elsewhere in this volume (Read, George, King and Masters 2004b).

Participants

Students in the CHEM1405 unit of study were surveyed in the final teaching week of Semester 1, 2003. The survey instrument used asks students for some general background information (gender, age, prior study of chemistry), their student identification number (SID—provision of which was optional) and for their opinion of the most and least positive aspects of the unit of study. Respondents are then asked to rate how well they believe they understand, and can apply that understanding to answer questions about, 17 nominated areas of the unit of study on a 7 point Likert scale (1 = no understanding to 7 = excellent understanding).

The exam scripts for this unit of study were then analysed. For those students who provided an SID, exam performance was compared to their perceived understanding of the areas nominated in the survey instrument. The performance of the respondents was compared to the CHEM1405 cohort as a whole, to check that the respondents constituted a representative sample of the student population. As a preliminary exercise, an in depth interview was also conducted with one student, 'Karen', to provide some further insight into the students' perspective. It is intended that more interviews will be conducted with CHEM1405 students from 2004 during semester 2.

Results and Discussion

Participants and Response Rate

93 students completed CHEM1405 in Semester 1, 2003, of whom 52 participated in the study by returning the survey. A total of 36 respondents provided both an SID and sufficient information for the full analysis to be completed, and thus the overall response rate for full participants was 39 %.

Responses show that most of the respondents were recent school leavers (69% aged 17-19), and most had a good background knowledge of chemistry (78% had completed HSC chemistry or its equivalent). Only 5% had no prior knowledge of chemistry, and a further 5 % had the bridging course as their only prior study of chemistry.

Sample Representivity

Figure 1 shows the distribution of exam grades achieved by the full participants, and also the distribution for the whole student cohort. Exam results are unscaled, and it should be noted that these represent only the result from the end-of-semester exam, and so grades do not include other parts of the assessment program, such as tutorial quizzes and laboratory work. There is no statistically significant difference between the exam grade distributions of the respondents and non-respondents ($\chi^2 = 0.69$, df = 4, p = 0.95), nor is there any difference in distribution of background knowledge ($\chi^2 = 6.4$, df = 3, p = 0.093) nor gender ($\chi^2 = 0.20$, df = 1, p = 0.65).

Figure 2 shows the distribution of exam grades for different levels of background knowledge, for both full participants (Figure 2(a)) and the whole cohort (Figure 2(b)). In Figure 2, students have been categorised according to their level of prior study of chemistry. Students who had only completed a bridging course were classified as having a poor background; students who had studied chemistry to HSC level (or its equivalent) were classified as having a good background; and, students who had undertaken prior university level study of chemistry were classified as having an excellent background. There are no statistically significant differences between these distributions, and from this and the above statistical tests we can confidently conclude that the full participants are indeed a representative sample of the student cohort as a whole.

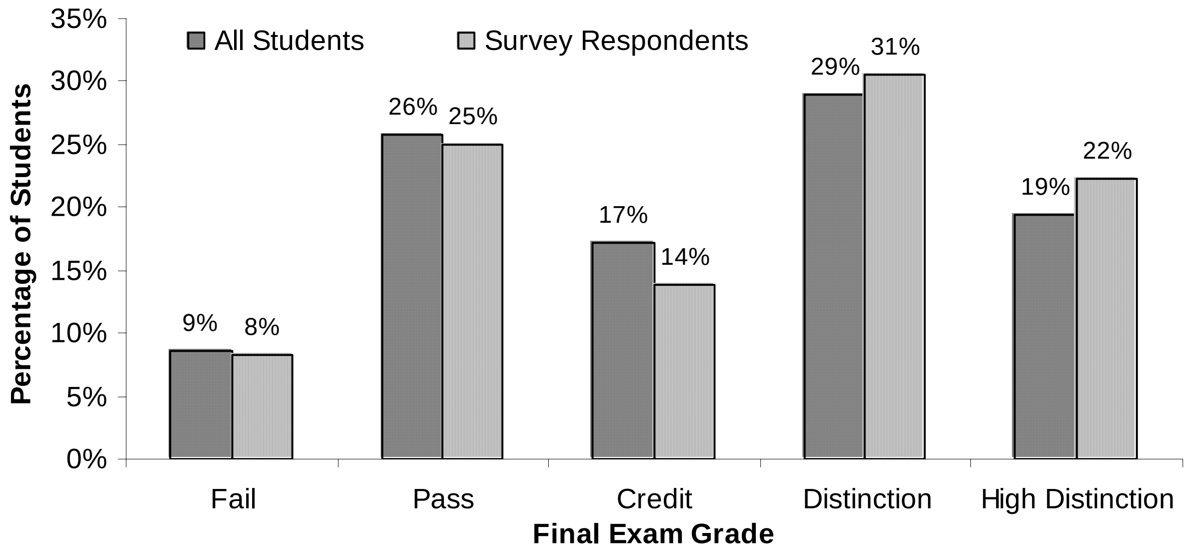

Figure 1. Distribution of exam grades for survey respondents and the whole CHEM1405 student cohort

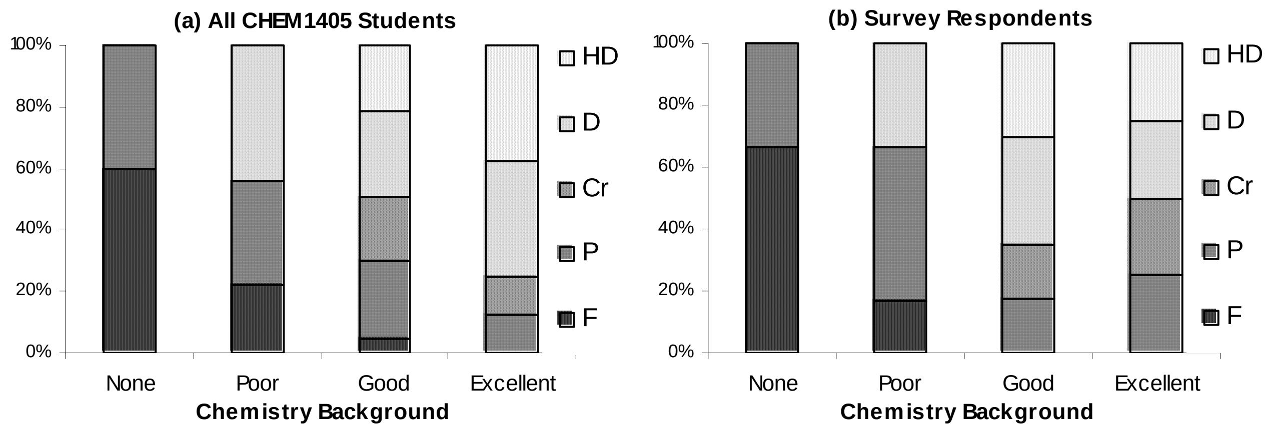

Figure 2. Distribution of exam grades for different levels of prior knowledge for (a) all CHEM1405 students and (b) the full participants in the survey

Figure 2 also shows that exam performance improves substantially with prior study – a finding that is not particularly surprising. However, the difference in performance between those students for whom a bridging course was their only prior study (poor background), when compared to those with no prior chemistry study, is surprisingly large. The bridging course that these students undertook comprised 13 hours of lectures, plus 26 hours of small group tutorials, spread over a 7 day period. The fact that such an intense but short period of study is associated with a reduction in exam failure rates by a factor of about 3, and seems to provide the basis for students to complete university level study at distinction standard, is a testament to its value, and is the subject of on-going investigation.

Individual Correlations

The first analysis performed on the data sets from the full participants was aimed at determining the strength of the correlation between each participant's perception of their understanding of nominated areas of chemistry, and their performance in those areas. Figure 3(a) shows the scatterplot for Karen, which is representative of most of the results obtained. Figure 3(b) shows the scatterplot for the student with the strongest correlation between perception and performance data.

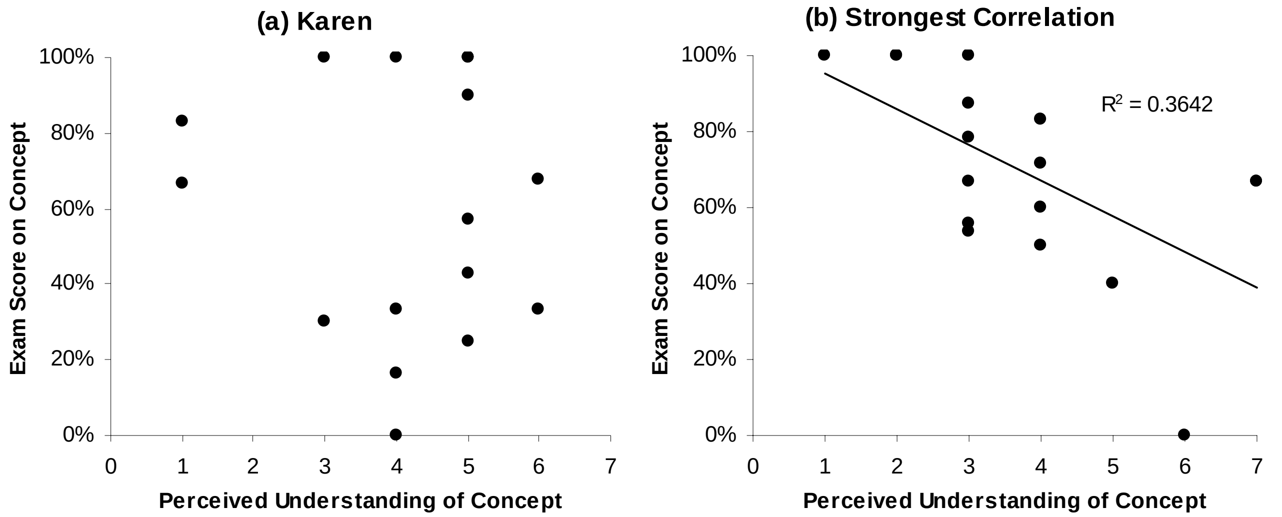

Figure 3. Scatterplots of perceived student understanding of nominated concepts against performance in exam questions on those concepts for (a) Karen, a typical CHEM1405 student, and (b) the participant with the strongest correlation

The results shown in Figure 3 were unexpected, especially given our starting hypothesis. In typical cases, there is no correlation between these factors (average $R^2 = 0.079$). In the case of the participant with the strongest correlation, whose R^2 is the largest by a considerable margin – the next largest was $R^2 = 0.255$ – although a weak correlation is present, it is a negative correlation. That is, it indicates that the student performed best in those areas where they believed they had the least understanding, and performance decreased as perceived understanding increased.

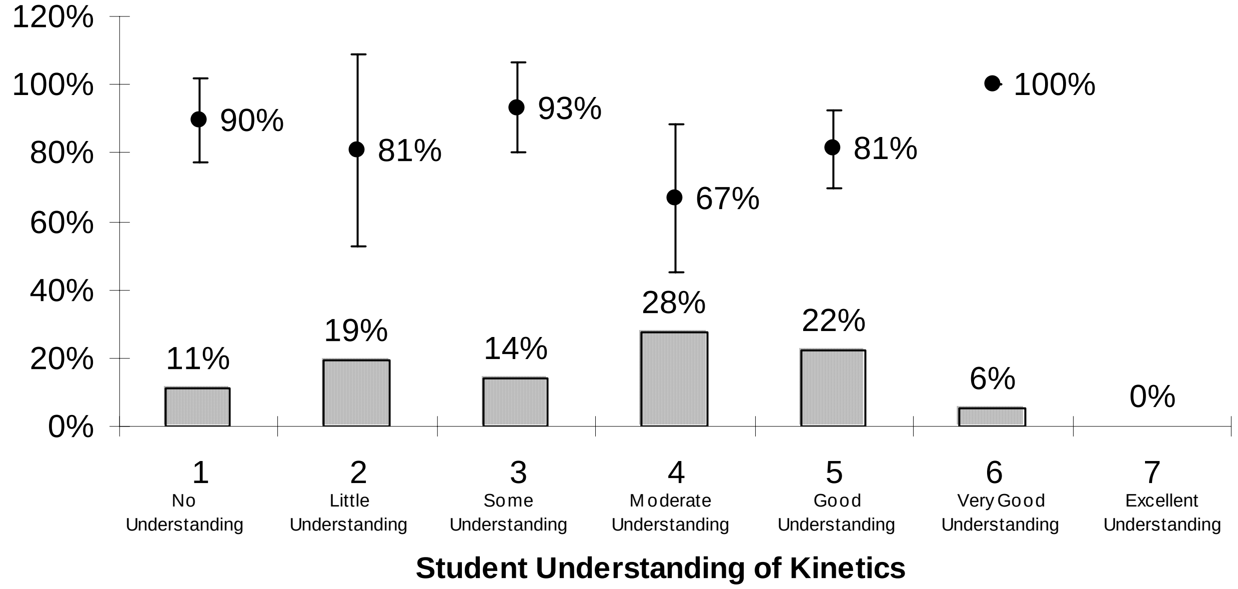

Figure 4. The distribution of reported student understanding of kinetics (bar graph, ▒) and 95% confidence intervals (●) of mean exam performance for each level of understanding

Since the analysis suggests that perceived student understanding and exam performance in individual areas are unrelated, it was decided to examine performance in a number of individual areas of the unit of study. One such area was kinetics, which the full participants rated one of the least understood concepts in CHEM1405. Figure 4 shows the distribution of reported levels of understanding of kinetics as a bar graph. For each of these levels of understanding, the 95% confidence interval for the mean exam performance has been constructed and plotted.

Figure 4 indicates that (for example) 11% of respondents reported that they had no understanding of kinetics. Despite this fact, the average exam score in the kinetics questions on the exam was 90 ± 12 % for these (perceived understanding = 1) students. Looking at the mean exam score over the range of reported levels of understanding, it is clear that exam performance on kinetics questions is independent of perceived understanding of kinetics. This was an unexpected result. Analysis of areas with higher reported perceived understanding also showed no substantial correlation between perceived understanding of a concept and exam performance on that concept.

Do we conclude from such results that students – at least, this group of students – are unable to identify what they do and do not understand? To address this question, we can look first to the interview with Karen. She was one student who rated her understanding of kinetics as '1', but she performed well on the kinetics questions (scoring 83%). When asked to comment, she said:

> That's because they repeated the question, and I looked at the exam from the year before, and I practiced those questions … if they hadn't of pretty much repeated the question, like, I would've been stuffed.

Karen went on to explain how she had now developed an understanding of kinetic phenomena, such as rate constants and half-lives, from second semester study. She gave a long example from a recent lecture concerning the use of thiopentone as an induction agent, discussing its biological half life in a number of different body tissues, and concluding:

> … and, I've now gone, oh my god, so this is just for this one drug that we've had a lecture on. And the *whole* mechanism by which it works, and its whole relevance to veterinary practice is based on its $t_{½}$. Like, because we know its $t_{½}$, we know how to use it.

It seems clear from these comments that Karen genuinely did not understand kinetics when she completed the first semester exam, relying instead on algorithmic and rote learning approaches to these exam questions. Qualitative data from exam script analysis suggests that such approaches were widely used by students in this unit of study (Read, George, King and Masters 2003; Read, George, Masters and King 2004a; Read et al. 2004b). There is ample evidence of algorithmic approaches and a lack of genuine understanding in the answers to questions in a number of areas of the exam. It seems likely, therefore, that the lack of correlation between student perception of understanding and exam performance in individual areas can, at least in part, be attributed to a failure of the exam to measure student understanding.

Group Correlations

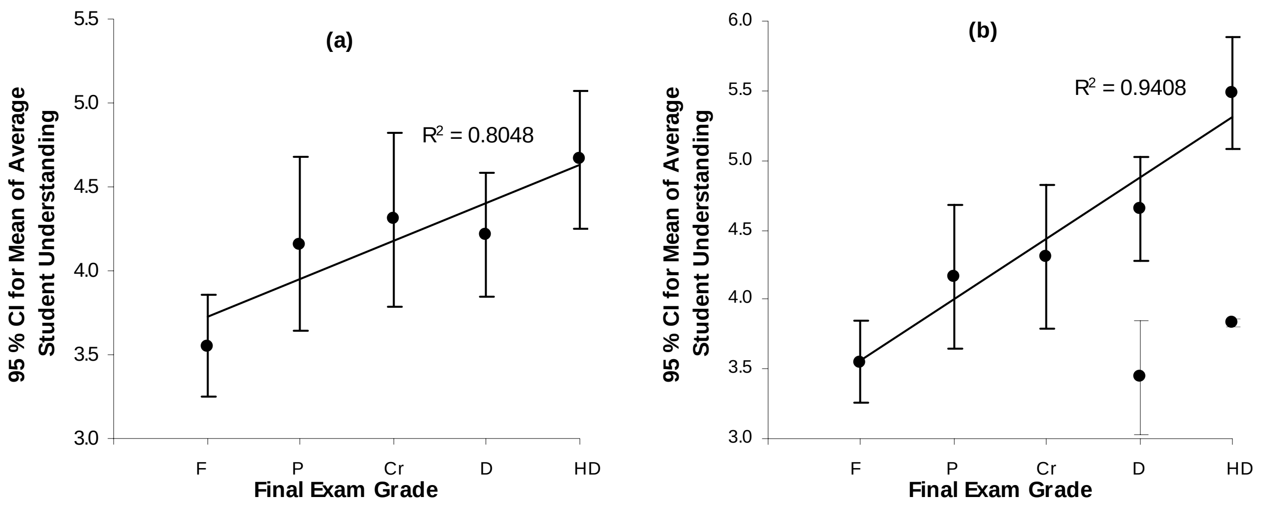

Figure 5. Correlation between average student understanding and exam grade grouped by grades for (a) all full participants and (b) with the perfectionist group separated

Since no correlation was found in individual areas of the unit of study, it was decided to check for any correlation between average understanding and overall exam performance. To do this, the reported levels of understanding for each of the nominated areas were averaged for each full participant. The data were then grouped by exam grades (rather than marks), and a 95 % confidence interval for the mean of average student understanding was constructed. When these confidence intervals were plotted against exam grades, a reasonably strong correlation emerged, Figure 5(a).

The correlation in Figure 5(a) is surprisingly strong, given the complete lack of correlation in individual areas of the unit of study, and especially since no correction has been made for exam weighting. That is, no correction has been made to compensate for the fact that the distribution of marks between the different nominated areas of the unit of study in the exam was not even.

When the data set within each exam grade is examined, it is apparent that the distributions are approximately normal for the lower grades, but bimodal for the upper grades. That is, there is a group of high achieving students, whom we term ‘perfectionists’, who report average levels of understanding much lower than would be expected given their overall performance. Figure 5(b) shows the correlation between average performance and understanding once this perfectionist group is separated. Performing this separation has two effects—it increases the gradient, so that average understanding is spread over a greater range, and it increases the strength of the correlation. (Investigations are continuing to try to identify the characteristics that make the perfectionist group different from other high achieving colleagues.)

The correlation seen in Figure 5(b) is remarkably strong, and further analysis was completed to check that this result is not an artefact created by the analysis. Since exam grade bands are not equally wide, the first check was to reconstruct Figure 5(b) using the same *y*-scale, but changing the *x*-scale to 95 % confidence intervals on the mean exam score for these groups, Figure 6.

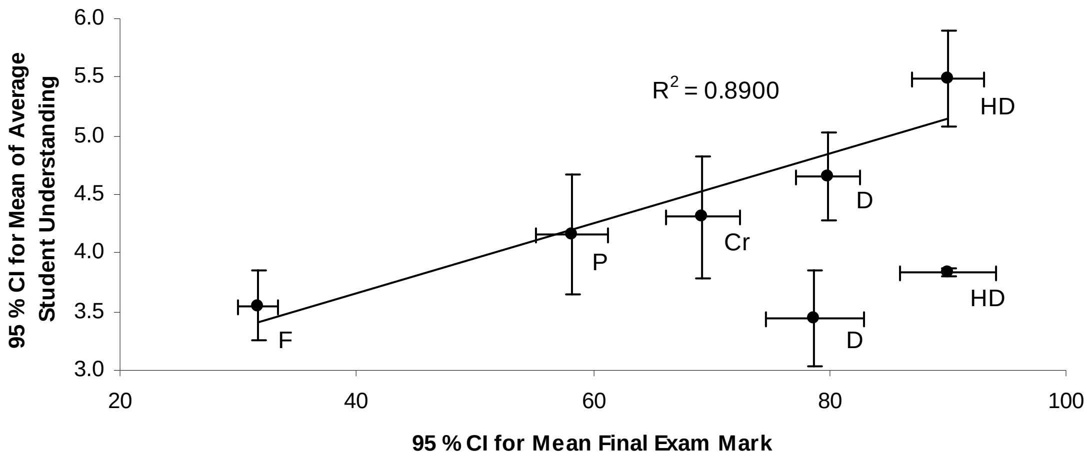

Figure 6. Correlation between average student understanding and average exam score

Figure 6 shows that the correlation between average perceived understanding and overall exam performance remains strong when marks are used. Results from other units of study (Read et al., 2003) show near identical results, with stronger correlations found in units of study with larger enrolments, where there were more full participants. It therefore seems unlikely that the observed correlation is merely a statistical artefact, although the reasons for its existence remain unclear. One possible explanation is that students who possess good exam technique, or higher self-efficacy, report generally higher levels of understanding—and further investigation is needed in this area.

Conclusion

The results from this study to this point show that there is no correlation between students' perceived understanding and their exam performance in individual areas of the CHEM1405 unit of study, a fact we believe can be attributed to the exam not measuring student understanding. The results also show that prior knowledge has a significant impact on exam performance. We have also shown that there is a strong correlation between average understanding and overall exam performance. However, the reason for the existence of this correlation, given our result that perception and performance in individual areas do not correlate, remains unclear.

References

Beyer, S., Riesselmann, M. and Warren, T. (2002) *Gender differences in the accuracy of self-evaluations on chemistry, english and art questions.* Paper presented at the Annual Meeting of the American Psychological Society, New Orleans, USA.

Chemers, M.M., Hu, L.T. and Garcia, B.F. (2001) Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology,* **93**(1), 55–64.

de Campos, R.C., Grinberg, P., Garcia, I., Parise, J. A., da Silveira, M. A. and Dumont, N. (1998, August) *Self-Image and expectations of first year engineering students at a Brazilian university.* Paper presented at the International Conference on Engineering Education, Taipei, Taiwan.

House, J.D. (2000) Student self-beliefs and science achievement in Ireland: Findings from the third international mathematics and science study (TIMSS). *International Journal of Instructional Media,* **27**(1), 107–115.

House, J.D. (2003) Self-beliefs and science and mathematics achievement in Hong Kong: Findings from the third international mathematics and science study (TIMSS). *International Journal of Instructional Media,* **30**(2), 195–212.

Read, J.R., George, A.V., King, M.M. and Masters, A.F. (2003) *Unpublished results*.

Read, J.R., George, A.V., Masters, A.F., and King, M.M. (2004a, February) *Correlation between individual student perception of understanding and examination performance in chemistry 1 for vet science.* Poster presented at the Annual Meeting of the Chemical Education Division of the Royal Australian Chemical Institute, Hobart, Australia.

Read, J.R., George, A.V., King, M.M. and Masters, A.F. (2004b, October) *Exam script analysis – A powerful tool for identifying misconceptions.* Poster presented at the Annual Symposium of UniServe Science, Sydney, Australia.

Vrugt, A.J., Langereis, M.P. and Hoogstraten, J. (1997) Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *The Journal of Experimental Education*, **66**, 61–72.

Zoller, U., Fastow, M., Lubezsky, A. and Tsaparlis, G. (1999) Students' self-assessment in chemistry examinations requiring higher- and lower-order cognitive skills. *Journal of Chemical Education*, **76**(1), 112–113.