The Journal of Scholarship of Teaching and Learning (JoSoTL)

The Journal of Scholarship of Teaching and Learning

Sponsored by UCET, FACET, and Indiana University South Bend

Volume 3, Number 1 (2002)

Development of a Simple Mathematical Predictor of Student Performance in General Chemistry

Nathan W. Bower

Chemistry Department
Colorado College
Colorado Springs, CO 80903

Abstract

Colorado College uses a sequential course structure exclusively in its calendar that is similar to those used in summer programs at other institutions. In this approach, a student takes, and a faculty member teaches, only one four-hour course each month. This format enhances longitudinal studies of the factors that affect student grade performance and retention. In this study, standard predictors of success, such as ACT and SAT scores, are compared with a simple mathematical background knowledge probe. Other factors that may impact student performance such as economic background, gender, learning style, and time between courses are also discussed.

Introduction

One of the major problems facing faculty who teach science courses that have a significant reliance on a mathematical foundation is determining whether individual students entering the class have the appropriate preparation in this ancillary area. Compounding the difficulty of determining and then enforcing pre-requisites are other factors that can enter the picture, such as math anxiety as a separable issue from math or science competence.[1] A simple background probe that can be administered in a few minutes and which is relatively free of confounding bias is needed to provide important feedback at the outset of the course. Such an instrument would allow the teacher to do some last minute fine-tuning of the course level as well as offering the opportunity for scheduling of individualized remedial help.

Background

The past decade has seen a fierce national debate over the validity of using standardized exams such as the SAT in college admission decisions, especially as affirmative action has come under attack.[2] Newsweeklies expound upon these tests and their role in our society, and it seems likely the Supreme Court will soon have to sit on their constitutionality.[3,4] In addition to being used for admissions and financial aid, these tests are also sometimes used to replace college-wide requirements and for placement in courses.[5] Some years ago, Pickering did an interesting long-term study using the SAT math score to identify students *a priori* who were expected to do poorly in General Chemistry. In a controlled experiment, he offered an intensive supplementary course in problem solving to a subset of the students with SAT math scores below 610. The modest improvement in grade noted (0.41 on a 4.0 scale for n = 43) versus the control group and the effect on their subsequent General and Organic Chemistry grades (0.17 and –0.08, respectively) for the same students raises questions about the long-term efficacy of such efforts.[6,7]

Despite the disheartening results in Pickering's study, science teachers continue to try to identify which students are likely to need help as early as possible. This seems especially important with the intensive course structure known as the "block plan" in which students take (and instructors teach) one course at a time for about a month. This structure is used at many

[1] Hembree, R., The nature, effects, and relief of mathematics anxiety, *Journal of Research in Mathematics Education,* **21**, 33-46 (1990).

[2] Mealer, B., Moves against affirmative action fuel opposition to standardized admissions tests, *Chronicle of Higher Eduction*, **48**(8) A40-A41 (Oct. 17, 1997).

[3] Cloud, J., What does SAT stand for? *Time*, **150**(20) 54-55 (Nov. 10, 1997).

[4] Lehman, N., Behind the SAT, *Newsweek*, **134**(10) 52-57 (Sep. 6, 1999).

[5] Coley, N., Prediction of Success in General Chemistry in a Community College, *Journal of Chemical Education*, **50**(9) 613-615 (1973).

[6] Pickering, M., Helping the High Risk Freshman Chemist, *Journal of Chemical Education*, **52**(8) 512-514 (1975).

[7] Pickering, M., The High Risk Freshman Chemist Revisited, *Journal of Chemical Education*, **54**(7) 433-434 (1977).

institutions for summer session courses, but it is used for the entire calendar of courses at Colorado College. One of the advantages of the block plan is that it fairly readily allows students to switch courses on the first day of class without increasing their years of matriculation if they can quickly determine what course is best for them. Even if a placement quiz or "knowledge probe" at the start of a course is not used to determine who may need extra help, it can provide students with additional information about whether they are in the appropriate course for their background and interest level.

Another use of such knowledge probes is to help in selecting groups for cooperative learning strategies. Because students exhibit a variety of learning styles,[8] it is useful to identify those that are stronger or weaker in the traditional algorithmic approaches. Depending on the tasks set by the instructor, it may then be desirable to form groups that are diverse in their abilities, or, if it makes sense is to spend more time with those groups needing additional help, it may be desirable to form groups with similar backgrounds.

At Colorado College, courses are limited to 25 students and there are up to eight different faculty teaching General Chemistry in any given year. Approximately 200 students take the introductory chemistry course each year, or about a third of each graduating class. No differentiation is made between students majoring in chemistry or any other field. General Chemistry I and II are offered almost every block, or nine and eight times per year, respectively. The courses are equivalent to the first and second semester of General Chemistry taught elsewhere, and each block is three and a half weeks long (followed by a half week break for grading and setting up the next course).

The unique nature of the block plan provides a "laboratory" for testing new ideas in education. With different teachers and so many students from varied backgrounds involved in the introductory courses, it is possible to obtain data on a host of variables with minimal confounding.

Because SAT and ACT scores are not routinely available to chemistry faculty at Colorado College for reasons of privacy, different "quizzes" have been devised that can be taken in a few minutes on the first day of class. The quizzes are taken without calculators, and they are designed to ask a few questions that are a little outside of the routine "algorithmic" approaches students learn, especially those that rely on a calculator. The results of these quizzes are shared with the students immediately so that they will be able to determine for themselves whether they need to arrange for additional tutoring or whether they should postpone the course until they are better prepared. Often the quizzes are simply exchanged with a neighbor and the scoring is covered in a couple minutes as a method of nearly instant feedback. In those instances, the quizzes may not even be collected, so the instructor may get no direct feedback on a given student's needs, and the student is given full responsibility for their own decisions regarding what to do with the results.

The purpose of this paper is to report on the efficacy of two of these quizzes, comparing them with other predictors and factors affecting long term "success". Success in this case is

[8] Felder, R. M., Matters of Style, *ASEE Prism*, **6**(4) 18-23 (1996).

measured by grades in subsequent chemistry courses, but other measures, such as retention and the number of subsequent courses taken are also considered. Personal reflections on what sorts of knowledge probes need to be developed in order to continue to improve through this classroom action research process are included in the ensuing discussion.

Experimental

Five General Chemistry II courses (n = 132 students) spread over five years (1995 to 2000) were randomly selected from those courses in which a two-question math quiz had been administered. Five additional courses (n = 117 students) were selected from those in which a seven-question quiz (Appendix 1) had been given. Three of these were General Chemistry I courses and two were General Chemistry II courses. The first two questions of the seven-question quiz are the same as those used on the two-question quiz. These courses involved two different teachers with varying degrees of cooperative learning strategies incorporated in their courses.

Additional retrospective information was obtained from student transcripts, such as the SAT and ACT scores for math and verbal reasoning, total financial aid, work-study grants, and grades in prior and subsequent chemistry courses. Information was also collected on how long students waited between courses, who taught each course, the format used in the course, length of time between courses, self-reported ethnic background, and the gender of the student. This information was correlated using multiple linear regression and ANOVA between the predictors and factors. Minitab version 12 was used throughout the analysis.

In one of the courses (a typical course of 25 students in spring of 2000), data was collected in class on the learning preferences of the students using the index of learning styles developed by Felder and Soloman.[9] The students self-scored this instrument, and made use of the suggested published strategies as they saw fit. Throughout the course, the instructor made the point of sharing with the students the various assessment tools being used and what was learned from them. Because the students saw themselves as involved in an experiment in this class, they seemed more involved in how the course was taught and in thinking about how to optimize their learning right from the start of the course. In the following year (spring of 2001), the same instructor in a matched class of students in terms of class size, content, text, diversity, and timing, repeated the experiment but without informing the students about the educational research aspects and assessments tools being used until the end of the course. Such classroom action research with individual courses is common, and these experiments add to the enthusiasm that both the instructors and the students feel in these courses.

Results

At Colorado College, students wait an average of 4.7 months between General Chemistry I and II, although they are advised to take them within the same semester. The average waiting period lengthens to 7.4 months between

[9] Soloman, B. A., and Felder, R. M., Index of Learning Styles, North Carolina State University, www2.ncsu.edu/unity/lockers/users/f/felder/public/ILSpage.html (2000).

General Chemistry II and Organic Chemistry I. Between the Organic Chemistry I and II, this period shortens to 2.6 months. Chemistry majors tend to take their courses somewhat closer together than this, and the Organic I and II progression reflects this as only Biochemistry and Chemistry majors (and those planning to attend medical school) are required to take the Organic Chemistry II. On the other hand, Geology majors (who are required only to take General Chemistry I) often wait three years between the first and second course. Only those Geology majors planning to continue into graduate school in the field return to take General Chemistry II.

Despite these widely varying times between courses, there is little evidence that students' *grades* were impacted by either putting off their chemistry courses or by taking them back-to-back. This may possibly be because most students who go on continue to mature in a parallel science and they bring that mental maturity with them. Alternatively, this may indicate that the learning strategies and motivation that the majority of students bring to their courses is more important to their overall success than the content we manage to impart in our chemistry courses, despite how sequential we think they are. (Students who skip a course or take them out of sequence *do* suffer, however, so there is something they are mastering even if we can't find an adequate test for it.) Whatever the reason, the gaps between courses are not drastically different from that experienced by students in a semester system, where the average time between material in successive semesters is 4.5 months (mid-fall to mid-spring) and 7 months (mid spring to mid fall).

The profile of the students entering General Chemistry II is typical of the student body as a whole, with 60% on financial aid and two-thirds of these on work-study. About 55% are female, 16% are a self-identified ethnic minority (about half are Hispanic), and the median SAT Math and English scores are 630 and 620 respectively. The average grade obtained in the first course is a 3.0 on a 4.0 scale. This drops slightly to a 2.9 in General Chemistry II and a little further to a 2.7 in Organic Chemistry I. The average returns to a 2.9 in Organic Chemistry II as primarily majors in chemistry, biochemistry, neuroscience, and pre-medical students take it.

Based on Soloman and Felder's four-dimensional Index of Learning Styles,[10] a given class will be moderately (but significantly) more visual than verbal and somewhat (but significantly) more active than reflective in their learning style preferences. The class will also be slightly (but not significantly) more sequential than global, and nearly equally balanced on the sensing versus intuitive dimension. Although the averages fall near the middle of the scale, at least a third of the students will have a strong preference for at least one learning style. Out of 25 students, each of the different dimensions were represented by at least two students with a strong preference for that style except the reflective and verbal dimensions, and even these had more than one student with a moderate preference in that direction. This profile matches the expected general science student population at Colorado College, which attracts outdoor-oriented, athletic students who like a balance of creative

[10] Soloman, B. A., and Felder, R. M., ibid, (2000).

outlets to go with the intensive study of the block plan.

These profiles indicate that using a variety of approaches on every major topic should be expected to be necessary in order to reach all of the students. At the end of the course in which the students were actively informed of the results of the various assessment tools, student evaluations indicated that they appreciated the efforts made to respond to their different styles of learning, and most students were more proactive than similar classes in trying to make the best use of the resources geared to their preferences. The following year's class (22 students), which had a slightly (but insignificantly) higher SAT-M score of 643.5 ± 40.3, 64% female, 14% minority (compared to 631.1 ± 53.2, 68% female, 16% minority), had a more typical (lower) level of engagement and interest in the class. The final standardized exams from the American Chemical Society, which are designed to have normal distributions around 50%, were also nearly equivalent (64.1% versus 62.0% for the second class compared to the first class, matching the ratio of the SAT-M scores). Although these results indicate the classroom involvement in the learning style research project had little or no impact on the student learning or grades achieved, the anecdotal evidence of retention and interest beyond the class is very much different: from the first class with the slightly lower SAT-M scores, five students (three women and two men) immediately selected the author as their academic advisor and two of these indicated an interest in majoring in chemistry, while only one from the second class did so in the in the two months following each course.

A multiple linear regression analysis of the various predictors (Table 1) shows that the two-question math quiz, graded 0 to 3 in half-units, is a better predictor of the grade a student will obtain in General Chemistry II than either the SAT or ACT math sub-tests. (Deviation from a normally distributed variable for the three predictors and the dependent "grade" is not significant despite their discrete functions.) General Chemistry II is one in which math ability plays a major role, as this is the course that deals with thermodynamics, acid-base equilibria, and kinetics. The ACT's better performance compared to the SAT may be due to a variety of factors. For example, the cram courses now available for improving a student's SAT test scores may be clouding the test's predictive power. Also, the ACT (in this study) seems to be taken by a larger percentage of students with a more diverse background of abilities and economic advantages compared to the SAT.

Table 1: National Test and Math Quiz Predictors of General Chemistry II Grades

Regression:

Predictor	Coef	StDev	T	P	ANOVA (F)
Constant	0.052	1.131	0.05	0.963	
Math-quiz	0.2549	0.1193	2.14	0.038	10.25
SAT-M	0.00027	0.0022	0.13	0.899	1.89
ACT-M	0.07982	0.0472	1.69	0.098	2.86

R = 0.504   R (adj. for d.f.) = 0.451 n = 48 (due to few students taking both tests)
Overall Regression: P = 0.005   (MANOVA F has 1 d.f./44 d.f.)

A regression of each predictor alone gives adjusted correlation coefficients of 0.381 (n=131), 0.257 (n=111), and 0.539 (n=65) for the Math pre-quiz, SAT-M, and the ACT-M tests respectively. As was noted above, the ACT test's apparently better performance is a result of the unique subset of students sampled, and Table 1 represents a better indicator of the relative merits of each test applied to the same subset of students despite the fact they are not truly independent variables.

Table 2: Math Quiz and Categorical Predictors of General Chemistry II Grades

Regression:

Predictor	Coef	StDev	T	P	ANOVA (F)
Constant	2.2372	0.3125	7.16	0.000	
Math-Quiz	0.30384	0.07654	3.97	0.000	22.94
Fin. Aid	-0.2869	0.1881	-1.53	0.130	2.92
Class	0.07920	0.09878	0.80	0.424	0.03
Instructor	0.0430	0.1448	0.30	0.767	0.22
Gender	0.04722	0.07062	0.67	0.505	0.68
Major	0.3691	0.1166	3.17	0.002	8.59
Ethnicity	-0.21308	0.09525	-2.24	0.027	5.00

R = 0.452   R (adj. for d.f.) = 0.397 (n = 129)   Overall regression: P = 0.000
(MANOVA F has 1 d.f./121 d.f.; Durbin-Watson stat. = 2.01; Lack of fit P > 0.1)

In Table 2, the two-question math quiz is coupled with a number of categorical variables and one continuous variable (financial aid, expressed as a fraction of the full cost of attending). The class variable has four levels (1 – 4), and the other variables have all been reduced to two levels (-1, 1). The underlying assumptions of the regression model are violated by departures from normality for these predictors, but the results are still useful for making some qualitative observations.

The results of this regression and MANOVA suggest that the class (first, second, third, or fourth year student), instructor, and gender of the student have little or no impact on the grade achieved in these courses. The major of the student (one of the chemistry options versus non-chemistry majors), the ethnic background of the student (Caucasian versus all others), and the financial need of the student are more important. As might be expected, chemistry majors tend to achieve higher grades, although at this stage of their careers only a small fraction have declared their major. Thus, cause and effect are still undifferentiated.

There is also a correlation of financial need and ethnic background, as students from more diverse backgrounds tend to have higher financial need at Colorado College. The financial need is usually (but far from always) coupled with more time spent on work-study, and this is time that may interfere with time spent on the course. Based on anecdotal student information, another contributor to the financial need effect that is largely independent of the student's ethnic background is that students who have a large loan often are under pressure from their families to transfer to a less expensive institution. This effect is most pronounced in the blocks taught at the end of the year as students begin to mentally disengage from the course and the institution.

While a complete analysis of the results of the seven-question pre-quiz will not be presented here because of the similarity of its results to the two-question pre-quiz, it should be noted that the seven-question pre-quiz doubles the range of possible scores (0 to 7 instead of 0 to 3). This improves its value for individual person diagnostics. It also adds a component that tests for recollection of chemical content from previous courses. As a result, it does a slightly better job of predicting the grade a student is likely to achieve. Although the results presented in this paper have focused on General Chemistry II, both the two-question and the seven-question quizzes have been administered in General Chemistry I with very similar results.

Even though memory of prior course content was not a variable emphasized in this study, comparison of the two-question prediction to the full seven-question prediction indicates that chemistry content memory is not as important as facility with math in predicting a subsequent course grade. The memory portion was also more subject to loss as the time interval between courses increased. There is little evidence that general chemistry is strictly sequential, as various textbooks order the material differently. Instead it seems there are a variety of valid starting points and the more grasp the student has of the global picture, the more easily new material can be placed in a meaningful context. Thus, the grade obtained in subsequent courses seems to be more closely connected to some longer lasting skills or a more global knowledge than it is to any specific content recollection. However, at Colorado College, content tests such as the American Chemical Society General Chemistry tests do correlate strongly with the course grade ($p = 0.000$) when they are administered at the end of the course, as does the GRE-subject test in chemistry ($p = 0.011$), which is taken by many of the majors at the end of their undergraduate career. The average grade on this latter test is comparable to the national scores for students from other schools, indicating that the block plan does allow an accumulation of content that can be measured to some extent.

Reflections and Future Directions

The efficacy of this simple knowledge probe for detecting those students who have math difficulties relevant to the course has been born out over the years. The positive impact of taking the time to involve the students in understanding their own learning processes is equally apparent, although more work is needed to find ways to make such involvement more time efficient in order to avoid additional loads on the faculty. Much less clear is what intervention measures to take with those students who have math difficulties in order to affect a long-term gain in chemistry. At a minimum, a good math review at the outset, additional tutoring outside of class, or a remedial math course is required. For some students this will prove to be adequate, but for a large percentage, something more is needed.

In a recent paper, Ashcraft and Kirk have provided some valuable insights into how math (or other anxieties) affect other performance.[11] By proving math anxiety is separable from math incompetence (and that they are independently treatable), they point the way to other testing that can indicate where students may obtain the help needed to overcome these two common hurdles. Not surprisingly, students with math anxiety often develop lower math aptitudes as they progress through their education, and as a society we must recognize (and treat) this problem in the same way we are beginning to recognize handicaps such as dyslexia.

The demonstration of the impact of math anxiety on the speed of mental processing for students who are competent in math despite having such anxiety indicates that giving these students longer to respond will allow them to reveal their actual level of competency on the subject matter of interest. Ashcraft and Kirk argue that problems that involve some form of math beyond the level of multiplication or addition tables, and which call upon other forms of memory at the same time, compete for "space" in the smaller "working memory" available to these students compared to others. An analogy might be comparing two computers, one with a smaller or "busier" CPU (due to interference from the anxiety) than the other. Both can solve the same complex problems, but because more shuffling of the data is needed in the smaller/busier CPU machine, longer time must be spent to achieve the same end result. Except for the time factor, both will achieve the same final goal. If one machine lacks the proper programs (math competence), it will not be able to solve the problem until such programming is provided, at which time it may be faster or slower depending upon multiple factors, including the "working memory" it has.

This anxiety preoccupation in the "CPU of the mind" also is relevant when the material is presented in the class if a mathematical presentation is involved. "Taking it in" will take longer for these students just as purely written (textual) presentations will take a dyslexic student longer to process correctly. The inherent abilities except for this time factor are in no way diminished. This suggests again that multiple modes of presentation are needed in order to reach the diverse population of students that we will encounter in our classes, especially if we are to help all of them achieve their full potential for contributing to society.

What does this suggest should be done in the way of modifying the knowledge

[11] Ashcraft, M. H., and Kirk, E. P., The Relationships Among Working Memory, Math Anxiety, and Performance, *Journal of Experimental Psychology: General*, **130**(2), 224-237 (2001).

probe described here? First, at least some measure of the level of math anxiety should be obtained so that the student may be steered to the appropriate source of help. Math anxiety is highly correlated with chemistry anxiety in general,[12] underscoring the need to determine the influence of this factor on student performance at an early stage. Two simple probes that may be of use for this include asking the students to report the number of high school math courses they have taken and to rank their own anxiety on a scale of 1 to 5. Both of these showed significant correlation ($p = 0.05$) with a much longer test of math anxiety that Ashcraft and Kirk employed, and this author intends to include at least these two in the next version of the background knowledge probe. As Claude Fuess once said, "I was still learning when I taught my last class."[13]

A slightly longer math test taken without calculators that includes a more active, non-mathematical, non-verbal visual test component (such as rotating or constructing actual stick and ball models of stereo isomers)[14] might also be useful for predicting academic success in Introductory Chemistry courses. This would offer a better range of responses so individuals can be more accurately diagnosed as well as keeping its predictive power for the group. It might also better test for the multiple intelligences that are correlated with the necessary skills to do well in Chemistry. The question of how sequential general chemistry really is, especially in the context of developing critical thinking skills, is still open for debate. Clearly, it would be very useful to proceed to background knowledge probes that can be administered on the first day that will determine what level of development students have achieved in the area of critical thinking skills, and then to select questions that help lead the students to move to the next level. Some work is beginning to be done in this area,[15] but a great deal of foundation still needs to be developed for a large percentage of students, including simply moving them beyond a state of anxiety.

Whether the instructor includes additional questions that probe other learning styles and intelligences, a short math quiz that does not allow the use of a calculator "crutch" taken at the start of the course seems to be a better predictor than the national ACT-M or SAT-M test. The short quiz described here with its immediate availability of results for either the students or the instructors (or both) provides a viable alternative to the much-maligned national tests.

[12] Eddy, R. M., Chemophobia in the College Classroom: Extent, Sources, and Student Characteristics, *Journal of Chemical Education*, **77**(4), 514-517 (2000).

[13] Claude M. Fuess, After 40 years at Phillips Academy, *Independent Schoolmaster*, Atlantic Monthly Press 52, http://www.bartleby.com/63/30/2530.html.

[14] Habraken, C. L., Perceptions of Chemistry: Why is the Common Perception of Chemistry, the Most Visual of sciences, So Distorted?, *Journal of Science Education and Technology*, **5**(3), 193-201 (1996).

[15] Kogut, L. S., Critical Thinking in General Chemistry, *Journal of Chemical Education*, **73**(3), 218-221 (1996).