
In-Course Optimization of Teaching Quality

Tony Bastick
University of the West Indies

Abstract

This paper demonstrates an alternative method of using students' evaluation of teaching (SETs) that circumvents many of the problems associated with traditional SETs. In particular, it shows how in-course feedback, consisting of only eight ratings, can be used to optimize post-course academic attainment.

Traditional Student Evaluations of Teaching (SETs) are feedback forms returned by students at the close of a course. Institutions intend that data from these forms be used to improve the quality of teaching and as an assessment of quality of teaching for deciding faculty promotion and tenure decisions. Although it is recognized that students can offer valuable information on the appropriateness of teaching quality, it has also been recognized that these traditional SETs are likely to have negative effects on the quality of teaching. These negative criticisms are quite extensive and range from 'dumbing down' of courses to restrictions on academic freedom (Gillmore, & Greenwald, 1999). One patently obvious criticism is that the information given by one group of students at the end of a course cannot be used to improve the teaching on that course for the very students to whom the feedback applies. Similarly, it can only be useful to future students to the extent that future groups of students are similar to the feedback group and to the extent that the course and teaching remain similar. However, courses and teaching methods hopefully evolve and the constituent subgroups of a student cohort can change considerably from one year to the next.

This paper introduces an alternative method of allowing students to assess the quality of teaching that circumvents many of the problems associated with traditional SETs. In particular it allows feedback to be used for optimizing teaching quality during the course for the whole class, for individuals or for identified subgroups of students within the whole group. The feedback is quick and cheap to process - as it requires only eight ratings from each course member. This more timely feedback and optimization allows teaching choices for keeping teaching on track for the very students who are giving the feedback. An added advantage is that the method can be used by administration at the close of a course to calculate a single indicator of teaching quality that can be used for comparative promotion and tenure decisions across the institution.

The paper out-lines the method and the theory behind it. The reliability and validity of the method is demonstrated with actual data from a course assessment. The method operationally defines three educational process objectives - Skills, Understanding and Attitudes. These three objectives are emphasized to a determined amount in the teaching and assessment of the course. Feedback forms used during the course give data on the lecturer's and students' expectations for change in these objectives. This data allows for calculations of the alignment between the lecturer's and the students' expectations for change. The theory is that academic success is maximized when students and their lecture are working towards the same changes. The theory is re-validated with each course by correlations of alignments with results, which show that in-course alignment predicts post-course academic success. This paper describes how the data are also used during the course to determine the changes that will best align in-course student/lecture expectations and so maximize predicted post-course academic attainment for the whole group or for different student sub-groups who are taking the course.

The educational importance of this alignment method is that it offers a cheap, efficient and effective alternative to the widespread problematic use of traditional SETs for quality control of teaching in tertiary institutions.

Introduction

This article briefly reports an alternative system for assessing quality teaching in tertiary institutions and focuses on the student feedback part of the system. The traditional method of assessing quality of teaching has been by questionnaires that ask students to anonymously rate the quality of teaching on a 4 or 5 point Likert scale from strongly disagree to strongly agree. In the literature the use of these forms is called Student Evaluations of Teaching (SETs).

SETs have been used in universities for more than thirty years as part of the Quality Assurance Cycle to assess the quality of teaching and as an indicator of successful teaching for promotion and tenure decisions. Unfortunately, their use has been accompanied by many counter-productive effects such as discouraging innovative teaching, and deterring instructors from challenging students (Damron, 1995; Murray, 1984; Ruskai, 1996). Although their outcomes are intended to improve teaching, a major negative effect of also using them to varying degrees for promotion and tenure decisions has been to contribute to the lowering of academic standards. Results of analyses of SETs and expected grades suggest that instructors can “buy” better evaluations via more lenient grading (Krautmann, & Sander, 1999). In the copious literature on the subject, this effect is referred to as ‘grade inflation’ or ‘dumbing down courses’ and some universities who use SETs now make statistical adjustments for these effects (Gillmore, & Greenwald, 1999). SETs have become known as little more than ‘smile sheets’ measuring popularity and ‘customer satisfaction’ (Altschuler, 1999), and lecturers have developed many methods for improving their SET scores that do not necessary improve their teaching (Crumbley, 1995). Its seems that one reason SETs continue to be used is that there has not been an expedient alternative. This article reports such an alternative - the alignment method.

The Alignment method

Cohen (1994) has introduced to education the term ‘instructional alignment’, meaning the alignment of teaching, assessment and objectives. Instructional alignment has been found to compare favourable with the use of other commonly used strategies intended to improve learning performance, such as criterion-referenced tests, curriculum-based measurement, direct instruction, learning strategies, peer tutoring, self-instructional training, cooperative learning and computer-assisted instruction (Ippolito, 1990; Redding, 1992; Vergason, & Anderegg, 1991). Elia (1994) and Walker (1998) found that instructional alignment had an unusually high positive learning effect producing substantial improvement on the achievement levels of disadvantaged and low performing low socio-economic level school students. Similarly, results of a controlled experiment with community college learners comparing verbal mediation, feedback monitoring, and instructional alignment, showed that alignment had the greatest overall effect on achievement (Breitsprecher, 1991). Biggs (1999) highly recommends educational alignment for improving the quality of university teaching. It has been found that aligning teaching, assessment and objectives in tertiary education can increase learning performance by up to two standard deviations (Cohen 1987, 1991; Cohen, Hyman, Ashcroft, & Loveless, 1989).

There are many psychometric instruments that use what is referred to here as ‘alignment methods’. In an alignment method a respondent’s current state is assessed and his/her ideal state is also assessed on the same indicators. The difference between the current and ideal states is the alignment. Where the difference is large, there is poor alignment which is indicative of problems. Where the difference is small, alignment is good which indicates that the current state is close to the ideal. Improved alignment can also be used as a measure of successful intervention. What is crucial to the alignment method is the choice of indicators measured to assess the current and ideal states. For each course, Cohen (1987, 1991) used three different objectives that were derived from the course instructor’s idiosyncratic values. However, to allow comparisons, this alignment method, following Bloom’s (1956) taxonomy of educational objectives, defines three standard process objectives that are emphasized to different degrees in quality teaching and learning.

These are Skills, Understanding and Attitudes operationally defined here as:

1. Skills – learning of facts/processes. Assessed by speedy accurate reproduction.
2. Understanding – professional competence. Assessed by justification of novel application and
3. Attitudes – professional values. Assessed by demonstration.

The method uses alignment on these process objectives as indicators of quality teaching. It should be noted that critical thinking is expected to be promoted by teaching and assessment of professional competence. This is because answers are not assessed as being right or wrong; only justifications that offer evidence of critical thinking are assessed. Alignment is not based on the course objectives. Similarly to common applications of Bloom’s taxonomy, which also emphasise different levels within domains, course objectives and content are used as vehicles for emphasizing the desired degrees of Skills, Understanding and Attitudes. This emphasis will vary according to the course level and culture of the subject taught.

What are aligned are ‘changes expected by the lecturer’ and ‘changes expected by the students’ in each of these three process objectives. Numerically stated: Alignment = changes expected by Lecturer - changes expected by students. Zero is the perfect score, the theory being that students achieve higher standards if they and their lecturer are working towards the same changes. Figure 1 shows the eight core questions that the lecturer and each student answers for the alignment to be calculated.

Figure 1: Five minute feedback form

Course Assessment - Skills, Understanding and Attitudes			
Estimate, <u>for you personally</u> , how much this course emphasises, and should emphasise (i) Skills, (ii) Understanding and (iii) Attitudes? Do this for both how the course is now , and for how the course should be - write a number in each box.			
	As it is now on this course	As it should be on this course	
(i) Emphasis on Skills <i>(getting it right)</i>	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<i>Write a Number in each box</i>
(ii) Emphasis on Understanding <i>(knowing why)</i>	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<i>Write a Number in each box</i>
(iii) Emphasis on Attitudes <i>(relevance to your life)</i>	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<i>Write a Number in each box</i>
How much so far have you enjoyed your experience of the teaching on this course?		<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	<i>Write a Number in each box</i>
Before you started this course, how much did you expect that you would enjoy your experience of the teaching on this course?		<input style="width: 50px; height: 25px;" type="text"/> Your estimate out of 100	

These forms are confidential, not anonymous. When students enrol, the process objectives are explained with generalised examples related to teaching and assessment. As part of their orientation they take a test to earn the right to be considered as informed assessors. At the start of each course, their lecturer gives subject specific examples as part of the introduction to the course. It has been found that exemplary university teachers find their own different effective teaching dimensions and strategies to achieve excellence (Hativa, Barak, & Simhi, 1999). Hence, although staff development units may advise, the teaching techniques for attaining these goals are left as a matter of informed professional choice to the lecturer.

Using data from the forms, individuals' alignments can be calculated and grouped to find the mean alignment of any student sub-group of interest - males v females, experienced v novice students, older v younger students, option 1 v option 2 students, etc.

Two alignment scores are calculated;

- Alignment of Scope (changes in absolute quantity) and
- Alignment of Proportions (changes in relative quantity)

However, these formative alignment indicators, that are measured during the course, are only predictors of quality teaching. They are not the criteria of quality teaching. The two post-course summative criteria of quality teaching are:

- Academic standards and
- Enjoyment of learning

Validation of the theory

When the courses are over and the academic results are compared with the alignment scores, it is possible to validate the theory for each course, and for each sub-group of students taking each course by correlating the 'Alignment of Scope' with 'Academic standards' and by correlating the 'Alignment of Proportions' with 'Enjoyment of learning'.

Further, when the course has finished it is possible to use sensitivity analyses on the data to calculate those lecturer's changes that would have most aligned the teaching and thus, according to the alignment theory, maximized the academic results and enjoyment of the students. It is seen from actual alignment data illustrated in Figure 2, that the choice of these preferred changes would have increased the correlations between alignment and academic standards, thus further validating the theory that alignments are predictors of quality teaching.

Illustrative Results

Figure 2: Example input, analysis and results for the alignment method

Part 1 For the Lecturer		Course	ED40C	Date	15/11/00	Name		Variables of Interest					For Validation: Academic Results	
		Skills		Understanding		Attitudes		Enjoyment	Sex	Age	Program	Option		Years Teaching
		is now	should be	is now	should be	is now	should be							
Start		30	60	40	45	60	50	50						
Best		30	30	40	44	60	71							
ALIGNMENT ASSESSMENT - DATA SHEET OPTIMISATION OF TEACHING														
Part 2 For the Students		# in class	36	# present	20									
		Skills		Understanding		Attitudes		Enjoyment	Sex	Age	Program	Option	Years Teaching	For Validation: Academic Results
		is now	should be	is now	should be	is now	should be							
#		50	50	95	98	90	95	95	2	30	3	2	6	56%
12		35	70	60	80	80	88	55	2	25	3	2	3	57%
25		80	40	20	60	40	50	30	2	25	3	2	3	43%
35		40	80	70	90	70	90	75	2	21	3	3	0	71%
30														
Part 3 Summary Results		Lecturer's changes					Alignment Predictors		Mean post-course results		Validation Correlations			
		Scope		Proportion			MScope	MProp	Enjoy	Acad	Scope	Prop		
		Skills	Underst	Attitudes	Skills	Underst							Attitudes	
		Whole	Class			n=20								
Start		100%	13%	-17%	16%	-2%	-14%	1.499	0.380	69.5%	63.0%	-0.265	-0.225	
Best		0%	10%	19%	-3%	0%	3%	0.674	0.154			-0.307	-0.594	
		Option	2	History		n=3								
Start		100%	13%	-17%	16%	-2%	-14%	1.861	0.458	60.0%	52.2%			
Best		0%	33%	10%	-4%	5%	-1%	1.221	0.299					
		Option	4	Modern Languages		n=4								
Start		100%	13%	-17%	16%	-2%	-14%	1.200	0.293	72.5%	75.7%			
Best		13%	3%	10%	1%	-2%	0%	0.223	0.065					

Figure 2 illustrates a typical data input, analysis and results sheet for the alignment assessment of a course. The method is designed to be generalised across courses, lecturers, subjects, levels, institutions and cultures. However, this illustrative data is from a postgraduate teacher education course called 'Psychology for teachers' that was taken by 36 mainly mature students.

Figure 2 is in three parts, Part 1 'For the Lecturer', Part 2 'For the students' and Part 3, the 'Summary Results'. In Part 1, the 'Start' row shows the first seven numbers input from the Lecturer's alignment form displayed in Figure 1. The eighth numbers were not used for this analysis. Part 2 'For the Students' shows just a selection of four rows #12, #25, #35 and #30, from the alignment forms of all the students in the course. As well as the first seven numbers from the students' alignment forms, these rows have been extended to show other variables of interest. The last column for the students shows their academic results. These are entered after the course and are used to validate the predictions from the in-course alignment indicators and to further validate the alignment theory. The 'Summary Results' section, Part 3, shows the calculated changes resulting from the lecturer's start position. It is seen, for example, that the lecturer's starting position for Skills of 30 for 'is now' and 60 for 'should be' requires a 100% change, i.e. $(60 - 30)/30 = 100\%$

Although not shown, the student changes were calculated in the same way and each compared with the lecturer's changes. This was done for changes in both Scope and Proportion. Scope is calculated for each student as the total of the student/lecturer absolute differences in the raw numbers given for the three objectives. Proportion alignment for each student is calculated in a similar way but using proportional changes calculated from the six numbers. The Alignment indicators shown in the results section of Figure 2 are the average of students' scope alignments. From this start position the average scope Alignment for all students was 1.499. Students' scope alignments were correlated with their academic results at -0.265 Also their Proportions alignments were correlated with their reported enjoyment levels at -0.22. This means that the more aligned students and their lecturer were on Scope then the higher were the students' academic results. Also, the more aligned students and lecturer were on Proportion then the more the students enjoyed the course. These results agree with the theory.

We can do a similar analysis for any subgroup of interest. Figure 2 also shows this analysis for two separate option groups in Part 3 - for Option 2 'History' students and for Option 4 'Modern language' students. The Scope alignment and Proportion alignment for these two sub-groups is given along with their respective mean enjoyment scores and mean academic results. It will be noticed that the Modern Language students were more aligned in both Scope and Proportion than were the History students. Correspondingly, we find that the mean academic results and enjoyment of the Modern Language students were higher than those of the History students. Although the sizes of the sub-groups were small, these comparative sub-group results are also in agreement with the alignment theory.

Calculating changes that optimize teaching and learning

We can use a simple linear programming algorithm to find those lecturer changes that would result in the best/minimum alignment. There are many options for this calculation that depend on (i) the indicator of most interest, (ii) the malleability of student expectations and (iii) the lecturer's freedom to adjust course expectations. The first example in Figure 2 minimises the mean Scope alignment for all students of 1.499 down to 0.674 and calculates the 'Best' changes that would give this minimised Scope alignment. The resulting 'Best' changes are shown in Part 3 of Figure 2 as 0% for Skills, 10% for Understanding and 19% for Attitudes. Given the lecturer's initial evaluation of the whole class as 30, 40 and 60, these optimum changes imply that the lecturer should have been working towards 30 (no change), 44 and 71. If the lecturer had done so then the students' alignment scores would have changed and their correlations of academic attainment with the new Scope alignments and Proportion alignments would have improved to -0.307 and -0.594 respectively. This sensitivity analysis again demonstrates (i) the predictive validity of the alignment indicators, in that Scope alignment and Proportion alignment correlate with attainment and enjoyment, and (ii) the validity of the theory that the better the scope alignment then the higher is the academic attainment and the better the proportion alignment then the more the students enjoy the course.

A similar sensitivity analysis was done for the History and Modern Language option groups by minimising their alignments to find the best changes for these two sub-groups. Figure 2 shows that the best alignments for these two groups are 1.221 and 0.223 respectively. The mean proportion alignments for these two option groups would then also have improved to 0.299 and 0.065 respectively, validating the theory for these student-subgroups.

Traditional SETs are a ‘post mortem’ assessment, collected at the end of the course when it is too late to use this feedback to help the students who made the assessments. However, a lecturer does not have to wait until the course is over to optimize teaching using the alignment method. The data collected in-course can be processed by the same type of sensitivity analysis to calculate the optimum changes that should be made by the lecturer during the course to maximize the students’ post-course academic attainment and/or course enjoyment.

Administrative decision point assessment of quality teaching

The lecturer may utilize the five-minute alignment form many times during the course to keep his/her teaching on track. The administration uses it just once near the end of the course to calculate the final alignment score for that lecturer’s quality of teaching. This results in a single decision point number that can be compared across the institution and used for promotion and tenure decisions.

It will be noticed from Figure 2 that the minimum alignment that is possible for this group is 0.674. Remember that the best alignment is the one closest to the perfect score of zero. The minimum possible alignment for the History students was much higher at 1.221 than for the Modern Language students at 0.223. These minimum possible alignment scores illustrate the best teaching/learning that is possible with these sub-groups of students and reflects the fact that students are not all equally amenable to required educational changes in Skills, Understanding and Attitudes. To give the lecturer some protection from such intransigence the decision point measure of quality teaching is taken as the actual alignment less this minimum/best possible alignment. In the example, given the alignment score for the whole class is 1.499 and the best possible alignment for the whole class is 0.674, the quality teaching score for the whole class is $1.499 - 0.674 = 0.825$. These results from Figure 2 are summarised in Table 1.

Table 1: Quality Teaching (QT) scores for whole class and for student sub-groups

	Alignment scores			Academic Attainment
	Mean Scope	Best Scope	QT Score	
Whole Class	1.499	0.674	0.825	63.0%
Sub-Groups				
Option 2	1.861	1.221	0.640 high	52.2%
Option 4	1.200	0.223	0.977 low	72.5%

Table 1 illustrates a novel research application of the alignment method. It measures, for the first time, the differential effort that teachers expend in teaching mixed ability students. Educators generally accept that teachers expend more effort teaching ‘less-able’ students than in teaching the ‘more-able’ students in their classes. It is thought that the ‘more-able’ students, being more independent, can manage more on their own. For example, as students improve they become less dependent on teacher assistance (Davis, 1998) and higher ability students use their time more productively while waiting for teacher assistance (de la Cruz, 1995).

It can be noticed from Table 1 that the quality of teaching was skewed more towards the needs of the more intransigent group. The quality of teaching experienced by the History Students (option 4, QT=0.640) was 0.337 better than that experienced by the Modern Language students (option 2, QT=0.977). Table 1 shows that this greater teaching quality produced a mean academic attainment of only 52.2% compared to the mean academic attainment of 72.5% for the Modern Language students who experienced a 53% lower quality of teaching.

Discussion

This article has only touched on the classroom assessment use of the Alignment Method. It has not discussed the staff and course development aspects of the method or the many benefits the method is designed to offer for Quality Assurance compared to traditional SETs.

This alignment method of assessing teaching quality is designed to offer 10 main benefits:

- 1 It identifies the quality of teaching experienced by each individual student.
- 2 It can be used to identify groups of students that might be disadvantaged by the teaching.
- 3 It offers detailed diagnostic reports to help the lecturer.
- 4 It only takes 5 minutes to administer and the analysis is quick and low-cost.
- 5 It can be given several times in-course resulting in optimum recommendations to keep teaching on track.
- 6 It is sensitive to criteria considered important in different subject areas and by different Faculties and levels of students.
- 7 It maintains lecturer/student trust and promotes higher quality teaching and higher quality learning.
- 8 It protects academic freedom, is non-threatening and has built in protection for Faculty who teach intransigent students and difficult courses.
- 9 It uses one standard form and gives one single decision-point number that can be used in institutional evaluations for comparing quality of teaching across the university e.g. for Quality Audits, teaching awards and for promotion and tenure decisions.
- 10 Post-course correlations with academic standards evidence the reliability and validity of the instrument for each course and for subgroups of students taking each course on which it is used.

This alignment method can be flexibly piloted at different levels within an institution - at the level of full institutional evaluation, at the level of staff and course development within Faculties, Schools or Departments and at the level of individual lecturers who are interested in improving the quality of their own teaching for their own students. The method is being continually up-graded and Web-based Alignment software is now being developed that will enable lecturers and administrators from tertiary institutions worldwide to avail themselves of the benefits of using the alignment method in their own institutions.

References

- Altschuler, G. (1999), Let me edutain you, *The New York Times*, Education Life Supplement, April 4.
- Biggs, J. B. (1999). *Teaching for Quality Learning at University: What the Student Does*. Buckingham, England: Open University Press.
- Bloom, . S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York: McKay.
- Breitsprecher, C. H. (1991, May). *Relative Effects of Verbal Mediation, Feedback Monitoring, and Alignment on Community College Learners' Achievement*. Unpublished Ed.D. Dissertation, University of San Francisco.
- Cohen, S. A. (1987). Instructional Alignment: Searching for a Magic Bullet. *Educational Researcher* 16(8), 16-20.
- Cohen, S. A. (1991). New Alignment Experiments: Using Outcome-Based Instruction to Teach Transfer of Learning. *Outcomes: The Quarterly Journal of the Network of Outcome-Based Schools*, 10(3), 11-16.
- Cohen, S. A. (1994). *Instructional Alignment*. In International Encyclopaedia of Education: Research and Studies, T. Husen & T. N. Postlethwaite (Eds.). Volume 5, Second Edition. London: Pergamon Press, 2852-2856.
- Cohen, S. A., Hyman, J. S., Ashcroft, L., & Loveless, D. (1989). *Comparing Effects of Meta-cognition, Learning Styles, and Human Attributes with Alignment*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Crumbly, D.L. (1995), Dysfunctional effects of summative student evaluations of teaching: Games professors play, *Accounting Perspectives*, 1(1), 67-77.
- Cruz, R. E. de la (1995, April). *Teacher Perspectives of the Social Skills Development of Children with Learning Disabilities*. Paper presented at the 73rd Annual International Convention of the Council for Exceptional Children. Indianapolis, IN, USA.
- Damron, J.C. (1995). *The three faces of teaching evaluation*. Unpublished manuscript, Douglas College, New Westminster, British Columbia.
- Davis, A. P. (1998). Performance Achievement and Analysis of Teaching during Choral Rehearsals. *Journal of Research in Music Education*, 46(4), 496-509.
- Elia, J. I. (1994). An Alignment/Transfer Experiment with Low Socioeconomic Level Students. *Teacher Education Quarterly*, 21(3), 113-24.
- Gillmore, G. M. & Greenwald, A. G. (1999). Using statistical adjustment to reduce biases in student ratings. *American-Psychologist*, 54(7): 518-519
- Hativa, N., Barak, R., & Simhi, E. (1999, April). *Expert University Teachers: Thinking, Knowledge and Practice Regarding Effective Teaching Behaviors*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Krautmann, A. C. & Sander, W. (1999). Grades and Student Evaluations of Teachers. *Economics of Education Review*, 18(1), 59-63.
- Murray, H. G. (1984). The Impact of Formative and Summative Evaluation of Teaching in North American Universities. *Assessment and Evaluation in Higher Education*, 9 (2), 117-132.
- Redding, S. (1992). Common Experience. *School Community Journal*, 2 (2), 43-51.
- Ruskai, M.B. (1997), Evaluating student evaluations. *Notices of The American Mathematical Society* 44(3), 308.
- Vergason, G. A. & Anderegg, M. L. (1991). Beyond the Regular Education Initiative and the Resource Room Controversy. *Focus on Exceptional Children*, 23(7), 1-7.
- Walker, M. H. (1998). The Fundamentals of a Quality Education. *Streamlined Seminar*; 16(3), 1-4. ■
-