

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**USING STUDENT EVALUATIONS FOR  
TEACHER DEVELOPMENT IN NINE URBAN  
SECONDARY SCHOOLS**

STEPHEN EARL IRVING

1996



## Abstract

---

Student evaluations are becoming more common in education, particularly in the tertiary sector. They have been used by students and by the authorities to make judgments about teachers and courses. Their use for helping teachers to improve their teaching is a recent phenomenon which has barely touched secondary schooling, let alone primary schools.

Teacher evaluation data collected for summative purposes have had little effect on teaching performance as there is usually little or no feedback designed to help the teacher to improve. Teachers need to know not only what to change but also how to change their behaviours in a desired direction. This research set out to develop a questionnaire and methodology that could be used by secondary teachers to evaluate their teaching using the students as the source of information, and then use that information to help the teacher to improve their teaching. The methodology draws heavily on the work of Wilson (1986), and Marsh and Roche (1993, 1994) who constructed a process that supplemented feedback with a collegial consultation to help the teacher interpret the data in a meaningful way and then act on it. This methodology has been shown to be the best practice in this field, even though the results of carefully researched studies are modest.

A questionnaire appropriate to the New Zealand secondary school environment was constructed and administered in nine urban secondary schools to 344 students. The subject teachers they evaluated were from a wide cross-section of curriculum areas. Most were experienced teachers. At the same time, the teachers completed a self evaluation using the same questionnaire. The teachers received the results of the evaluation with notes on how to interpret the tables and graphs. This was followed by a consultation with the researcher, using a methodology developed from appraisal interviewing techniques.

An action plan was devised during this consultation. The teacher then put this into action, and the students were re-surveyed after approximately thirteen weeks. The results of the two surveys were compared to see whether this process was beneficial in improving teaching, as perceived by the students.

Overall, the results showed a rather modest improvement across the board. There was a noticeable difference between two groups of teachers dependent on the difference between their own self-evaluation and the average student response. Teachers whose self-evaluation was similar to the student evaluation, or whose self-evaluation was worse than the student evaluations changed little between the two administrations of the questionnaire. On the other hand, if the self-evaluation was better than the average student evaluation, then there were significant improvements in the student evaluations on the second administration. This finding is in keeping with the theory of cognitive dissonance, first espoused by Festinger (1957). When the teacher has a positive self-evaluation but the students rate that teacher poorly, then the teacher is motivated to change their teaching behaviours so that the next student evaluation is favourable.

As part of a teacher development programme, students and teachers similarly felt that this form of evaluation is valuable and has a place in appraisal schemes designed to help teachers improve their teaching. There is still considerable reluctance on the part of teachers for this type of evaluation tool to be used for the purposes of promotion, tenure and reward. In light of the requirements for schools to implement performance appraisal schemes, and the need for appraisals to be based on "objective" data, student evaluations can provide the desired information.

## Acknowledgments

---

This research has only been possible with the assistance of a very large cast who have given of their time, effort and energy.

The participants are the people without whom this research could never have taken place. I would like to express my thanks to them all for responding to the calls that I have made on them, often at what seems the most inconvenient time in the annual cycle of school life. The school principals have been very generous in allowing me access to their students, teachers and classrooms. The key people were the teachers. They have been the willing guinea pigs in this project, and were extremely receptive to my work with them. I believe that I have learned more from them than they could ever have gained from me. The students were crucial. The students in the nine schools were wonderful, and working with them on this research confirmed my optimism about our young people. Secondary school students are not yet used to being asked for their opinion about their teachers, and all of the student participants undertook their role with considerable thought and interest in (hopefully) promoting a better learning environment. Without their thoughtful input, this research would not have proceeded.

My supervisors have been invaluable in encouraging this novice through the preparation and production of this research. Dr Mollie Neville has always been available to assist me and offer invaluable advice, and Associate Professor Wayne Edwards has been a tower of support, inspiration and encouragement. I would have been hopelessly lost on the statistical analysis without the willing assistance of Dr Denny Meyer. Brian Steel of Brian Steel Research has been amazingly willing to assist me with the intricacies of data processing using SPSS, and has often been available at short notice to remove bugs from the setup.

I was indeed fortunate to receive a funded study award from the New Zealand Schools Trustees Association. This made it possible for me to have the time to work on the research without having to test the goodwill of my colleagues at Onehunga High School who have been really supportive throughout the year. I certainly appreciated the opportunity this award gave me to spend time on research, quiet reflection, visiting schools and classes, writing or just taking a break.

Finally, I would like to acknowledge my wife and daughters for their support and encouragement. They have been tolerant through all the jubilations and trials that this research has generated, allowing me to be selfish in completing this project.

My thanks to all of you for making this project possible.

# Contents

---

Abstract	iii
Acknowledgments	v
Contents	vi
List of Tables	ix
List of Appendices	xi
Chapter One: Introduction	1
Chapter Two: Literature Review	5
Part One: Student Evaluation of Teacher Performance	6
Historical Context	
Arguments For and Against Student Evaluations	
Validity	
Do student evaluations accurately reflect student opinions?	
Do student evaluations accurately reflect instructional	
effectiveness?	
Multi-section studies	
Meta-analysis	
Generalisability	
Educational seduction	
The dimensionality of teaching	
Reliability	
Stability	
Utility of Student Evaluations	
Part Two: Using Student Evaluation Feedback for Teacher Development	20
Simple feedback research	
Augmented feedback studies	
Feedback with consultation	
Summary	30

Chapter Three: Methodology	31
Procedural Overview	
Human Ethics Considerations	
The Participants	
The Schools	
The Teachers	
The Classes	
The Students	
The Questionnaire	
The Consultation	
The Follow-Up Questionnaires	
Time	
Summary	56
Chapter Four: Results	58
Part One: Quantitative Results	58
Student Evaluations: Descriptive Information	
Factor Analysis	
Did the Teacher Improve?	
Size Effect, d	
Student t-Tests on Items	
Student t-Tests on Factors	
Teacher Self Evaluations	
Student Reactions	
Time Considerations	
Reliability and Validity	
Part Two: Qualitative Results	80
Time One Student Comments	
Time One Teacher Comments	
Time Two Student Comments	
The Evaluation Process	
The Questionnaire	
Classwork	
The Teacher	
The Students	

Teacher Post Analysis.	
The Questionnaire	
Written Feedback	
The Consultation	
Teacher Development	
The Teacher Reactions	
Project Evaluation	
Teacher Appraisal	
Summary	100
Chapter Five: Discussion	102
The Research Design and Student Questionnaire	
The Consultation Conference	
Teacher Appraisal and Evaluation	
Teacher and Student Reactions	
Implications for Future Research	
Summary	113
Bibliography	115
Appendices	138

# List of Tables

---

2.1	Correlation between Teacher/Course and the Factors of Skill, Rapport, Structure and Difficulty.	18
3.1	Categories of All Schools and the Selected Schools in Region.	38
3.2	Roll Numbers, Student Characteristics and Number of Teaching Staff in Selected Schools.	40
3.3	Selected Schools by Type.	41
3.4	Summary of Classes Selected by Form Level and Subject.	46
3.5	Number of Students Completing the First, Second or Both Administrations of the Student Questionnaire.	47
3.6	Classes, showing the Number in the Class, Number who Consented, and Number who Completed the Two Administrations of the Questionnaire.	48
4.1	Rank Order for Response Categories 1 and 2, separately and combined, Time One and Time Two combined.	59
4.2	Rank Order for Response Categories 4 and 5, separately and combined, Time One and Time Two combined.	60
4.3	Rank Order of Response Categories 3 and 6 for Time One and Time Two combined.	60
4.4	Rank Order of Mean and Standard Deviation for Time One and Time Two combined.	61
4.5	Number of Beneficial and Non-Beneficial Changes in Item Scores between Time One and Time Two for All Items	67
4.6	Number of Beneficial and Non-Beneficial Changes in Size Effect d in Item Scores between Time One and Time Two for Targeted Items.	68
4.7	Size Effect, d, on the Overall Rating Items for Individual Teachers/Classes.	69
4.8	Student t-test for the Differences on the Means for the Overall Rating Items for Individual Teachers/Classes.	70
4.9	Student t-Test Scores for Each Factor Score for Individual Teachers.	70

4.10	Summary of One-Way Analysis of Variance on Factors as a Function of Ability, Class, Ethnicity, and Gender.	71
4.11	Summary of One-Way Analysis of Variance on Factors as a Function of Best Class, Best Teacher and No Home Help.	73
4.12	Number of Positive Differences between Time One Student and Teacher Ratings, Number of Significant and Highly Significant Changes between Time One and Time Two.	74
4.13	Spearman Correlation Coefficients between Mean Teacher/Class Average Rating Difference and the Total Number of Significant Changes and the Number of Highly Significant Changes.	75
4.14	Time Two Student Responses to Project Evaluation.	76
4.15	Time Taken to Complete Time One and Time Two Questionnaires as a Function of First Student Finished, First Ten Students Finished, and All Students Finished.	77
4.16	Pearson and Spearman Correlations for Each Response Category between Time One and Time Two.	78
4.17	Correlation for Combined Response Categories between Time One and Time Two.	79
4.18	Correlation for Opposite Response Categories between Time One and Time Two.	79
4.19	Correlation for Combined Opposite Response Categories between Time One and Time Two.	79

## List of Appendices

---

3.1	Sample Printout of Time One Results as presented to Teachers, with Instructions on Interpretation.	138
3.2	Item Pool for Questionnaire.	148
3.3	Student Questionnaire for Time One.	160
3.4	Student Questionnaire for Time Two.	164
3.5	Framework for Consultation Conference.	165
3.6	Teacher Questionnaire for Time Two.	166
4.1	<i>A Priori</i> Factors: Wording of Items.	178
4.2	Rotated Factor Matrix.	180
4.3	<i>Post Facto</i> Factors: Wording of Items.	181
4.4	Printout of Time Two Results for Each Teacher showing Size Effect, d.	183
4.5	Summary of Student t-Tests for Difference between Time One and Time Two Mean Rating on Individual Items for Each Teacher.	215
4.6	Coding of Time One Student Comments according to <i>A Priori</i> Factors.	219
4.7	Coding of Time One Student Comments according to <i>Post Facto</i> Factors.	220
4.8	Coding of Time One Teacher Comments according to <i>A Priori</i> Factors.	221
4.9	Coding of Time One Teacher Comments according to <i>Post Facto</i> Factors.	222
4.10	Results of Teacher Post Evaluation Questionnaire.	223