

1987

An Empirical Classification Of University Lecturing Styles

Denise Frances Mclean

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Mclean, Denise Frances, "An Empirical Classification Of University Lecturing Styles" (1987). *Digitized Theses*. 1601.
<https://ir.lib.uwo.ca/digitizedtheses/1601>

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

An Empirical Classification of
University Lecturing Styles

by

Denise Frances McLean

Department of Psychology

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario

London, Ontario

April, 1987

© Denise McLean, 1987

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-36067-4

ABSTRACT

The goal of this research was to develop and validate an empirical classification of naturally occurring teaching styles in university undergraduate lecture courses. Teaching style was described in terms of patterns of observable classroom teaching behaviour.

In the first study, the classification was developed with reference to an initial sample of 298 university lecturers. Each instructor was observed by student raters who judged the frequency of occurrence of 100 low-inference behaviours. Student ratings were averaged within classes and eight scale scores were computed. A cluster analysis identified six groups of instructors, with members of each group homogeneous in terms of their classroom teaching behaviour patterns. The replicability of the clusters was assessed within this sample through split-half samples, statistical tests, and a comparison of results based on item and scale scores.

In the second study, the typology was validated by investigating relationships between teaching style and other relevant variables, such as instructor characteristics and attitudes, course characteristics, and global student ratings of teaching effectiveness. For this purpose, a second sample of 135 classes was obtained. Significant differences in the frequencies of the six styles across

subject area and course level had been noted in the original sample. These differences in style frequencies were replicated in the second sample with regard to course level but not subject area. Overall elevation of the behavioural scale scores played a major role in determining relationships between teaching style and global student ratings. Instructor characteristics such as years of teaching experience, academic rank, and educational goals and attitudes did not display significant relationships with teaching style.

In the third study, additional validation evidence for the typology was gathered through examination of the stability of teaching styles across courses and across years. It was found that behaviour was relatively stable when courses were at the same academic level (i.e. introductory, general, honours), even if the courses were separated by as much as eight years.

In general there was evidence that the empirical classification approach is a useful method for describing teacher behaviour. It allows a definition of teaching style, or educational treatment, in terms of discrete, observable behaviours.

ACKNOWLEDGEMENTS

I wish to acknowledge the contributions made by several individuals. I am particularly appreciative of the support and advice offered to me by my supervisor, Dr. Harry C. Murray, throughout this endeavour. In addition, I thank Harry for allowing me to include much of his early data in this thesis. I also thank Dr. Richard Harshman for our many discussions on the reality of clusters and Dr. Tony Vernon for his helpful suggestions and comments. The observations made by members of my examining committee, Dr. Jim Neufeld, Dr. Nancy Vye, and Dr. Al Slemon, were helpful to me in viewing my research from slightly different perspectives. Many of my colleagues were also very helpful to me at various stages, and I thank P. Duschner, C. Fekken, T. Hill, S. Paunonen, and R. Stockwell for their varied contributions. Finally, I extend my gratitude to my family and friends for their understanding and encouragement.

TABLE OF CONTENTS

	Page
CERTIFICATE OF EXAMINATION.....	ii
ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	ix
LIST OF FIGURES.....	xi
LIST OF APPENDICES.....	xii
INTRODUCTION.....	1
Need for a Description of Lecturing Behaviour.....	2
Previous Classification of Teaching Styles.....	8
Informal/Observational Classifications.....	9
Formal/Theoretical Classifications.....	12
Empirical Classifications.....	14
Development of a Teaching Style Classification.....	19
Description of Teaching Behaviour.....	20
Identification of Teaching Styles.....	22
Validation of a Teaching Style Classification.....	24
STUDY 1.....	26
Method.....	26
Subjects.....	26
Measures.....	28
Cluster Analysis.....	32
Results.....	33

Psychometric Properties of Measures.....	33
Classification Results.....	34
Item-Scale Cluster Comparison.....	39
Cluster Replication Across Split-Half Samples.....	41
Wishart's Test for Hierarchical Clustering.....	44
Effect of Outliers.....	45
Effect of Reordering Cases.....	46
Relationship to Subject Area and Level of Course.....	47
Summary.....	50
STUDY 2.....	54
Method.....	58
Subjects.....	58
Measures.....	58
Teacher Behaviors Inventory.....	58
Student Rating Form I.....	58
Instructor Self-Rating Questionnaire.....	60
Procedure.....	65
Results.....	67
Psychometric Properties of Measures.....	67
Teacher Behaviors Inventory.....	67
Student Rating Form I.....	67
Instructor Self-Rating Questionnaire.....	69
Classification of Teaching Styles.....	77
Relationship of Cluster Membership to Other Variables.....	77
Instructor Demographic Characteristics.....	80
Class Characteristics.....	82

Instructor Attitudes.....	85
Instructor Roles.....	85
Instructor Self-Ratings.....	86
Global Student Ratings.....	87
Informational Value of Clusters versus Scales.....	92
Summary.....	94
STUDY 3.....	94
Method.....	100
Subjects.....	100
Sample A.....	100
Sample B.....	100
Measures and Data Collection.....	100
Results.....	101
Stability of Cluster Membership.....	101
Stability of Profiles.....	106
Stability of Individual Differences in Scores.....	108
Summary.....	110
DISCUSSION.....	114
Utility of Method.....	114
The Six Teaching Styles.....	115
Stability of Teaching Style.....	119
Relationships between Teaching Style and Other Variables.....	121
Future Research.....	125
REFERENCES.....	128
APPENDICES.....	136
VITA.....	155

LIST OF TABLES

Table	Description	Page
1	Definitions and Sample Items of TBI Scales.....	30
2	Mean, Standard Deviation, Internal Consistency, and Intercorrelations of TBI Scales (Study 1)....	35
3	Comparison of Cluster Memberships Using Scale and Item Data.....	40
4	Observed Intercluster Distances in Split-Samples.	43
5	Relationship between Teaching Style and Course Level (Study 1).....	48
6	Relationship between Teaching Style and Subject Area.....	51
7	Demographic Characteristics of Instructors and Courses (Study 2).....	59
8	Definitions of Teaching Styles Inventory Scores..	64
9	Mean, Standard Deviation, Internal Consistency, and Intercorrelations of TBI Scales (Study 2)....	68
10	Rotated Principal Components of Global Student Rating Form.....	70
11	Descriptive Statistics for Instructor Questionnaire.....	72
12	Descriptive Statistics for Teacher Attitude Variables on Instructor Questionnaire.....	73
13	Rotated Principal Components of Teacher Attitude Variables.....	76
14	Relationship between Teaching Style and Course Level (Study 2).....	83
15	Style, Level, and Faculty Effects on Global Student Rating Factors.....	89
16	Replication of Teaching Styles Across Classes Taught in Same Year.....	102
17	Similarity in Teaching Style Compared to Similarity in Course Level.....	104
18	Correlation Coefficients Assessing Stability of Profile Distance to Cluster Centroids.....	107

19	Mean Values of Profile Similarity Measures.....	109
20	Correlation Coefficients Assessing Stability of TBI Scale Scores.....	111

LIST OF FIGURES

Figure	Description	Page
1	Cluster Profiles (Study 1 Sample).....	36
2	Cluster Profiles (Study 2 Sample).....	78

LIST OF APPENDICES

Appendix	Description	Page
A	Teacher Behaviors Inventory.....	136
B	Global Student Rating Form I.....	141
C	Instructor Self-Rating Questionnaire.....	143

The author of this thesis has granted The University of Western Ontario a non-exclusive license to reproduce and distribute copies of this thesis to users of Western Libraries. Copyright remains with the author.

Electronic theses and dissertations available in The University of Western Ontario's institutional repository (Scholarship@Western) are solely for the purpose of private study and research. They may not be copied or reproduced, except as permitted by copyright laws, without written authority of the copyright owner. Any commercial use or publication is strictly prohibited.

The original copyright license attesting to these terms and signed by the author of this thesis may be found in the original print version of the thesis, held by Western Libraries.

The thesis approval page signed by the examining committee may also be found in the original print version of the thesis held in Western Libraries.

Please contact Western Libraries for further information:

E-mail: libadmin@uwo.ca

Telephone: (519) 661-2111 Ext. 84796

Web site: <http://www.lib.uwo.ca/>

INTRODUCTION

The educational process is among the most complex human interactions studied by psychologists. As such, the need in this area for parsimonious yet operationally meaningful descriptions of behaviour is of tantamount importance. For many of the agents in the educational process, this need for description is on the verge of being met: characteristics of effective schools can be assessed (Good & Weinstein, 1986), textual passage difficulty can be estimated (Klare, 1974-1975), the knowledge structure of course material can be organized into hierarchies (Gagne, 1977) or networks (Donald, 1983), and relevant student aptitudes, attitudes, and achievements can all be measured with a high degree of accuracy. The one component upon which measurement attention has not yet been effectively focused is the medium through which students receive most of their formal education, the teacher.

The program of research described here addressed this remaining need. The three studies outlined below comprise the development and initial validation of an empirical classification of classroom teaching styles in lecture courses at the postsecondary level.

Need for a Description of Lecturing Behaviour

The need for a valid multidimensional description of teacher behaviour has become increasingly apparent in educational research over the past fifteen years. The recent emphasis on aptitude-treatment interaction (ATI) research (Cronbach & Snow, 1977) has only served to emphasize this lack. In a paradigm which predicts differential outcomes as a function of an interaction between student characteristics and educational treatments, it is imperative that we have effective ways of measuring, or describing, all of these components. Advances in item response theory (e.g., Hambleton, 1983) and criterion-referenced testing (Popham, 1978) have yielded great gains in the measurement of student ability and achievement. In contrast, the description of educational treatment, of which teacher behaviour is one major aspect, is as yet quite undeveloped. It has typically been limited to dichotomies such as lecture vs. discussion method, or student-centered vs. teacher-centered approaches, distinctions which are often insensitive to other relevant variables or to gradations within each category. This difficulty has often been cited as a possible reason for the disappointing record of ATI research (e.g., Tobias, 1979). Any differential effectiveness of various educational approaches for various groups of students will not be uncovered or replicated if we cannot adequately understand and describe the educational treatments being applied.

Increasingly, teacher behaviour is being recognized as a crucially important component of educational treatments, and as something that must be considered in educational research. At the elementary school level Brophy (1986) observed that the research focus has shifted over the past fifteen years from teacher traits to teacher behaviours. He summarized this research and concluded that a consistent body of knowledge has been developed, linking particular teacher behaviours to student achievement. In contrast similar research at the postsecondary level, though abundant in numbers, still remains very much a collection of isolated results. For example, although close to 800 studies were published between 1965 and 1980 on student ratings of teacher effectiveness (Murray, 1980), and although these ratings have been found to be statistically related to specific teaching behaviours (e.g., Murray, 1983a), student ratings of global instructor characteristics must be viewed as no more than remote approximations to a proper description of teacher behaviour.

Current attempts at describing teacher behaviour suffer from several disadvantages. For example, most of the research on the effects of teacher behaviour has failed to examine the multidimensional nature of teaching style. Instead, researchers have focused on unidimensional, single-behaviour constructs such as teacher questioning (Andrews, 1980; Gall et al., 1978; Merlino, 1977) or have collapsed a number of constructs into a "syndrome" of behaviours which

4

is then treated as a single dimension, such as student-centered vs. teacher-centered instruction (cf. Costin 1971). Although these two approaches can simplify experimental design and data analysis considerably important issues are often overlooked.

First, the teaching-learning act is, among other things, a complex social interaction between teacher and students. The social psychology literature tells us that the interpretation of another's behaviour is strongly affected by the context, both situational and behavioural in which a given act is observed. For example, it seems reasonable to postulate that high-level teacher questions (i.e., questions that call for interpretation or application rather than only recall of basic knowledge) may seem to the students to be indicative of genuine teacher interest when the questions are asked by a teacher with good student rapport. Alternatively, the same questions might be interpreted as threatening or intrusive in a low rapport situation. As a result one might well expect different outcomes resulting from the same behaviour in the two situations (e.g., Andrews, 1980; Hegarty, 1978). Consistent with this concern, Brophy, (1986) has recommended a shift in emphasis away from the counting of single molecular behaviours and toward an investigation of more molar or contextual instructional patterns.

Second, collapsing several variables into one composite to represent a syndrome or pattern of behaviour requires the

assumption that the effects of the individual behaviours are always additive. For example, this approach assumes that an instructor with a score of 10 on student-centeredness by virtue of scoring 5 on rapport and 5 on availability outside of class, is the same in students' eyes as an instructor who scores 10 on rapport and 0 on availability. Lacking any empirical evidence to justify the use of an additive index to measure teaching style, a better strategy would be to take a multivariate approach, which allows the examination of patterns of behaviour without losing individual trait data.

Third, regardless of whether a unidimensional or a multidimensional approach is taken to describe teacher behaviour, a more general concern is the lack of empirical data to support the teaching styles investigated. As elaborated below, types of teachers are often listed in the literature, along with the defining characteristics of each type, but no evidence is provided in support of either the hypothesized trait interrelationships in each type or the prevalence of each type among teachers in general.

A final problem with current descriptions of teacher classroom behaviour lies in the measures used. At the elementary and secondary-school levels, the tendency has been to collect detailed observational data using a time-sampling scheme (e.g., Flanders, 1970). Although the resultant data are usually of high quality, the complexity of the observational systems must be seen as a contributing

factor in the overemphasis on single behaviours. In contrast, teacher descriptions at the postsecondary level are frequently based upon student ratings of global instructor characteristics such as clarity, organization, or enthusiasm. These data have the disadvantage of requiring a substantial amount of inference on the part of student observers.

In summary, unidimensional approaches to teacher behaviour have failed to provide a complete understanding of the complexities of the teaching process, particularly at the postsecondary level. Multidimensional approaches, on the other hand, have in some cases been limited to a series of unidimensional analyses of traits, with no investigation of interactions or context effects; and in other cases have failed to include empirical data of any sort.

Because the logical first step in studying any phenomenon is finding a method of describing and measuring it, the present series of research studies was aimed at developing a description of lecturing behaviour at the postsecondary level.

Teaching was described through ratings from students, who must be viewed as the group of raters in the best position to judge the instructor's in-class behaviour. However, rather than global characteristics, students were asked to rate the frequency-of-occurrence of observable, low-inference classroom teaching behaviours, such as "exhibits facial gestures or expressions" and "criticizes

students when they make errors". Results reported by Murray (1980) indicate that student ratings of such behaviours agree well with similar ratings made by impartial trained observers. It was assumed that such ratings would provide an accurate, detailed description of behaviour, from which a more global, multidimensional description can be empirically derived.

The first goal of the research was to develop a typology of teaching styles. Specifically, Study 1 was conducted to derive empirically a set of naturally occurring teaching styles within university lecture courses. A teaching style is identified when a similar pattern of classroom teaching behaviours is exhibited by a number of different instructors. Dunkin (1983, 1986) has recently called for an increased emphasis on field studies of this sort in the investigation of lecturing behaviours. An understanding of the patterns of behaviour which instructors naturally exhibit is the first step toward understanding the causes and effects of those behaviours. These patterns can be identified by applying such data reduction techniques as cluster analysis to the student observations of a large number of instructors.

The second goal of this research was to validate the descriptive typology of teaching styles against criteria either internal or external to the classification. For example, the internal criterion of mathematical stability of the empirically derived teaching styles was evaluated in

Study 1. External criteria included the relationships between these styles and other relevant variables, investigated in Studies 1 and 2, and the stability of an instructor's style across time and across courses, examined in Study 3.

Previous Classification of Teaching Styles

There are a number of approaches to classification in general, and to classification of teacher behaviour in particular. At one level, probably all human beings engage in classification of the objects in their environment. Such informal grouping enables us to react to a wide variety of stimuli in an appropriate and efficient manner. However, classification of classroom teaching styles on the basis of informal observation and/or rational description is especially difficult. Teaching is in general a solitary profession with each instructor working in isolation in a classroom and seldom, if ever, observing the work of his or her colleagues (Sarason, 1982). Particularly at the postsecondary level, few individuals have the benefit of either formal training in teaching methods or systematic observation of other instructors' classroom behaviours. For this reason, the literature on university teaching styles is replete with "armchair theorizing", which has little or no observational basis. As pointed out by Dunkin (1983), "concepts of teaching styles and roles are much more to be found in popular parlance than in research on teaching in

higher education" (p. 74).

Informal/Observational Classifications. Several authors have developed teaching style typologies based on anecdotal observation and/or rational considerations. The five typologies presented below are examples of classification systems based upon what seem to be purely informal methods, in that no theoretical model is offered for the derivation of the "styles" and no supporting empirical data are presented.

Axelrod (1973) identified two general teaching styles based upon interviews with over 100 faculty members in the humanities. The didactic teaching style, used primarily by individuals labelled "teacher-craftsmen", is completely teacher-centered and does not require or encourage inquiry on the part of students. The purpose of the didactic teaching style is to impart very specific knowledge or skills, as in language training. The evocative teaching style, used primarily by "teacher-artists", does encourage student inquiry and is the style advocated strongly by Axelrod. Within the evocative teaching style, Axelrod describes four instructor prototypes: principles-and-facts, instructor-centered, student-as-mind, and student-as-person. The first two prototypes are subject-centered and the last two are student-centered. The principles-and-facts prototype is the most common of the four models and embodies an emphasis on cognitive knowledge. The instructor-centered prototype is the second most common model and also includes

an emphasis on the transmission of knowledge. It is distinct from the principles-and-facts model in its focus on the viewpoints of the instructor, who acts as role model and in its emphasis on the presentation of a diversity of approaches and controversies as opposed to universally-accepted facts. The student-as-mind prototype focuses on the intellectual development of the student. The student-as-person prototype emphasizes the personal development of the whole student, not just his or her mind. Axelrod's description of the four prototypes focuses on instructor goals and values, rather than on instructor behaviours. He emphasizes that within each of the styles, there are very effective teachers and very ineffective teachers. Axelrod states that there is no strong relationship between teaching styles and type of course. Subject-centered styles are not limited to lecture courses and student-centered styles are not limited to discussion courses.

Utilizing an intensive case study approach, Mann et al. (1970) described college teaching in terms of six teacher roles: Expert, Formal Authority, Socializing Agent, Facilitator, Ego Ideal, and Person. These roles can be interpreted as strategies or patterns of behaviours which might be utilized at various times by the same instructor. However, one would expect that instructors will differ in the relative amount of time they devote to each of these roles and in the manner in which they utilize these.

strategies in the classroom.

Three additional typologies have been offered in short, user-friendly, non-technical articles. Although they undoubtedly reach a wider audience than the work of Axelrod and Mann, they are even more lacking in scientific rigour. No explanation of the derivation of the classification is offered in any of these typologies. These papers are included here as somewhat extreme examples of the informal/observational approach.

Bonnici (1978) described two types of teachers: the Mechanic, who is a highly structured person valuing efficiency and organization, and the Artist, who is an unstructured and creative person. He also offers a 20-item questionnaire which the reader can use to answer the question, "Which type are you?"

Fischer and Fischer (1979), in an article emphasizing elementary and secondary school teaching, discussed six teaching styles: task-oriented, cooperative planner, child centered, subject centered, learning centered, and emotionally exciting vs. rational. They emphasize that a teaching style is distinct from a teaching method. A teaching style is "a classroom mode; a pervasive way of approaching the learners that might be consistent with several methods of teaching" (p. 251).

Hamachek (1972) classified teaching in terms of four neurotic styles expressed in a classroom context: obsessive-compulsive, paranoid, hysterical, and impulsive.

The instructor's particular neurotic style will determine his or her choice of teaching method. Additionally, Hamachek states that classroom situations may have the effect of exaggerating the behaviours characteristic of particular personality styles.

Formal/Theoretical Classifications. In contrast to the above typologies, other researchers have taken the approach of using a theoretical model to describe teaching styles. Variables are selected which according to the model should be relevant and empirical data are collected to test the fit of the model.

Abbott (1970) developed a theoretical model of college teaching style which specifies three distinct domains: cognitive, affective, and manipulative. The components of the model are: behaviours, traits, dimensions, and styles. These components were postulated to be developmental stages of style attainment. Using a selection of traits and dimensions from the teaching effectiveness research literature, Abbott constructed six written cameos of hypothetical teaching styles. She found some empirical support for the model in a factor analysis of student ratings of the written cameos. No data were presented to support the existence of these styles in an actual population of teachers.

Newport (1980) provided a set of components which can be used to operationally describe and identify teaching styles. The seven components are: specification of

objectives (behavioural, vague, none, or a combination),
determiner of learner activities (teacher or student),
pacing of activities (class, small group, or individual),
initiator of activities (teacher or student), determiner of
the direction of activities (teacher, student, or other),
focus of activities (objectives, task, subject matter, time,
interest, experience, or a combination), and evaluation
procedures (norm-referenced tests, criterion-referenced
tests, informal teacher observation, or a combination). The
style of any particular teacher would be fully described by
classifying his or her behaviour on each of these
dimensions.

Three studies have defined teaching style as a variant
of cognitive style. Wyett (1967) described both cognitive
styles and teaching styles in terms of three components:
symbol system (qualitative vs. theoretical), perception base
(family, associates, or individualism), and modalities of
inference (magnitude, difference, relationship, or
appraisal). Instructors were also described in terms of the
consistency of their teaching style (predominant vs.
switcher). Studying 27 student teachers, Wyett found that
performance decreased in teaching situations which required
the instructor to use an approach inconsistent with his or
her preferred cognitive and/or teaching style.

Raines (1976) described teaching style by use of the
Teaching Styles Inventory developed by Canfield (cited in
Raines, 1976). This measure, an adaptation of the earlier

Learning Styles Inventory by Canfield and Lafferty (cited in Raines, 1976), focuses on the cognitive style preferred by the instructor. Raines (1976) used Canfield's two measures to describe and compare student learning styles and instructor teaching styles within mathematics courses. He found that the learning styles of high-achieving students were more similar to the instructor's teaching styles than were the learning styles of low-achieving students.

Crookes (1977) compared instructors in the Arts and Science faculty to instructors in the Applied Arts and Science faculty in a community college. He identified three different teaching style groups. He found that the two faculties differed only in the prevalence of authoritarian (vs. permissive) teaching style.

Generally, the formal/theoretical approach to describing teaching style has not been characterized by an emphasis on observable behaviour. Abbott's focus was on more global traits and dimensions. Newport's measures are more accurately described as teaching methods, and Wyatt, Raines, and Crookes examined cognitive preferences. The extent to which teaching styles described in this manner are reflected in classroom behaviour is unknown. A description of teaching styles that is not linked to classroom behaviours will be of limited value in furthering our understanding of the teaching process.

Empirical Classifications. Other investigators have established typologies of teaching style through purely

empirical classification methods. This is the category in which the present research can be placed. Two groups of researchers have conducted studies in this area.

Bennett and his associates (Bennett, 1976; Aitkin, Anderson & Hinde, 1981; Aitkin, Bennett & Hesketh, 1981) conducted a large-scale study of teaching styles in British elementary schools, in which 12 distinct styles of teaching were identified among 468 fourth grade teachers, and teachers using each style were compared on student achievement measures. The original report of the study (Bennett, 1976) concluded that formal teaching styles yielded greater student gains in basic skills than informal styles. The conclusion generated much controversy in both the popular press and the statistical journals (Cray & Satterly, 1976; Bennett & Entwistle, 1976; Aitkin, Anderson & Hinde, 1981; Aitkin, Bennett & Hesketh, 1981). Some of the controversy revolved around the method used to compare gain scores in student achievement in intact classrooms. Aside from these issues, the identification of teaching styles in the Bennett study is relevant to the present research. Clusters were identified on the basis of a self-report questionnaire containing 38 yes-no items. The questionnaire focused on six major areas of teacher classroom behaviour: classroom management and organization, teacher control and sanctions, curriculum content and planning, instructional strategies, motivational techniques, and assessment procedures. It should be noted that these

items did not assess discrete behaviours such as "gestures with hands" or "puts outline on blackboard", but rather focused on general teaching strategies or methods such as "pupils allocated to seating by ability" and "emphasis on separate subject teaching". The reanalyses of the Bennett data by Aitkin and his colleagues in 1981 identified only three teaching styles, which could be defined as Formal, Informal, and Mixed. The difference between the two analyses can be attributed to the number of questionnaire items included in the analysis and the particular clustering technique applied.

Another group of researchers has used empirical methods to develop a typology of university lecturing styles. Brown, Bakhtar, and Youngman (1984) had 258 lecturers complete a 42-item self-report questionnaire comprised of 6 scales: Information giving, Structured lecture, Purposive lecture, Visual presentation lecture, Self-doubt lecture, and Presentation. Most items assessed specific lecturing behaviours (e.g., providing a list of headings to students), although some were more attitudinal in nature (e.g., "Is lecturing an economical method of teaching?"). Items were selected from an original set of 60 items and assigned to scales on the basis of an oblique factor analysis. Scales ranged in length from 4 to 12 items, with alpha reliabilities ranging from .52 to .63. A relocation cluster analysis was performed on the scale scores, but Brown et al. specify neither the distance measure nor the clustering

criterion chosen. Five clusters of teacher styles were identified:

(1) Oral Lecturer. These instructors were above average on structure, and below average on visual presentation and self-doubt. They were described as depending primarily upon oral presentation in their lectures. They typically did not use overheads and blackboards and did not present lecture notes or summaries of objectives to students.

(2) Exemplary Lecturer. These instructors were below average on information giving and self-doubt, and above average on the remaining four scales. They were described as being opposed to reading the lecture from notes and to giving too detailed information to students.

(3) Information Providers. This cluster was marked by high scores on information giving, structure, and visual presentation, and by low scores on self-doubt and presentation. These lecturers were described as "newscasters" who provided too much detail in their lectures.

(4) Amorphous Lecturers. This group of instructors had the lowest mean profile, with below average scores on the structure, purposive, and presentation scales. According to Brown et al., they had no selected objectives for their lectures, and did not structure lectures well.

(5) Self-doubters. These instructors showed high scores on the self-doubt scale, and scores which did not differ significantly from the average on the remaining scales. They admitted to having difficulty in selecting and structuring material and in keeping to their notes when lecturing. Consequently, they stated that they often felt that they had not met their objectives for a given lecture.

Brown et al. found significant differences in the prevalence of various styles across different academic disciplines. Oral lecturers were more common in humanities and social sciences, whereas Exemplary lecturers were more common in biomedical sciences, and Information Providers and Amorphous lecturers were more common in applied sciences such as engineering. Generally, instructors of applied sciences differed most from the other groups. These instructors were more likely to write down the whole lecture when preparing, to use complex diagrams, and to provide full notes to students, either on the blackboard or in printed form. Neither years of experience nor academic rank was found to be significantly related to teaching style.

The present research will clarify and extend the results reported by Brown and his colleagues. First, the description of teaching styles reported here relied upon discrete classroom behaviours, as observed and rated by students. Second, the behaviours were observed and rated by

more than one student, and teacher mean scores were used as the unit of analysis. Such data will be more reliable, and possibly more valid, than self-report measures. Third, the present research sought to validate the obtained classification through statistical replication as well as through investigation of empirical relationships with theoretically relevant variables.

Development of a Teaching Style Classification

As outlined above, teaching styles can be classified through an informal/observational approach, a formal/theoretical approach, or an empirical approach. Each of these methods of classification is exemplified in the literature reviewed above. Non-empirical approaches are problematic in any field where current understanding of the phenomena under study does not allow for clear-cut descriptions and typologies. Teaching behaviour is one such field.

Hence, the emphasis in the proposed research is on the formal, empirical approach to classification, also referred to as numerical taxonomy (Sneath & Sokal, 1973).

Statistical methods exist for the identification of naturally occurring patterns of observations, and such methods serve as useful adjuncts to the informal/observational and formal/theoretical approaches. They are particularly useful in applications where our understanding of the observations and their co-occurrences is not

sufficiently well-developed to generate formal and rational classification systems. The goal of these methods is to identify individuals who can be considered similar in some way. In the present research on classroom teaching behaviour, the similarity between instructors was determined by first describing each instructor's behaviour in a quantitative manner and then computing a numerical index of similarity between all possible pairs of instructors. This index of similarity provided the basis for identifying a small number of patterns of classroom teaching behaviour, or teaching styles.

Description of Teaching Behaviour. The present research relied upon a multidimensional description of teaching, with attention restricted to specific and observable classroom lecturing behaviours. These descriptions were derived from student responses to a standardized questionnaire arranged in a behavioural frequency rating format.

As in any multidimensional approach, the focus of measurement (the instructor in this case) can be described by a vector representing scores on several attributes (in this case, behaviours). When this set of scores is depicted as a profile of scores, each individual's profile can be completely described in terms of three characteristics: elevation, shape, and scatter (Cronbach & Gleser, 1953).

Elevation is the mean level of the profile, or the extent to which the individual's scores are generally high.

or low. In teacher evaluation, profile elevation will represent overall teaching effectiveness if all variables in the profile are positively related to student outcome.

The shape of the profile refers to the pattern of high and low scores. Profile shape may summarize the style of teaching, such as student-centered or teacher-centered.

Profile scatter can be measured by the standard deviation of each individual's scores and indicates the amount of dispersion around the individual's mean level. Scatter may reflect the degree to which an individual displays his/her preferred style. A small amount of scatter among the scores would represent a mild or fairly neutral example of the style, such as someone who is "marginally enthusiastic and slightly disorganized". A large spread in scores might represent an exaggerated expression of even a caricature of the style, as in "extremely enthusiastic but totally disorganized".

The search for similar individuals becomes a search for similar profiles -- those which have similar elevation and/or shape and/or scatter. Any of these components could be eliminated from the similarity measure by the appropriate preprocessing of the data. In the present research, profile elevation, shape and scatter were considered simultaneously in the identification of those individuals who might be considered to be similar to each other.

An alternative approach is to give the profile or vector of scores a geometric interpretation. In this sense,

the vector gives an individual's location in the multi-dimensional space. The search for similar individuals thus becomes a search for individuals who are placed close to each other in the multidimensional space. A common metric for this purpose is the Euclidean distance between the two data points, which is calculated as the square root of the sum of the squared differences in the corresponding scores. Clearly, this value ranges from zero to some maximum value determined by the scale of the measures used.

Identification of Teaching Styles. Patterns of teaching behaviour, or teaching styles, were used in this study to summarize and describe the observed behavioural similarities between different instructors. There are two major classes of statistical methods for summarizing or grouping data in order to identify these patterns. Both approaches are based essentially on the similarities of pairs of data points. The first method identifies underlying dimensions, whereas the second identifies clusters of contiguous data points.

The dimensional approach includes multidimensional scaling and R- and Q-type factor analysis. These methods are the ones most often associated with psychology, largely due to the early factor analytic work on the nature of intelligence. The goal is to find the minimum number of dimensions which will reproduce the original data with a high degree of fidelity, and then to explain the observations in terms of these dimensions.

The second approach uses clustering methods to identify homogeneous sets of individuals who are more similar to each other than they are to other individuals in the sample. Typically a distance function is calculated from the raw data for each pair of individuals, and clusters of similar individuals are created on this basis. The centroid of each cluster can be considered to be the point which best represents the group, or an "ideal type" described by the group mean profile. A new individual can be classified by determining his similarity to, or distance from, each cluster centroid. Classification can then be made solely on the basis of the individual's closest ideal type, or his characteristics can be described as a mixture of two or more ideal types (Skinner, 1979).

The appropriateness of dimensional approaches for classification is unclear. When the aim of the research is to describe or classify individuals according to their position along one or more axes (e.g., an introversion-extraversion continuum), then certainly these are the methods of choice. However, when the goal is to describe individuals in terms of one or more homogeneous groups to which they are similar, then clustering methods would seem to be more appropriate. Skinner (1979) has suggested a two-stage hybrid approach which calls for the use of factor analysis first as a data reduction and error-purging technique, then the application of cluster analysis to factor scores. An analogous approach was utilized in the

present research, where scores on teacher behaviour scales, rather than individual behaviours, served as input to the cluster analysis. The idiosyncracies (unique and/or error variance) of individual items can, to some extent, be controlled in this manner. It was assumed that results obtained in this way would be more stable, interpretable, and theoretically meaningful than results based on individual items.

Validation of a Teaching Style Classification

Empirical methods of developing a classification scheme will always produce a grouping of individuals, regardless of the natural structure of the data. It is essential, therefore, that such groupings be carefully validated.

Skinner (1981) has described two stages in this validation process. Internal validation is conducted by applying alternative empirical methods to the sample data used to derive the original classification. These methods include replication of the classification in split halves of the sample, in alternative subsets of variables, and through mathematically different methodologies. Such replications are described along with the original development of a typology of lecturing styles in Study 1 below. The second stage in validation, external validation, requires additional experimental or correlational support for the meaningfulness of the typology. Investigations in this vein are reported in Studies 2 and 3 below. Relationships

between teaching styles and other relevant educational variables were examined in Study 2, whereas consistency in instructors' teaching styles across years and across courses was examined in Study 3.

STUDY 1

Study 1 was designed to develop empirically a classification of teaching styles through identification of naturally occurring patterns of classroom behaviour exhibited by university lecturers. Within this development process, analyses were conducted to assess the replicability of the clusters identified across different samples of instructors, different types of variables, and different analytical techniques. These investigations fulfill the need for internal validation of the typology, as outlined by Skinner (1981). Finally, the teaching styles identified were compared on two other available measures: subject area and level of course. These comparisons form part of the process of external validation of the typology. More importantly, however, they allow us to examine the extent to which teaching style varies across discipline and course level.

Method

Subjects

Over a period of three successive academic years (1976-1979), student observations of instructors' low-inference classroom teaching behaviours were collected from 298 classes at the University of Western Ontario.

These data were collected in the context of research projects conducted by Professor H. G. Murray and/or by the present author. Instructors in 130 of these classes had agreed to have ratings collected as part of Professor Murray's original development work on the measure of classroom teaching used in this study, namely the Teacher Behaviors Inventory. These instructors were selected from the faculties of Arts, Science, and Social Science, and included individuals of varying overall teaching effectiveness as measured by student instructional ratings over the previous five academic years. Instructors in other classes ($n=123$) had volunteered to participate in subsequent studies involving feedback regarding specific classroom behaviours, whereas other instructors ($n=45$) were asked to participate in studies because they taught class sections in large multi-section courses with common final examinations.

Instructors who participated in more than one of these studies appear more than once in the present data set. With the exception of two classes whose instructors were evaluated twice, once in each semester, all cases in the data set are different classes. Thus, the 298 cases represent 296 classes taught by 237 different instructors.

All instructors were rated in lecture-style courses with enrollments of 30 or more. Some instructors ($n=69$) were rated by trained outside observers, mostly senior psychology students, who observed the instructor for a minimum of three class sessions before completing the

Teacher Behaviors Inventory. Most instructors ($n=229$), however, were rated by students enrolled in their own class. In these cases, students were randomly selected to receive the behaviour-rating questionnaire. In 36 of these 229 classes, students were selected from among those present in class on the day of the researcher's visit. In the remaining 193 classes, students were randomly chosen from class lists and approached by mail. The number of raters per class ranged from 1 to 21, with an average of 6.93.

Measures

Student observations of classroom behaviours were collected using the Teacher Behaviors Inventory or TBI (see Appendix A) developed by Murray (1977) and described more fully by Murray (1983a). The TBI consists of 100 items, grouped into nine categories: mannerisms, explanation, organization, interaction, speech, interest, disclosure, rapport, and affect. Each item refers to a specific, observable classroom teaching behaviour, such as "puts outline of lecture on blackboard" or "moves around class while lecturing." Concrete behaviours such as these, which can be recorded by direct observation, have been labelled "low-inference" behaviours by Rosenshine and Furst (1971). Students or outside observers were asked to rate the frequency with which their instructor exhibited each of these behaviours, on the following 5-point scale: (1) never, (2) rarely, (3) sometimes, (4) often, (5) always.

Student or observer responses were averaged within each class to yield mean scores on each of 100 items for each instructor.

Previous research conducted with the TBI (e.g., Erdle & Murray, 1986; Murray, 1983a, 1983b; Murray, Rushton, & Paunonen, 1986) has amply documented its psychometric characteristics. Interrater reliabilities for each of the 100 items have typically ranged from .4 to .9, with a mean of approximately .7 (based on 7 observers). These results indicate good agreement among observers and, therefore, good generalizability of results across different groups of students within a class. Also, the behaviours included in the questionnaire have been found to distinguish among instructors receiving consistently high, medium, or low student ratings of overall teaching quality (Murray, 1983a), and to correlate with a variety of student outcome measures, including final examination performance (Murray, 1983b).

To facilitate interpretation of instructor profiles, TBI scale scores were computed. Eight scales had been developed in previous research on the basis of factor-analytic results and rational consideration of item content. Item scores were first standardized with reference to normative data available for the TBI items. Then scale scores were computed as unweighted means of the appropriate items. Table 1 presents the name, definition, and a number of relevant items for each scale.

Table 1

Definitions and Sample Items of TBI Scales

Disclosure (DS): instructor's explicitness and helpfulness regarding course requirements and grading criteria (9 items)

makes students aware of overall objectives of course
tells students exactly what is expected of them
on tests, essays, or assignments
provides sample exam questions

Expositional Style (ES): Behaviours serving to explain, clarify, and/or generate student interest in subject matter (19 items)

repeats difficult ideas several times
gives everyday, real-life examples to
illustrate concepts or principles
clearly indicates transition from one topic
to another

Expressiveness (EX): Instructor's use of movement, gesture, and vocal inflection to capture student attention (12 items)

moves back and forth in front of class
gestures with hand or arms
speaks in a "dramatic" or expressive way

Organization (OR): Ways in which instructor attempts to organize subject matter; methods used to create structure and coherence (11 items)

explains how each topic fits into the course
as a whole
puts outline of lecture on blackboard or overhead
screen
periodically summarizes points already made

Rapport (RP): Instructor's attitudes and feelings toward students; interpersonal relations between teacher and students (27 items)

praises students for good ideas
knows individual students by name
tolerant of other points of view

Relaxed Manner (RM): Indications that the instructor is comfortable and at ease in front of the class (8 items)

exhibits distracting mannerisms (negatively keyed)
explains subject matter in informal, colloquial language
reads lecture from prepared notes or text, thus reducing spontaneity

Speaking Ability (SA): Qualities of instructor's voice; manner of talking or speaking (7 items)

pauses in mid-sentence (negatively keyed)
says "um" or "ah" (negatively keyed)
does not speak clearly (negatively keyed)

Use of Class Time (UT): Manner in which class time is utilized; pacing of lecture (5 items)

dwells excessively on obvious points (negatively keyed)
digresses from major theme of lecture (negatively keyed)
sticks to the point in answering students' questions

Note: Two of the 100 TBI items are not included in any scale.

Cluster Analysis

A cluster analysis was conducted to identify groups of instructors with similar teaching behaviour profiles. Scores were standardized in each analysis to avoid the possibility of the cluster solution being unduly influenced by variables (e.g., TBI scale scores) with large variance. No other data preprocessing was done. Consequently, elevation, scatter, and shape effects were combined in the measurement of profile similarity.

The cluster analysis had three major components, each conducted with the CLUSTAN package (Wishart, 1978). First, a hierarchical clustering of the intersubject Euclidean distances was conducted, using Ward's minimum variance method. This particular clustering approach starts with individual subjects as clusters and successively merges pairs of clusters which produce the least increase in the total error sum of squares. Error is defined as the distance from each individual to the centroid of its parent cluster (Wishart, 1978, p. 3). The resultant clusters tend to be small-variance and spherical. Monte Carlo studies reviewed by Punj and Stewart (1983) suggest that clusters identified by hierarchical methods may not be optimal. Therefore, the second component of the analysis was a K-means iterative relocation of the individual subjects, to optimize the hierarchical solution. Starting with an existing classification (e.g., the hierarchical solution), the K-means algorithm moves individuals from one cluster to

another to improve the solution. The criterion for improvement of the solution was again a decrease in the error sum of squares. The third component was an identical K-means analysis, but with a random cluster configuration used as a starting point. This offers some evidence regarding the mathematical stability of the cluster solution within the sample.

Results

Psychometric Properties of Measures

Interrater agreements for the 100 items of the TBI were computed using the intraclass correlation approach advocated by Ebel (1951). Analysis of variance procedures were used to estimate between-teacher effects (ANOVA "treatments") and within-teacher effects (ANOVA "subjects"). Ebel's formula, $(MSB - MSW)/MSB$, was then used to estimate the reliability of the mean rating obtained from k raters (mean $k=6.9$ in the present data set). The reliabilities of the mean ratings for the 100 TBI items ranged from .42 to .92, with a mean of .69 and a median of .68, indicating that there is agreement in the ratings provided for the same instructor by different observers and that an instructor's mean rating on an item is measured reliably.

Overall means, standard deviations, internal consistency (alpha) reliabilities, and intercorrelations of

the 8 TBI scales in the present sample of 298 classes are given in Table 2. Internal consistency estimates were generally high, ranging from .65 to .94, indicating that each scale is composed of a homogeneous set of behaviours. It may be noted that scale intercorrelations were generally positive in direction and moderate to high in magnitude. The one exception was Use of Class Time, which had low correlations with the other scales.

Classification Results

Hierarchical clustering, subsequently optimized by a K-means procedure, was performed on TBI scale scores. Six clusters appeared to replicate across the two K-means starting positions. Solutions with greater numbers of clusters did not replicate clearly. Replication was assessed through examination of cluster membership (i.e. which subjects were allocated to each cluster) and examination of cluster mean scores on the eight scales. The six cluster mean profiles are presented in Figure 1, as deviations from the grand mean. Evidence presented below confirms that six clusters can be interpreted with confidence. A description of the clusters at this point will facilitate understanding of later sections.

Cluster A individuals are good speakers who use class time efficiently and who are relaxed and confident in front of a class. Compared to other instructors, they avoid "spoon-feeding" students -- that is, they give limited

Table 2

Mean, Standard Deviation, Internal Consistency
and Intercorrelations of TBI Scales (Study 1)

	Scales							
	DS	ES	EX	OR	RP	RM	SA	UT
DS	89							
ES	40	90						
EX	27	52	88					
OR	49	61	22	87				
RP	46	70	51	51	94			
RM	25	58	45	35	67	65		
SA	32	54	47	41	55	64	86	
UT	04	26	07	32	18	31	45	78
Mean	.44	.11	.00	.23	.02	.09	.01	.19
SD	.73	.54	.65	.68	.61	.52	.67	.74

Notes: (1) Decimals have been omitted from the correlations.
(2) Diagonal elements are alpha reliabilities.

Figure 1

Cluster Profiles (Study 1 Sample)

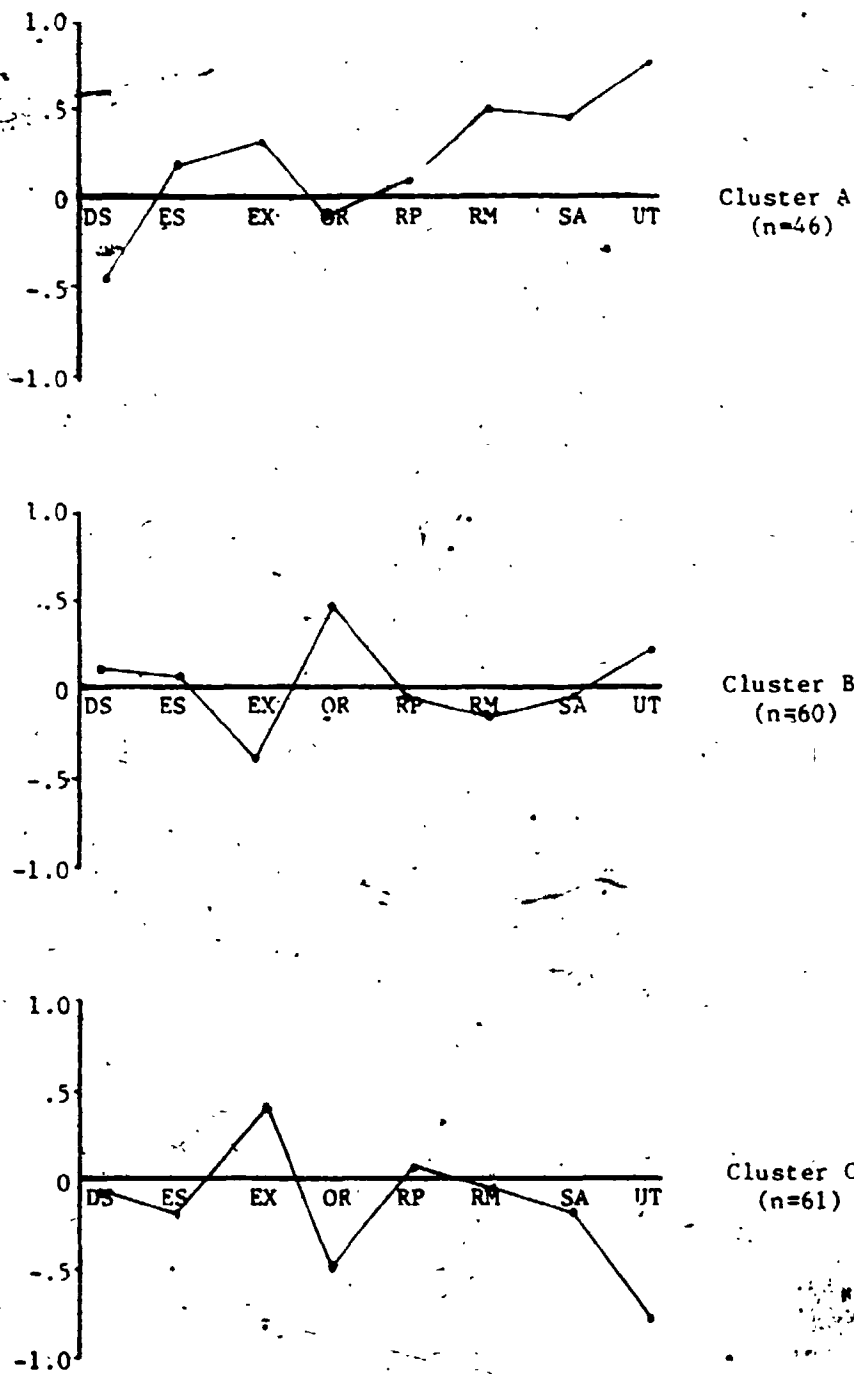
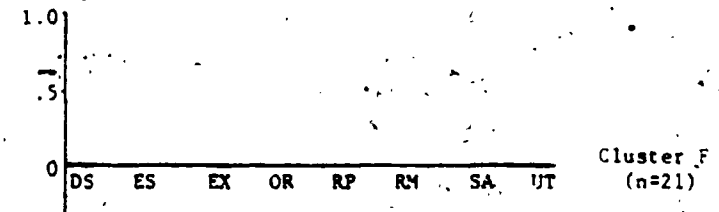
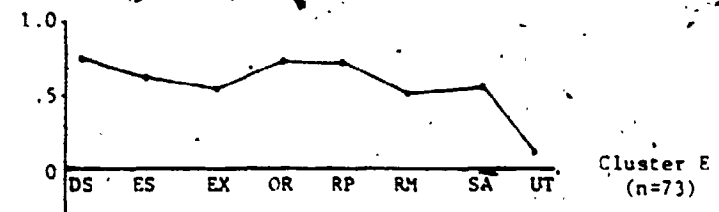
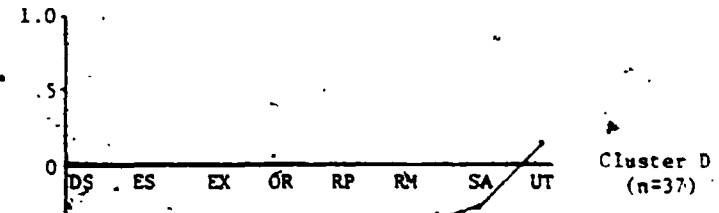


Figure 1 (cont'd.)



38

information regarding sample exam questions and grading criteria, and do not remind students of deadlines or provide helpful hints for learning the course content. They can be labelled Speech Givers.

Cluster B contains instructors who are low on Expressiveness, indicating that they tend to show very little movement or gesture in class and tend to speak in a slow, droning monotone. They are, however, above average on Organization, indicating that they clearly state the objectives of each lecture, explain how it fits into the course as a whole, and organize the material within each lecture according to a clearly defined structure. These instructors can be called Structured but Boring.

Cluster C instructors are almost the mirror image of the Cluster B group. They are expressive and enthusiastic lecturers who reveal very little organization or structure in their presentation of course content. In a given lecture, they are likely to use time inefficiently by digressing from the main topic and covering very little material. These individuals, then, are Entertaining but Unstructured.

Cluster D instructors have a profile of scores that are below average in everything except Use of Class Time. On the basis of Murray's (1983b) data indicating generally positive relationships between TBI scores and various measures of teaching effectiveness, the instructors in this cluster can be characterized as Ineffective Lecturers.

Cluster E instructors are generally above average. Because their profile is almost the mirror image of the Cluster D profile, they can be described as Effective Lecturers.

Cluster E contains individuals who receive extremely low scale scores, particularly on Speaking Ability. They are described by students as apprehensive speakers who frequently say "um" or "ah", mumble, speak softly, and pause or stop in mid-sentence. Although they are below average in all areas, these instructors are perhaps best characterized as Mumblers.

Item-Scale Cluster Comparison

Cluster analyses identical to those described above were conducted using item scores rather than scale scores as input data. Cluster memberships were compared for classifications based on items and scales. These results are presented in Table 3. In general, instructors who were grouped together on the basis of one set of data tended to be similarly grouped on the basis of the other. The two classifications were statistically related, $\chi^2(25) = 729.30$, $p < .001$, and inspection of the entries in Table 3 confirms that there was a high degree of agreement in the allocation of individuals to clusters. As further confirmation of the similarity between the two cluster solutions, the order in which pairs of individuals (and clusters) were merged in each hierarchical analysis was

Table 3
Comparison of Cluster Memberships
Using Scale and Item Data

		Clusters Based on Item Data						
		A	B	C	D	E	F	Total
Clusters Based on Scale Data	A	23	14	6	0	3	0	46
	B	3	50	1	3	3	0	60
	C	0	13	39	8	1	0	61
	D	0	0	3	27	0	7	37
	E	6	1	0	0	66	0	73
	F	0	0	2	3	0	16	21
	Total	32	78	51	41	73	23	298

Note: Table entries are numbers of subjects classified into each cluster

examined. The cophenetic correlation between the two hierarchical solutions was .63, indicating substantial agreement in the structure and ordering of allocation of cases to clusters. Because of the similarity of the results obtained with item and scale data, only the analyses conducted on scale data are reported.

Cluster Replication Across Split-Half Samples

The total sample of 298 subjects was split (on an odd-even basis) into two half-samples of 151 and 147 subjects. The data from each half-sample were analyzed independently in order to test the replicability of the clusters obtained. Cluster analyses identical to those conducted with the total sample were conducted with each half-sample. Hierarchical clustering, subsequently optimized by a K-means procedure, was performed on TBI scale scores. Within each half-sample, six clusters replicated in the two K-means analyses (based on two different starting positions). Solutions with larger numbers of clusters did not replicate well, and therefore only the six-cluster solution was examined further. The cluster means from the two half-samples were subsequently input as raw data to the same cluster analysis procedure to further assess replication across samples. As expected, each of the six clusters in each half-sample clustered with only one corresponding cluster from the other sample. This finding provides additional support for the "reality" of the

clusters (teaching styles) identified.

A Monte Carlo study was undertaken to assess the degree to which replication of clusters across the two split-halves of the sample could be due to capitalization on chance. Scale scores were simulated for 49 samples of 298 cases, with scale intercorrelations designed to reproduce correlations obtained in the observed sample data. This was accomplished by generating for each subject 8 orthogonal scores from the normal distribution, premultiplying this score matrix by the full-rank principal components solution of the eight scale scores in the original data, and then scaling the obtained scores to reproduce the standard deviations of the original data. In this manner, 49 new samples of 298 "individuals" were generated. In each sample, cases were allocated to one of two half-samples and cluster solutions obtained.

The distances between cluster centroids in each pair of half-samples of simulated data were computed and compared to the distances obtained with observed data. The distances between cluster centroids in the observed data are reproduced in Table 4. It can be seen that the replication was quite clear in the observed data, with each cluster in one solution being close to only one cluster in the other solution. The average distance between "matching" clusters (i.e., the average of the diagonal elements in Table 4) was 644. The average distance between "non-matching" clusters (i.e., the average of the off-diagonal elements in Table 4)

Table 4

Observed Intercluster Distances in Split-Samples

		Sample 1					
		1	2	3	4	5	6
Sample 2	1	.39	2.1	1.7	1.6	3.2	3.8
	2	2.0	.52	1.7	1.7	2.2	2.1
	3	2.0	1.8	.54	1.5	2.3	3.4
	4	1.6	1.2	1.4	.57	1.8	2.6
	5	2.8	1.1	1.8	1.5	.87	1.5
	6	4.6	2.8	3.8	3.4	2.2	.98

was 2.24. In comparison, of the 49 pairs of cluster solutions of simulated data, only one had an average diagonal element less than that obtained in the observed data (distance = .571), and only three had average off-diagonal elements greater than that obtained in the observed data (distances = 2.29, 2.28, 2.63). The ratio of the diagonal to the off-diagonal can be taken as an estimate of how close the replicated clusters are to each other, relative to their distances to other clusters. Smaller values indicate greater replication. In the observed data, this ratio is .287 (.644/2.24). None of the values obtained with simulated data was lower than this value (Mean = .483, SD = .104).

Hope (1968) has discussed this type of comparison of a single observed value to several simulated values as a permutation test. Under the null hypothesis, the probability that the single observed value is larger than all s simulation values is $s!/(s+1)!$ or $1/(s+1)$. This suggests that the replication of clusters derived in the current data set could be considered "statistically significant" at the $p < .02$ level. These results provide a certain added measure of confidence in the replication of the clusters in the observed data.

Wishart's Test for Hierarchical Clustering

A "stopping rule" for the number of clusters in a hierarchical analysis has been proposed by Wishart (1982).

The test is somewhat analogous to the scree test in factor analysis. It identifies a point of discontinuity in the particular criterion being optimized by the algorithm, indicating a noticeable improvement/decrement in fit. For example, in the solution of TBI scale data, the error sum of squares for 11 clusters was almost 3 standard deviations above (i.e., worse than) the predicted value based on the linear trend of solutions with decreasing numbers of clusters, indicating that 12 clusters, rather than 11, should be selected. Slightly different estimates were obtained by varying the number of points used to establish the linear trend. Results consistently showed strong support for up to 7 clusters ($z=4.8$ approx.), whereas support for 11 or 12 clusters was somewhat weaker. Very similar patterns of results were found with the item scores. As discussed earlier, however, replicability within and across samples indicated that a more conservative number of 6 clusters can be interpreted with confidence.

Effect of Outliers

Outliers were defined as individuals lying more than 2 distance units from their cluster centroid. Sixteen outliers were identified and eliminated from a trial reclassification of scale data. No major changes were noted in the cluster centroids or memberships, with the exception of Cluster F (Mumblers). This cluster of very low-scoring individuals had a large variance relative to other clusters.

and contained nine of the sixteen outliers. The elimination of these cases led to the re-allocation of 15 low-scoring cases from other clusters into Cluster F, giving a mean profile for the resulting 27 cases which was less extreme than the original on all scores. On this basis it can be concluded that the solution adopted in this study was relatively robust with respect to outliers. Therefore, the outlier cases were retained in all further analyses.

Effect of Reordering Cases

The particular algorithm utilized for the K-means clustering procedure, the RELOCATE subroutine in the CLUSTAN package (Wishart, 1978), updates each cluster centroid after reallocating any individual case from its original cluster to a closer cluster. One effect of this approach is that the obtained solution is dependent, to some extent, upon the ordering of the cases in the data set. A very brief investigation was undertaken to assess the extent of this effect in the present study.

Twelve different orderings of the subjects were analyzed, using a hierarchical minimum-variance algorithm optimized by the K-means procedure with the same minimum-variance criterion. The 6-cluster solutions were compared across the twelve analyses.

Six different solutions were obtained from the twelve orderings. The within-cluster sum of squares did not differ appreciably from one solution to another. However,

allocation of cases to clusters did vary across solutions. Three of the six clusters -- i.e., Cluster A (Speech Givers), Cluster E (Effective Lecturers), and Cluster F (Mumblers) -- appeared to be quite stable. In contrast, the numbers of cases in Cluster C (Entertaining but Unstructured) and Cluster D (Ineffective Lecturers) varied by as much as 40%. Despite variation in allocation of cases to clusters, the mean profiles of the six clusters were very similar across all solutions.

In conclusion, it appears that the identification of an individual instructor's teaching style may be somewhat unstable. However, the six teaching styles which have been identified reappear consistently across all analyses.

Relationship to Subject Area and Level of Course

For each of the 298 classes in the present data set, the level of the course and the department in which it was offered could be determined. Clusters were compared on each of these variables to determine whether teaching style was dependent on class characteristics.

Classes were classified into four levels; (1) first year introductory survey courses; (2) second year general interest courses designed for students not majoring in the subject area; (3) second year honours courses; and (4) third and fourth year honours courses. The relationship between instructor's teaching style and course level for the 298 classes used in this study is presented in Table 5.

Table.5

Relationship between Teaching Style and Course Level

(Study 1)

Course Level	Teaching Style Cluster						Total
	A	B	C	D	E	F	
Introductory	27	24	19	15	26	4	115
General	6	7	20	7	10	11	61
Junior Honours	9	23	16	13	24	5	90
Senior Honours	4	6	6	2	13	1	32
Total	46	60	61	37	73	21	298

Note: Table entries are numbers of glasses..

There was a statistically significant relationship between these two variables. $\chi^2 (15) = 37.61, p < .001$, indicating that teaching style was not independent of course level. In examining the entries of Table 5: it can be seen that instructors in typically large, multi-section introductory (Level 1) courses were more likely than other instructors to be Speech Givers (Cluster A). In the Level 2 (general interest) courses, instructors tended to show the Entertaining but Unstructured style (Cluster C), perhaps reflecting their need to maintain the interest of large classes of students who are not expected to study further in the area. There was also a large proportion of Mumbler's in the Level 2 courses. Although the cause for this cannot be determined from the data, it might reflect a tendency for departments to assign inexperienced and/or below-average lecturers to these lower priority courses. In second-year honours (Level 3) courses, there was a tendency for instructors to be Structured but Boring (Cluster B) rather than Speech Givers or Entertaining but Unstructured. Whether this is due to a conscious choice of teaching style by the instructor in order to adapt to the course, or whether it is a function of differential assignment by departments cannot be ascertained. Although the number of senior honours (Level 4) classes was small, there was a tendency for Effective Lecturers (Cluster E) to be overrepresented at this level.

Classes were also classified into one of three major subject areas: (1) Arts, including English, Music, Modern Languages, and Philosophy; (2) Sciences, including Computer Science, Engineering, Medicine, Anatomy, Chemistry, Physics, and Mathematics; and (3) Social Sciences, including Psychology, Economics, History, Geography, Sociology, Anthropology, Political Science, and Administrative Studies. The Social Science area accounted for 191 of the 298 classes, or almost two-thirds. As shown in Table 6, there was a statistically significant relationship between teaching style, as defined by cluster membership, and subject area. $\chi^2(10) = 25.45, p < .005$. In examining the entries in Table 6, it can be seen that Arts instructors were more likely than instructors in other fields to be Entertaining but Unstructured (Cluster C), and also were more likely to be Effective Lecturers (Cluster E). Science instructors, on the other hand, were slightly more likely to be Ineffective Lecturers (Cluster D) or Mumlbers (Cluster F) and less likely to be Speech Givers (Cluster A).

Summary

A variety of classification methods were employed in Study 1 to group 298 instructors into clusters on the basis of 100 behavioural ratings. The six-cluster classification based on an optimized minimum-variance hierarchical clustering of 8 TBI scale scores was the best-replicated and best-supported of the solutions. Very similar cluster mean

Table 6

Relationship between Teaching Style and Subject Area

Subject area	Teaching Style Cluster						Total
	A	B	C	D	E	F	
Arts	5	5	15	3	16	4	48
Science	4	8	15	12	13	7	59
Social Science	37	47	31	22	44	10	191
Total	46	60	61	37	73	21	298

Note: Table entries are numbers of classes.

profiles were obtained from different starting positions and in different subsets of the data. The grouping of instructors into clusters was similar whether the data were scale scores or item scores.

Cluster A instructors, the Speech Givers, were identified as good speakers who use class time efficiently but do not give much information regarding course requirements. Many instructors teaching large introductory courses were found to fall into this category. The Cluster B mean profile showed a style that might be described as Structured but Boring, in other words high scores on Organization and low scores on Expressiveness. The pattern of scores for Cluster C was almost a direct reversal of the Cluster B profile, a style that might be labelled Entertaining but Unstructured. Many instructors teaching Arts courses and general interest courses were found to fall into this cluster. Cluster D instructors were below average on all scales except Use of Class Time, and thus can be described as generally Ineffective Lecturers. Instructors in Cluster E, above average on all scales, can be described as generally Effective Lecturers. The Cluster F mean profile showed extremely low scale scores, but was marked by particularly low ratings for speaking ability. The individuals in this group were labelled Mumbler.

It is clear that elevation of scores played a major role in determining cluster membership. This can be seen by referring to the mean profiles for clusters D, E, and F

presented in Figure 1. Cluster E was above average on all eight scores, whereas F was far below average on all eight, and D was moderately low on all but Use of Class Time. Additional support for the importance of elevation could be seen in comparing the six-cluster solution to others obtained when the effect of elevation had been largely removed. Cluster analysis of data from which subject means had been removed, as well as modal profile analysis of intersubject correlations, yielded classifications substantially different from that reported here.

Teaching style was found to be significantly and meaningfully related to both the level and the subject area of the course. The causes of these relationships cannot, of course, be determined from the available data. With regard to course level, the results indicated that instructors of introductory or general interest courses tended to utilize a teaching style which favours Expressiveness over Organization (Speech Giver or Entertaining but Unstructured). Instructors of second-year honours courses tended to display an opposite style (Structured but Boring). With regard to subject area, it was found that Arts instructors tended to be Entertaining but Unstructured, and also were more likely to be Effective Lecturers. The low profiles of the Ineffective Lecturers and the Mumbler were over-represented among Science instructors.

STUDY 2

Once a reliable typology of teaching styles has been identified, it is important to investigate the relationship of teaching styles to other relevant educational variables. Recent research has shown significant and predictable relationships between lecturer behaviours and several other relevant variables, including student outcome measures. In his review of the research on university lecturing, Dunkin (1983) concluded that although tentative conclusions can be drawn regarding the positive effects of particular lecture characteristics such as content coverage, clarity, and expressiveness, no generalizations can yet be made regarding lecturing styles, or patterns of behaviour.

There is some evidence that patterns of teaching behaviour differ across academic disciplines. Differences in the measures employed and lack of rigour in conducting and reporting analyses make it difficult to draw firm conclusions in this area, however. Crookes (1977) found that instructors of community college Arts and Science courses could be distinguished from instructors of Applied Arts and Science courses in the frequency of authoritarian vs. permissive teaching styles. The extent to which this is reflective of classroom teaching behaviours is not clear. In a study focusing on low-inference items and controlling for halo effect due to the student's overall liking for the instructor, Miller (1979) found significant differences.

across 20 disciplines in the level and shape of instructor profiles of all teaching dimensions. Halo effect did not alter the relative shapes of the profiles. Erdle and Murray (1986) compared university instructors in Arts, Social Science, and Science faculties on 7 classroom behaviour factors derived from outside observer ratings of the 100 behaviours listed in the Teacher Behaviors Inventory. They found that Arts teachers were significantly higher than Science teachers on interaction, interest, and rapport factors, while Social Science teachers were significantly higher than Science teachers on the interest scale, and Arts instructors were significantly lower than both other groups on organization. As described previously, Brown, Bakhter and Youngman (1984) also found a statistically significant interaction between lecturing style and subject area. Generally, Science instructors were more likely to provide students with detailed content information than were instructors in other areas.

In summary, then, there is reason to expect that classroom teaching behaviours will in fact be seen to vary across academic disciplines. This hypothesized relationship could be confirmed by significant differences in scores on teaching behaviour scales, or by significant differences in the frequency of various teaching styles or clusters across subject areas. Dunkin and Biddle (1974) have described this type of research as fitting a presage-process paradigm, in which stable instructor characteristics such as personality

traits, or in this case academic discipline (presage variables) are related to measures which describe classroom teaching behaviour (process variables).

In a process-product study Murray (1983b) found consistent patterns of correlation between classroom teaching behaviour factors as assessed by the Teacher Behaviors Inventory and six different measures of overall teaching effectiveness including student instructional ratings, final examination achievement, and subsequent course enrollment. Similarly Murray (1983a) found that several TBI behaviours differentiated significantly among instructors who had consistently received high, medium or low student ratings over the previous five academic years. These findings are in keeping with the substantial body of research literature supporting the validity of student ratings, and provide further evidence that classroom teaching styles can be expected to show meaningful relationships with other educational variables such as global student ratings of teaching effectiveness.

Feldman (1986) recently reviewed the research on the relationships between instructor educational attitudes (presage variables) and overall teaching effectiveness (a product variable). The results of these studies are mixed. For each significant relationship there seems to be at least one negative result with very similar variables. The interpretation of this research has been hampered by the omission of mediating process variables, or teaching behaviours.

Erdle, Murray and Rushton (1985) investigated relationships among presage, process and product teacher variables within a single study. They found that classroom teaching behaviours as measured by the TBI accounted for approximately 50% of the covariance between instructor personality and rated teaching effectiveness in university lecturers. Considering that their behaviour measure was limited to two factor scores representing only 29 of the 100 TBI teaching behaviours, this is indeed a substantial relationship. This finding encourages investigation of the role of classroom teaching behaviour in the relationship between instructor characteristics and student outcomes.

Study 2 was designed to investigate relationships between teaching styles and other educationally relevant variables. In Skinner's (1981) terminology this study sought to continue the external validation of the typology developed in Study 1. A new sample of instructors was identified and their teaching styles classified into one of the six "types" described above. Relationships were investigated between teaching styles on the one hand and course and instructor characteristics, instructor self-ratings, and global student ratings on the other. Whereas previous research has examined these latter variables in relation to individual teaching behaviours or factor scores, no previous study has investigated these variables in relation to patterns of teacher behaviour, or in other words, teaching styles.

Method

Subjects

All full-time faculty members at the University of Western Ontario were approached by letter in the winter of the 1984-1985 academic year, and were asked to participate in a study of teacher characteristics, teacher attitudes, and teaching effectiveness. Instructors were asked to identify an undergraduate lecture course in which they agreed to have student ratings collected. They were encouraged to volunteer two or more courses, if possible. Agreement to participate was obtained from 97 instructors teaching a total of 135 lecture-type classes. The demographic characteristics of the 90 instructors who completed the instructor questionnaire are presented in Table 7. Also presented are descriptive statistics for the classes in which ratings were collected.

Measures

Teacher Behaviors Inventory. As described under Study 1, the Teacher Behaviors Inventory (TBI) assesses the frequency of occurrence of 100 low-inference teaching behaviours (see Appendix-A). Frequency ratings were obtained from students enrolled in the 135 classes taught by the 97 participating instructors.

Student Rating Form I. A 16-item rating form was distributed to students to assess more global student

Table 7

Demographic Characteristics of Instructors and Courses
 (Study 2)

Instructor Characteristics

Sex	78 male 12 female
Academic Rank	16 lecturer 24 assistant professor 27 associate professor 23 full professor
Years of Experience	12.4 Mean 8.5 SD
Age	41.4 Mean 8.7 SD

Course Characteristics

Course Level	33 introductory 34 general interest 34 junior honours 34 senior honours
Subject Area	21 Arts 42 Science 72 Social Science
Class Size	74.1 Mean 68.4 SD

attitudes toward the course and instructor (see Appendix B). The first 15 items asked students to indicate their agreement or disagreement on a 7-point Likert scale, with statements concerning course work load and difficulty, impact of course on student thinking or knowledge, future-enrollment plans, and overall evaluation of the course and the instructor. The last item asked students to indicate the percentage of lectures they had attended, with response alternatives ranging from "less than 50%" to "all lectures" in increments of ten percent.

Instructor Self-Rating Questionnaire. Each participating instructor was asked to complete a questionnaire divided into the following four sections: instructor information, course information, classroom teaching style, and attitudes toward teaching (see Appendix C). The first section asked the instructor to provide his or her departmental affiliation and academic rank, sex, age, years of teaching experience, and current teaching load.

The second section of the questionnaire requested information regarding the class in which ratings were being collected: grading format, sources of course content in addition to lectures, percentage of students for whom the course was mandatory, number of times instructor had taught same course in the past, and instructor's current enjoyment of the class. Instructors who had volunteered two or more classes were asked to provide this information for both classes. Additionally, they were asked three questions

which assessed the degree of similarity between the two classes.

The third section of the questionnaire solicited instructor self-ratings on a number of teaching style dimensions. The items referred to the instructor's teaching in general, across various courses. However, instructors were encouraged to provide multiple responses to the items in any case where they felt that the response would vary by class. Only two instructors provided more than one response to any item, and their averaged responses to each were used. Instructors first provided ratings of their classroom teaching on each of the eight scales of the TBI, using a 7-point Likert scale ranging from "much below average" to "much above average". Next, instructors were asked to indicate the approximate percentage of classroom time they spent on each of the six teaching roles discussed by Mann et al. (1970). Although responses were requested as percentages, it was acknowledged that the total might not equal 100. Instructors were next provided with 13 alternative methods which might possibly be used to improve teacher effectiveness (ranging from making inferences from student academic performance to substantial redesign of a course), and were asked to indicate those methods they had tried and their satisfaction with the results. The final item in the third section of the questionnaire presented brief descriptions of "teacher types" corresponding to each of the clusters identified in Study 1. (Note that the

descriptions in the questionnaire are not identified as to cluster, but appear in the following order: Cluster A, Cluster D, Cluster C, Cluster B, Cluster E, Cluster F.) Insofar as possible, descriptions were written to be neutral in desirability but to present the salient points of the profile of each cluster. Instructors were asked to indicate how similar they thought their teaching style was to each of the clusters, using a 7-point Likert scale ranging from "extremely dissimilar" to "extremely similar".

Included in the fourth and final section of the questionnaire were three sets of questions designed to assess the instructor's attitudes toward teaching. One set asked instructors to indicate the amount of time they devoted to teaching in relation to research and service activities, and to rate the importance to them of each of these areas, and their satisfaction with their performance in each area. Instructors were also asked to rate their overall teaching effectiveness relative to other instructors at their rank, on a scale ranging from "much below average" to "much above average". The second set of questions in the fourth section presented a list of 19 potential goals of teaching, and asked instructors to rate, on a 7-point scale, the relative importance of each of these. The final set of questions contained 36 items adapted from the Teaching Styles Inventory developed by Canfield (reproduced in Raines, 1976). The 36 items were grouped into 9 "question sets", each consisting of a common stem, followed by four

response alternatives, such as:

In my opinion, one of the most important responsibilities of a teacher is to:

- a) organize course content and programs in a logical and meaningful order.
- b) motivate students to do the best work of anyone in the class.
- c) get students to cooperate and help one another.
- d) encourage the students to make their own decisions as to what they can accomplish.

For each of the nine question sets, instructors were asked to indicate the extent of their agreement with each of the four response alternatives on a 7-point scale ranging from strongly disagree to strongly agree. These 36 items were scored to yield 12 scale scores. The scale names and definitions are presented in Table 8. Each item (i.e., each response alternative within a set of four) is keyed to a different scale, so that a given question set (stem plus options) relates to four different scales. In the example given above, the options are scored on the following scales: (a) structure-organization, (b) student competition, (c) student-student affiliation, (d) student goal-setting. An initial investigation of the internal consistency of the 12 scale scores identified one question set (i.e., one stem followed by four alternatives) in which the items showed poor correlations with those items which were designed to measure the same constructs in other question sets. It is

Table 8

Definitions of Teaching Styles Inventory Scales

Scale	Definition
Student-Student Affiliation	Student teams, working with other students, etc.
Student-Teacher Affiliation	Knowing the teacher personally
Structure-Organization	Course work and assignments logically organized, sequential, meaningful
Structure-Detail	Specific information on assignments, requirements, etc.
Student Goal Setting	Students setting own objectives; making own decisions
Student Independence	Students working alone and independently
Student Competition	Performance comparisons among students
Teacher Authority	Classroom discipline; knowledgeable teachers
Teacher Attribution	Poor students have the ability; teacher responsible for failure
Shared Attribution	Poor students have some ability; students and teacher share responsibility
Effort Attribution	Poor students not trying hard enough
Ability Attribution	Poor students lack ability

possible that these low correlations were due to an inappropriate revision of the source item in the present research. Raines' (1976) original item asked respondents to consider the "most likely negative consequence" of evaluation of students, and rank order the four consequences provided. For the present research, this was changed in such a way that instructors rated the likelihood of each of the four consequences independently. The focus was thus on the likelihood of the consequences (all of which are negative), rather than on their potential degree of negative impact. As such, the four items in this question set did not appear to be tapping the same dimensions as other items on the same scales. This question set of four items was not scored; as a result, four of the twelve scales are based on two items rather than three items.

Procedure

An official class list was obtained from the university registrar for each of the 135 participating classes. A random sample of students who would be asked to complete teaching evaluation questionnaires was then selected from each class list. For most classes, a sample of 30 students was selected, with 15 asked to complete the Teacher Behaviors Inventory and 15 the Global Student Rating Form. In classes smaller than 40, all students were given questionnaires, with each of the two forms randomly assigned to an equal number of students. In classes larger than 200,

a sample of 40 students was selected, with half of the sample randomly assigned to each questionnaire.

Questionnaires and instructions for their completion were placed in sealed envelopes identified with the student name. For most classes, these were distributed in class by the instructor during the last three weeks of the academic term. For 29 of the 135 classes, the TBI was distributed approximately four weeks earlier. In 5 of these 29 classes, global questionnaires were not distributed. Students were instructed to return completed questionnaires to the researcher by campus mail.

Of all students identified, 1751 (approximately 46%) returned completed questionnaires. The overall return rate for TBI questionnaires was 42%, with number of raters ranging from 1 to 12 per class, and averaging 6.04 per class (809 students in 134 classes). The return rate for the global questionnaire was 51%, with number of respondents ranging from 1 to 16 per class, and averaging 7.30 per class (942 students in 129 classes). - It was impossible to ascertain the number of students who did not receive their questionnaires due to absence from class. Absent students were treated as non-respondents, which caused the return rate to be underestimated by an unknown amount. For six classes, no global rating data were available, and for one class, no completed TBI forms were returned. These classes were treated as missing data in statistical analyses of these variables.

Results

Psychometric Properties of Measures

Teacher Behaviors Inventory. Using the same procedures as in Study 1, interrater reliabilities were computed for the 100 TBI items for 131 classes in which 2 or more students completed the questionnaire. The obtained mean rater reliabilities ranged from .26 to .92, with a mean and median value of .66. There were ten items with reliabilities of less than .50. These items were among the 22 items with the smallest variance across teachers. It appears, then, that the low reliability of these items may be attributable more to homogeneity in teacher behaviours across instructors than to lack of agreement in student ratings within classes.

Internal consistency reliability coefficients for the eight TBI scales were similarly recomputed for the new sample of 134 classes. Table 9 presents the reliability results, along with overall means, standard deviations, and scale intercorrelations. The reliabilities were slightly lower than those obtained with the earlier sample of 298 classes, ranging from .59 to .93.

Student Rating Form I. Interrater reliability coefficients were computed for each of the 16 items of the global student rating form, using the 124 classes in which two or more students responded to the questionnaire. Reliabilities ranged from .51 to .85, with a mean of .71 and

Table 2.
Mean Standard Deviation Internal Consistency
and Intercorrelations of TBI Scales (Study 2)

	Scales							
	DS	ES	EX	OR	RP	RM	SA	UT
DS	83							
ES	57	88						
EX	31	47	84					
OR	55	60	15	85				
RP	51	59	29	49	93			
RM	47	51	15	26	61	61		
SA	43	45	33	28	38	56	83	
UT	13	12	26	28	23	24	27	59
Mean	.38	.03	.13	.14	.20	.25	.16	.34
SD	.74	.59	.67	.72	.61	.53	.69	.63

Notes: (1) Decimals omitted.
 (2) Diagonal elements are alpha reliabilities.

a median of .74, indicating moderate to strong levels of interrater agreement on global aspects of course and instructor effectiveness.

A principal components analysis was conducted to reduce the number, and the amount of intercorrelation, of the global rating variables. Classes were used as the unit of analysis. Three components with eigenvalues greater than 1.0, together accounting for 75.1% of the total variance, were rotated to a VARIMAX criterion. Rotated loadings are presented in Table 10. The first two components can be interpreted as representing an overall impact/interest factor and a workload/difficulty factor, respectively. The third component is more difficult to interpret, but seems to tap a tendency for some courses to be well-attended and useful to students, but not necessarily enjoyable enough to encourage students to enroll in further courses in the same area. Courses high on this factor are likely to be mandatory core courses. This explanation was supported by an observed correlation of .41 between scores on this factor and the instructor's estimate of the percentage of students for whom the course was mandatory.

Instructor Self-Rating Questionnaire. Completed questionnaires were received from 90 of the 97 instructors. The results of the demographic and course characteristic questions are summarized in Table 7. Descriptive statistics for the teacher behaviour and teacher role questions are presented in Table 11. Instructors reported that the

Table 10

Rotated Principal Components of Global Student Rating Form

Rating Item	Component Loading		
	I	II	III
Think differently now	76	03	-06
Learned a lot from course	86	32	-04
Learned a lot from instructor	92	12	14
Course material interesting	76	-03	-39
Improved analysis and problem-solving ability.	57	14	50
Instructor increased my interest	95	-04	-05
Factual knowledge increased	82	18	-13
Lectures increased understanding	91	12	14
Would take another course, this area, this instructor	92	-04	-09
Instructor overall	91	08	19
Course overall	87	27	-19
Heavy work load	03	89	07
Difficult material	-04	84	15
Lot of material covered	27	73	-06
Would take another course, this area, different instructor	35	13	-78
Attendance	21	28	58

Note: Decimals omitted.

greatest proportion of their classroom teaching time (59.7%) was spent in the "teacher-as-expert" role. Results for questions asking instructors to rate their similarity to the six cluster profiles are shown in the middle of the table. As may be seen from the mean ratings for these questions, the attempt to minimize the social desirability of the cluster descriptions was not entirely successful.

Instructors consistently rated Clusters D and E, which have generally low profiles, as being quite dissimilar to their own teaching style. For this reason, subsequent analyses with the cluster similarity ratings used scores which had been standardized across the sample to have a mean of zero and a variance of 1.0. Descriptive statistics for questions assessing the relative importance of Teaching, Research, and Service roles are presented at the bottom of Table 11.

Although instructors reported spending a slightly greater percentage of their time in teaching than in research (43.3% vs. 37.5%), they attached more importance to their research activities than to teaching.

The fourth section of the instructor questionnaire contained a number of attitudinal items, the results of which are summarized in Table 12. The revised items of the Teaching Styles Inventory (TSI; Raines, 1976) were scored according to the scoring key presented by Raines (1976). As described above, one question set of four items was not scored. Thus, four of the twelve scales consist of only two items, while the remaining eight scales contain three items.

Table 11

Descriptive Statistics for Instructor Questionnaire

TBI Scales (1=below average, 7=above average)	Mean	SD
Disclosure	5.13	1.06
Expositional Style	5.00	1.08
Expressiveness	4.91	1.43
Organization	5.28	.93
Rapport	5.29	1.23
Relaxed Manner	5.49	1.17
Speaking Ability	5.08	1.31
Use of Class Time	5.05	1.01

Teaching Roles (in percentage of time)	Mean	SD
Expert	59.66	23.62
Formal Authority	8.51	5.48
Socializing Agent	4.45	5.26
Facilitator	10.93	12.97
Ego Ideal	9.12	10.35
Person	8.95	6.70

Similarity to Teaching Style Clusters (1=extremely dissimilar, 7=extremely similar)	Mean	SD
Cluster A	4.88	1.39
Cluster B	4.37	1.26
Cluster C	3.36	1.67
Cluster D	2.57	1.40
Cluster E	5.08	1.28
Cluster F	1.70	1.17

Time Spent in Teaching, Research, and Service (in percentage of time)	Mean	SD
Teaching	43.32	19.71
Research	37.50	21.70
Service	14.30	11.00

Personal Importance of Teaching, Research, and Service (1=not at all important, 7=extremely important)	Mean	SD
Teaching	5.29	1.54
Research	6.17	1.25
Service	3.62	1.75

Note: The order of variables differs from that of the questionnaire in Appendix C.

Table 12

Descriptive Statistics for Teacher Attitude Variables on
Instructor Questionnaire

	Mean	SD	No. of Items	Alpha
Teaching Styles Inventory Scales				
Student-Student Affiliation	9.12	2.55	2	.70
Student-Teacher Affiliation	14.89	2.78	3	.49
Student Goal-setting	7.19	2.55	2	.58
Student Independence	12.25	3.45	3	.59
Structure-Organization	11.99	1.65	2	.74
Structure-Detail	15.09	3.08	3	.58
Student Competition	9.09	2.09	2	.05
Teacher Authority	13.92	2.75	3	.32
Teacher Attribution	11.56	3.42	3	.65
Shared Attribution	11.81	2.64	3	.19
Effort Attribution	11.84	2.96	3	.33
Ability Attribution	12.68	3.71	3	.56
Educational Goals				
Stimulating student thinking	6.30	.96		
Factual information	5.49	1.35		
Enjoyment	4.16	1.60		
Independent study	5.32	1.48		
Technical skills	3.89	1.70		
Problem solving ability	5.68	1.27		
Understanding parts-whole	5.59	1.25		
Student confidence with subject	5.10	1.42		
Transmitting own interest	5.13	1.44		
Ability to argue logically	5.13	1.35		
Ability to use subject knowledge	5.63	1.05		
Ability to use prior knowledge	5.30	1.30		
Ability to generalize	5.03	1.16		
Technical vocabulary	3.98	1.61		
Interest in subject	5.43	1.16		
Knowledge of reference materials	4.27	1.76		
Oral communication skills	3.41	1.75		
Written communication skills	4.86	1.78		
Particular theory or method	4.13	1.61		

Descriptive statistics for the 12 scales of the TSI are reported at the top of Table 12. The highest scores were associated with conditions that center on the teacher: structure-organization, structure-detail, and student-teacher affiliation. It would appear that the instructors in this sample attached greater importance to the teacher's responsibility for providing structure to students than for encouraging students to set their own individual goals (student goal-setting). Table 12 also presents internal consistency estimates (coefficient alpha) for the twelve TSI scales. Five of the scales have coefficients below .50, indicating a lack of consistency among instructor responses to the various items on each of these scales. Internal consistency measures of the magnitude reported in Table 12 are not uncommon for 2- or 3-item scales. However, scales with low internal consistency are unlikely to show significant relationships with other variables.

The bottom half of Table 12 shows descriptive statistics for instructor ratings of the importance of 19 potential educational goals. It may be noted that stimulating student thinking was rated as the most important goal. The lowest ratings were given to oral communication skills, technical skills, and technical vocabulary.

To reduce the number of attitudinal variables to fewer but more reliable composites, the 12 Teaching Styles Inventory scales and 19 educational goal ratings were

subjected to a principal components analysis. Three components accounting for a total of 39.7% of the variance were rotated to a VARIMAX criterion. Rotated loadings are presented in Table 13. The three factors were interpreted and named through an examination of the variables with salient loadings. The first component, Importance of Enjoyment, is marked by four TSI scales, namely, student-student affiliation, student-teacher affiliation, student goal-setting, and student independence; as well as by educational goals of student enjoyment, interest, and self-confidence. This factor reflects the instructor's desire to foster good interpersonal relations among students and between teacher and students. It also reflects a desire to encourage students in making their own decisions, in becoming more independent, and in enjoying the class. The second component, Importance of Skills, is marked by goals such as stimulating student thinking, fostering independent study, and developing students' abilities to argue logically, use their prior knowledge, and communicate orally and in writing. The third component, Importance of Knowledge, characterizes instructors who emphasize specific subject matter content, as opposed to generalizable skills. It has high loadings for the TSI structure-organization and structure-detail scales, and for the goals of transmitting factual information and technical vocabulary. Scale scores were computed for each instructor by summing the items with loadings greater than .40 in the rotated solution.

Table 13

Rotated Principal Components of Teacher Attitude Variables

	Components		
	I	II	III
<u>Teaching Style Inventory Scales</u>			
Student-Student Affiliation	61	35	-04
Structure-Organization	05	-09	71
Student Goal-Setting	39	23	-49
Student Competition	25	-07	-01
Student-Teacher Affiliation	67	-08	17
Structure-Detail	21	-03	70
Student Independence	55	31	-51
Teacher Authority	-14	27	51
Teacher Attribution	73	-12	-20
Shared Attribution	46	-11	-02
Effort Attribution	-21	-08	11
Ability Attribution	-41	06	42
<u>Educational Goals</u>			
Stimulating student thinking	06	67	-05
Factual information	00	-18	67
Enjoyment	56	02	02
Independent study	04	67	-10
Technical skills	-03	13	25
Problem solving ability	01	16	15
Understanding parts-whole	34	50	34
Student self-confidence with subject	57	24	15
Transmitting own interest	55	21	17
Ability to argue logically	-23	53	-10
Ability to use subject knowledge	08	40	35
Ability to use prior knowledge	-07	77	11
Ability to generalize	-04	58	33
Technical vocabulary	40	30	52
Interest in subject	49	51	21
Knowledge of reference materials	43	40	05
Oral communication skills	14	70	01
Written communication skills	19	66	-10
Particular theory or method	36	23	-36

Note: Decimals omitted.

Classification of Teaching Styles

The TBI profile of each instructor/class combination was classified into one of the six styles identified in Study 1. Euclidean distances were computed between each individual and the centroids of the six clusters from the earlier study. Each of the eight scales was standardized with respect to the Study 1 sample, so that large-variance scales would not unduly affect classification. Individuals were classified into the cluster whose centroid was closest to their own TBI profiles.

The proportions of instructors classified into each of the clusters in Study 2 differed from those of Study 1. The numbers of cases and the mean profiles of the groups are presented in Figure 2. The number of low-elevation profiles decreased from Study 1 to Study 2, and the number of high-elevation profiles increased. Specifically, Mumbler (Cluster F) were much less common in the Study 2 sample, including only 3 cases. Conversely, Effective Lecturers (Cluster E) were somewhat more common in the Study 2 sample, comprising 35% of cases as compared to 24% in the earlier sample.

Relationship of Cluster Membership to Other Variables

The following sections report relationships between teaching style, as defined by cluster membership, and six other types of educational variables: instructor demographic characteristics, class characteristics,

Figure 2
Cluster Profiles (Study 2 Sample)

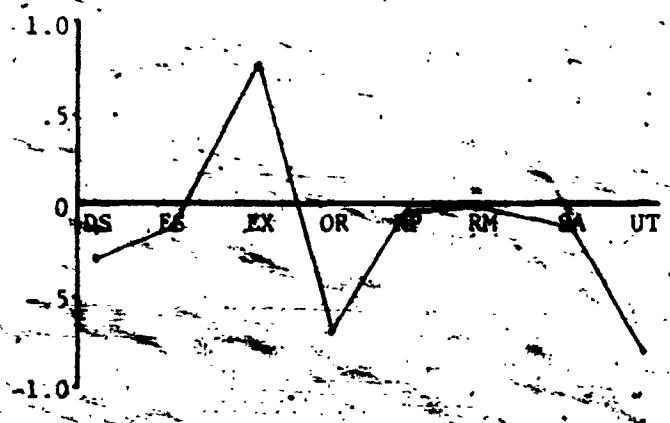
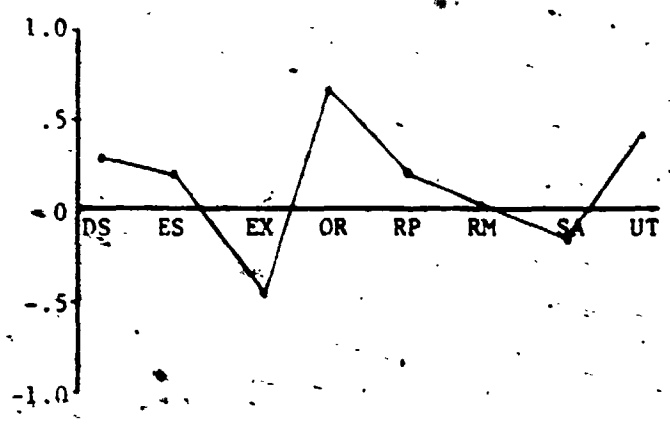
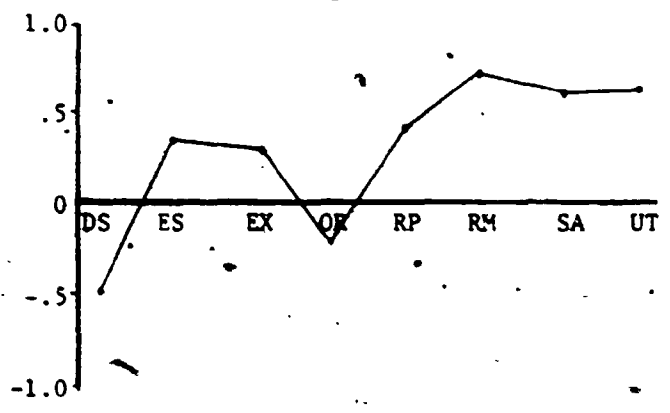
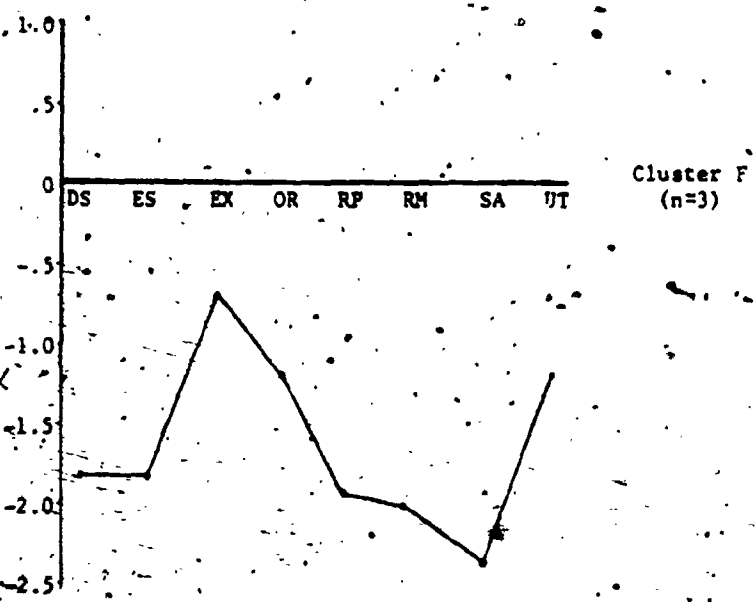
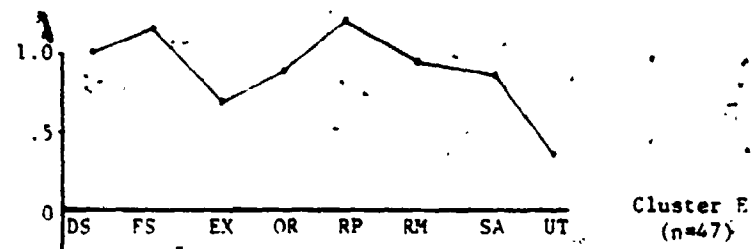
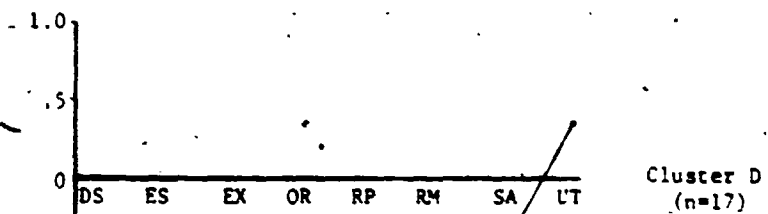


Figure 2 (cont'd.)



instructor attitudes, instructor roles, instructor self-ratings, and global student ratings. As part of the process of external validation (Skinner, 1981), these analyses were designed to examine the meaningfulness of the teaching style typology developed in Study 1, with reference to other relevant variables. The unit of analysis was the class section. Thus, instructors who taught more than one class in Study 2 were included more than once in the analysis. Parallel analyses were performed with such duplicate instructors removed but no substantial differences in results were noted. Thus results are presented here for the complete data set.

Instructor Demographic Characteristics. Instructors falling into each of the six teaching style clusters were compared on three categorical and five continuous instructor variables: department, academic rank, sex, age, years of teaching experience, and three measures of current teaching load.

The first categorical instructor variable to be considered, departmental affiliation, was collapsed into three faculty areas: Arts, Science, and Social Science, defined as in Study 1. No statistically significant relationship was found between cluster membership and faculty affiliation, $\chi^2(10) = 11.632$. This result fails to replicate the significant relationship found with faculty affiliation in Study 1. The reason for this is not entirely clear. It could be a symptom of insufficient power due to

the smaller sample size in the present study.

Alternatively, it might reflect differences between a sample composed entirely of volunteers (Study 2) and one including instructors who were recruited ($n = 175$ of Study 1).

The second instructor variable to be considered, academic rank, also showed no significant relationship to cluster, $\chi^2(15) = 14.021$. Similarly, sex of instructor was not significantly related to cluster, $\chi^2(5) = 5.048$.

In the analysis of continuous or quantitative instructor variables, the six teaching style clusters were treated as six levels of a single-factor independent variable in a multivariate analysis of variance, with instructor variables serving as dependent measures. The first analysis examined age and years of teaching experience. Cluster differences in mean age and mean instructor experience were not statistically significant, $F(10, 242) = 1.26$. The second analysis examined three measures of the instructor's current teaching load, namely total number of students taught, total number of hours of teaching per week, and total number of different courses taught. Cluster differences in teaching load of the instructor were found to be statistically significant, $F(15, 360) = 2.74, p < .001$. Examination of univariate results showed that only total number of students taught differed significantly across clusters, univariate $F(5, 120) = 6.38, p < .001$. Total number of hours of teaching per week and total number of courses taught did not differ significantly

across clusters. $E(5, 120) = 1.99$ and $E(5, 120) = 1.24$, respectively. Further examination of cluster means revealed that the significant result for number of students taught was entirely due to one class reported to have over 600 students enrolled.

In summary, only one statistically significant relationship was found between instructor teaching styles and instructor demographic characteristics, and it seems to be attributable more to one outlier in class size than to any meaningful relationship. Although a significant relationship was found in Study 1 between subject area and teaching style, it did not appear in this new data set. Thus, there is little evidence that demographic characteristics of instructors are related to their displayed teaching style.

Class Characteristics. In the next set of analyses, which focused on course characteristics, cluster membership was first compared across levels of courses. As in Study 1, courses were classified into four levels: Introductory, General Interest, Junior Honours, and Senior Honours. Table 14 shows the number of instructor/class combinations at each course level assigned to each of the six teaching style clusters. The relationship between teaching style (i.e., cluster membership) and course level was statistically significant, $\chi^2(15) = 28.304, p < .02$. An examination of the entries in Table 14 indicates that instructors in introductory courses were more likely than other instructors

2



2

MICROCOPY RESOLUTION TEST CHART
NBS - 1010a
(ANSI and ISO TEST CHART No. 2)

1.0		2.8	2.5
1.1		3.2	2.2
1.25		3.6	2.0
		4.0	1.8
		4.5	
		5.0	
		5.6	
		6.3	
		7.1	
		8.0	
		9.0	
		10	
		11.2	
		12.5	
		14	
		16	

Table 14

Relationship between Teaching Style and Course Level(Study 2)

Course Levels	Teaching Style Clusters						Total
	A	B	C	D	E	F	
Introductory	5	2	10	2	14	0	33
General	6	7	7	3	9	1	33
Junior Honours	4	6	2	8	14	0	34
Senior Honours	11	6	1	4	10	2	34
Total	26	21	20	17	47	3	134

Note: Table entries are numbers of classes,

to be Entertaining but Unstructured (Cluster C) and less likely to be the opposite, Structured but Boring (Cluster B). In Study 1, Speech Givers (Cluster A) were found to be over-represented at the introductory level. This earlier result is not entirely inconsistent with the present finding, given the similar pattern of Expressiveness exceeding Organization which is displayed both by Speech Givers and by the Entertaining but Unstructured group. In second-year honours courses, the current sample had fewer Entertaining but Unstructured instructors than would be expected, consistent with Study 1 results. Also, there were slightly more Ineffective Lecturers (Cluster D) at the junior honours level than would be expected from the marginal frequencies. In senior honours courses, there were more Speech Givers and fewer Entertaining but Unstructured lecturers than would be predicted from marginal frequencies.

Clusters of teachers were next compared on three continuous variables describing the class: class size, percentage of students for whom the course was mandatory, and percentage of course content which was available only through instructor lectures. In order to control for any effect due to course level, a separate two-way analysis of variance, with cluster and course level as independent variables, was conducted for each of the three aforementioned dependent variables. In all three ANOVA's, there was no statistically significant main effect due to cluster and no significant interaction of cluster with

course level, all F 's < 1.0 .

In summary, among the class characteristics studied here, only the level of the course had a significant relationship with cluster membership. The nature of this relationship differed somewhat from that found in Study 1, although some consistent patterns were noted. Although the causes of this relationship are as yet unexplained, it appears that teaching style and course level are in fact meaningfully related.

Instructor Attitudes: Clusters of teachers were next compared on the three instructor attitude scales derived from factor analysis of instructor responses to TSI items and educational goals. These scales, described in Table 13, reflect the instructor's values and objectives with respect to the educational process. Mean differences across teaching styles (clusters) were not statistically significant for any of the three attitude dimensions: Importance of Enjoyment, $F < 1.0$, Importance of Skills $F < 1.0$, and Importance of Knowledge, $F = 1.09$. These results indicate that instructors displaying different teaching styles did not differ significantly in their attitudes regarding educational goals.

Instructor Roles. Instructors were asked to rate the percentage of classroom time spent in various teaching roles derived from the work of Mann et al. (1970). Instructors also rated the percentage of their total working time spent in the three academic roles of teaching, research, and

service. A multivariate analysis of variance was performed with cluster as the independent variable and ratings of classroom and academic roles as dependent variables. The MANOVA indicated that there were significant differences across the six teaching style clusters for all dependent variables combined, $F = 1.42$, $p < .05$. Although none of the univariate analyses attained statistical significance, the greatest differences were in the amounts of time spent in Formal Authority roles, Facilitator roles, research activities, and service activities. Inspection of structure coefficients suggested that differences across clusters could be described by one linear function of the dependent variables. The significant effect was due primarily to the three members of Cluster F (Mumblers), who reported that they spend more time than other instructors as facilitators and in service activities, and less time in formal authority roles and research activities.

Instructor Self-Ratings. Instructors had been asked to rate the degree of similarity of their own teaching to written descriptions of the six teaching style profiles. The validity of teaching style clusters would be supported by negative correlations between similarity ratings and the observed dissimilarity (i.e., Euclidean distance) between the instructor's TBI profile and the clusters. In computing correlations between similarity ratings and Euclidean distances for the six clusters, duplicate representation of instructors who taught more than one class was eliminated by

arbitrarily selecting only one class (the class at the higher level) for such instructors. Observed correlations were as follows: for Cluster A, Speech Givers, $r = .06$; for Cluster B, Structured but Boring, $r = .03$; for Cluster C, Entertaining but Unstructured, $r = .09$; for Cluster D, Ineffective lecturers, $r = -.23$; for Cluster E, Effective Lecturers, $r = -.05$; for Cluster F, Mumlbers, $r = -.30$. Correlations were in the predicted negative direction only for the three teaching styles whose mean TBI profiles showed pronounced non-zero elevations (i.e., Clusters D, E, and F), and only the low-elevation profiles, Cluster F (Mumlbers) and Cluster D (Ineffective Lecturers) showed correlations significantly different from zero. These results could be due to a number of factors. The social desirability of the written descriptions, which to a large extent may reflect the desirability of the various styles, could have interfered with the similarity ratings. Alternatively, instructors may not be particularly accurate at describing their own classroom behaviour, except in very general terms such as above or below average in overall effectiveness. In other words, instructors may be more sensitive to profile elevation than to profile shape. It might be possible to disentangle these effects by examining instructor self-ratings on the eight TBI scales and using these scores to compute a perceived similarity measure.

Global Student Ratings. In a final set of analyses, the six teacher clusters were compared with respect to mean

scores on the three global student rating factors identified through a principal components analysis of the global student rating form. The three factors were interpreted as: overall impact/interest, workload/difficulty, and attendance/utility. The first factor was composed largely of items reflecting instructor quality, impact, and overall effectiveness. For each factor, an analysis of variance was conducted to examine the effects of teaching style (cluster), course level, and faculty affiliation on factor scores for instructor/class units. To reduce the number of missing cells, the three instructors in Cluster F (Mumblers) were combined with those in Cluster D (Ineffective Lecturers) for the purposes of these analyses. The results of the three analyses of variance are presented in Table 15.

For the first factor (Overall Impact/Interest), there was a significant main effect of teaching style, $F = 14.654$, $p < .001$. Ineffective Lecturers (Cluster D/F) had a mean effectiveness rating (-1.07) that was a full standard deviation lower than the grand mean. Speech Givers (Cluster A) and Effective Lecturers (Cluster E) had the highest mean effectiveness ratings (.39 and .43, respectively). Given the substantial body of research literature supporting a link between classroom behaviours and overall teaching effectiveness, these results are consistent with the high and low elevation levels of TBI profiles for Cluster E and Cluster D/F respectively. The only significant interaction in the ANOVA was between course level and faculty.

Table 15

Style, Level, and Faculty Effects on Global Student Rating Factors

Impact/Effectiveness Factor

Source of Variance	SS	DF	MS	F
Main Effects				
Level	1.63	3	.54	.86
Style	6.98	4	9.24	14.65*
Faculty	2.80	2	1.40	2.22
Interactions				
Level X Style	5.94	12	.50	.79
Level X Faculty	9.74	6	1.62	2.57*
Style X Faculty	9.23	8	1.15	1.83
Level X Style X Faculty	5.93	8	.74	1.18
Residual	52.99	84	.63	
Total	127.46	127	1.00	

* $p < .05$

Table 15 (cont'd.)

Workload/Difficulty Factor

Source of Variance	SS	DF	MS	F
Main Effects				
Level	4.16	3	1.39	1.87
Style	14.87	4	3.72	5.00**
Faculty	14.99	2	7.50	10.09**
Interactions				
Level X Style	13.84	12	1.15	1.55
Level X Faculty	10.17	6	1.70	2.28*
Style X Faculty	2.98	8	.37	.50
Level X Style X Faculty	3.20	8	.40	.54
Residual	62.43	84	.74	
Total	127.99	127	1.01	

* $p < .05$ ** $p < .01$

Table 15 (cont'd.)

Attendance/Utility Factor

Source of Variance	SS	DF	MS	F
Main Effects				
Level	6.87	3	2.29	2.64
Style	3.00	4	.75	.87
Faculty	2.36	7	1.18	1.36
Interactions				
Level X Style	9.71	12	.81	.93
Level X Faculty	12.65	6	2.11	2.43*
Style X Faculty	9.12	8	1.14	1.32
Level X Style X Faculty	10.22	8	1.28	1.47
Residual	72.84	84	.87	
Total	127.99	127	1.01	

* $p < .05$

$F = 2.573$, $p < .05$. This effect will not be interpreted here, as it is tangential to the present discussion. The interaction between teaching style and faculty approached statistical significance, $F = 1.830$, $p < .10$, reflecting a tendency for the Entertaining but Unstructured style to be particularly ineffective ($M = -1.47$) in the Arts faculty but slightly above average in effectiveness ($M = .20$) in the Social Science faculty.

The second factor (Difficulty/Workload) showed significant main effects of both faculty and teaching style, $F = 10.085$ and 5.001 respectively, $p < .001$ in both cases, plus a significant interaction of faculty and course level $F = 2.280$, $p < .05$. As noted above, it is the main effect of teaching style that is most relevant here. An examination of cell means revealed that students in classes taught with a Cluster D/F style (Ineffective Lecturers) perceived the course material to be most difficult and the workload heaviest ($M = .69$), whereas students taught with an Entertaining but Unstructured (Cluster C) style showed the lowest course difficulty ratings of any group ($M = -.42$).

There were no significant main or interaction effects involving teaching style in the analysis of variance of the third global student rating factor (Attendance/Utility).

Informational Value of Clusters versus Scales

A multiple regression analysis was conducted to determine whether a cluster (i.e., teaching style)

description provides any information over and above that available from simple teaching behaviour scale scores for purposes of predicting student ratings of overall teaching effectiveness. A separate hierarchical regression was carried out for each of the three global student rating factors. In each analysis, the predictor variables were: (1) style classification, consisting of 6 dichotomous, dummy-coded variables, plus distances between the instructor's profile and each of the six cluster centroids; (2) the eight TBI scale scores; and (3) course level. All obtained multiple Rs were adjusted for estimated shrinkage. For the first factor, Overall Impact/Interest, the TBI scale scores and course level yielded a multiple R of .581, accounting for 34% of the variance. The teaching style variables alone yielded a similar multiple R of .575, accounting for 33% of the variance. In a hierarchical regression, the multiple R for both sets of variables combined increased to .61, but neither set of variables added a significant amount of information to that predicted by the other set. $F(10, 115) = 1.46$ for teaching style, and $F(8, 115) = 1.38$ for TBI scale scores. For the second global rating factor (Workload/Difficulty), the variable set consisting of TBI scales and course level yielded a multiple R of .38, whereas the teaching style variables yielded a multiple R of .32. The two sets of variables combined yielded a multiple R of .42. When the teaching style variables were entered first, the multiple R for both sets

combined was .35. The teaching style variables did not add a statistically significant amount of information to the regression equation $F(10, 115) = 1.40$. However, TBI scale scores did contribute significantly, $F(8, 115) = 2.70$, $p < .01$. For the third global rating factor (Attendance/Utility), all regressions yielded non-significant F -ratios of less than 1.0 for amount of variance explained by both cluster and scale score variables. These analyses are therefore not reported or interpreted here.

The results of these analyses indicate that, for the prediction of global ratings of teaching effectiveness, a description based on the teaching style typology developed in Study 1 provides no more information than that already available to us through simple scale scores. By considering the course level and individual TBI scale scores, we can predict overall effectiveness ratings reasonably well ($R = .58$). Consideration of course level plus the six teaching styles or clusters yields approximately the same level of predictive power for overall impact measures. In other words, although each method of measurement has predictive validity, neither displays significant incremental validity.

Summary

Teaching style as defined by the typology developed in Study 1 showed relatively weak relationships with most of the educational variables included in Study 2. Significant

relationships found in Study 1 between teaching style and subject area were not replicated in Study 2. Relationships with instructor demographic characteristics, self-reported roles, educational values, and self-ratings of teaching style were statistically significant only in isolated cases.

As in Study 1, a significant relationship was found between teaching style and level of course. Generally, introductory courses were more likely to be taught with an emphasis on expressiveness rather than organization. This tendency reversed itself in honours level courses, where the "free-wheeling" Entertaining but Unstructured approach was less likely to be found.

Global student ratings, the only "product" measures, included in this research, were significantly different for the six teaching styles. The pattern of these differences for the overall impact/interest factor reinforces the importance of profile elevation. High profile clusters (e.g., Cluster E) were associated with high overall ratings, whereas low profile clusters (e.g., Cluster D and Cluster F) were associated with low overall ratings. For student ratings of workload/difficulty, an Entertaining but Unstructured teaching style was associated with low ratings of difficulty. This finding exemplifies the lack of causal explanation inherent in this research. It cannot be determined whether students perceive the subject matter to be easy because the instructor has a "freewheeling" approach, or whether the instructor takes this approach

because the subject matter is easy.

In the prediction of global impact/interest ratings, the typology approach utilized here fares no better, but no worse, than an approach relying solely on scale scores. • The examination of patterns of behaviour, i.e., teaching styles, may be warranted if they are conceptually easier to grasp. • The results of this study indicate that we will neither gain nor lose any predictive information related to overall teaching effectiveness.

STUDY 3

The purpose of Study 3 was to evaluate the stability of an individual's identified teaching style across courses and across years. Research on student ratings at the postsecondary level suggests that an instructor's ratings are highly stable across years for the same course, and at least moderately stable across different courses. Marsh (1982) and Marsh and Hocevar (1984) reported generalizability analyses on 9 student evaluation factors: learning/value, enthusiasm, organization/clarity, group interaction, individual rapport, breadth of coverage, examinations/grading, assignments, and workload/difficulty. They concluded that only about 20% of the reliable variance in teacher ratings was course-specific. Generalizable instructor effects were determined to be the major determinant of student ratings of teaching.

Smith (1979) emphasized the importance of distinguishing between items designed to assess instructor characteristics and items designed to assess course characteristics. This distinction has not been observed by Marsh and his colleagues. Smith pointed out that it is somewhat difficult, with instructor and course items mixed together, to make statements related specifically to the stability of perceived instructor characteristics across different courses. In the present study, however, the description of teaching styles was based solely upon

specific classroom behaviours exhibited by the instructor. Thus the examination of consistency across classes was an examination of instructor characteristics only.

Murray, Rushton and Paunonen (1986) reported that correlations for global student ratings of instructor effectiveness were quite high (.70 to .80) across years for the same course, but somewhat lower (.40 to .50) across different courses or types of courses taught in the same year. Additionally, their results showed that the between-course correlation decreased in magnitude with greater differences in course level or course type.

These studies, and others, have focused on more global, evaluative ratings made by students, rather than on specific teaching behaviours. Although research evidence suggests that global effectiveness ratings are likely to be determined by, or at least related to, specific teacher behaviours (e.g., Murray, 1983a), evidence for consistency of global ratings is not necessarily evidence for consistency of specific teacher behaviours. Shavelson and Atwood (1977) reviewed eleven studies of the stability of teacher behaviour across occasions in elementary and secondary schools. The results indicate that the duration and frequency of lecturing behaviour is stable across occasions, but the stability of other specific behaviours ranges from highly stable to highly unstable.

An additional limitation of previous studies of the stability of teacher behaviour and/or student ratings is

that they have focused solely on correlations of teachers across occasions. As Rogosa, Floden and Willett (1984) point out, this approach answers only the question: are individual differences in ratings (or behaviours) consistent across occasions? This is a valid question, particularly with respect to the use of rating or behavioural measures in a norm-referenced format for purposes of summative teacher evaluation or process-product research. These are the most common applications of student ratings at the postsecondary level. However, other applications of these measures, such as formative evaluation for teacher improvement or descriptive research of classroom processes, may require a somewhat different conceptualization of "consistency". The question to be answered in these situations is: are an individual teacher's ratings (or behaviours) consistent across occasions in absolute terms? Study 3 was designed to examine the stability of teacher behaviour in both relative and absolute terms.

More generally, the major goal of Study 3 was to assess one final aspect of what Skinner (1981) calls external validation of a typology. This study examined the stability of an individual's teaching style across years and across courses. Relationships between individual differences in teaching style consistency and other relevant instructor and course characteristics were also investigated.

Method

Subjects

Sample A. Instructors approached by letter in Study 2 had been encouraged to volunteer more than one course for research purposes. Of the 97 instructors who agreed to participate, 35 consented to have ratings collected in two undergraduate courses. This group of 35 instructors teaching 70 classes formed Sample A in Study 3.

Sample B. Of the 97 instructors who agreed to participate in Study 2, 25 had also participated at least once in earlier research utilizing the TBI. These instructors, therefore, have already served as subjects in both Study 1 and Study 2 of the present research. Ten of these twenty-five instructors were included in Sample A: nine of them taught two classes in Study 2 plus one in Study 1, and one instructor taught two classes in Study 2 plus two in Study 1. This group of 25 instructors teaching 61 classes constituted Sample B in Study 3.

Measures and Data Collection

For Sample A, all measures described above for Study 2 were available. For Sample B, only TBI data and information regarding the department and level of the course were available for both Study 1 and Study 2 courses. Methods of collecting and scoring data were as described above.

Results

Stability of Cluster Membership

Instructors were classified into one of the six teaching styles, separately for each of the two classes they taught. In Sample A, the crossed classifications represent the degree to which an instructor's teaching style is identified similarly in two different courses taught in the same year. Table 16 contains the crossed classifications for the 35 Sample A instructors. It may be noted that 18 of the 35 instructors were classified into the same teaching style in both courses, and the relationship between classifications in the two courses was statistically significant, $\chi^2(20) = 42.92, p < .005$.

Instructors whose teaching style, as defined by cluster classification, was constant across the two courses ($n = 18$) were compared to instructors whose teaching style differed across their two courses ($n = 17$). Chi-square tests or independent-groups t-tests revealed no statistically significant differences between these two groups in academic rank, sex, age, years of experience, teaching load (hours, students, courses), or instructor ratings of the degree of similarity of the two courses or of their teaching styles in the two courses. However, a statistically significant relationship was found between similarity of cluster membership and similarity in level of the courses.

Instructors teaching two courses at the same level (e.g.,

Table 16

Replication of Teaching Styles Across Classes Taught in Same Year

		Class B Teaching Style						
		A	B	C	D	E	F	Total
Class A Teaching Style	A	1	1	1	2	0	0	5
	B	0	3	0	2	0	1	6
	C	3	0	3	2	0	0	8
	D	2	0	0	9	0	0	11
	E	0	0	0	0	0	1	1
	F	0	2	0	0	0	2	4
Total		6	6	4	15	0	4	35

Note: Table entries are numbers of classes.

both Introductory) were more likely to be classified into the same teaching style cluster in the two classes than instructors teaching at different levels. $\chi^2(1) = 6.69$, $p < .01$. This relationship is displayed in Table 17.

Accordingly, subsequent analyses in both Sample A and Sample B were performed separately for "same-level" and "different-level" pairs of classes.

Crossed classifications for Sample A teachers were investigated separately for pairs of courses at the same and at different levels, taught within the same year. For pairs of courses at different levels ($n = 20$), the between-course stability of teaching style classification was not statistically significant. $\chi^2(16) = 17.22$. The degree of stability can be assessed by Cohen's (1968) weighted kappa coefficient. Kappa is generally a coefficient of agreement between two sets of categorical ratings or two classifications. Weighted kappa allows a differential weighting of disagreements. In other words, certain disagreements can be scaled as more "serious" than others. In the present research, disagreements were weighted by the distance between the corresponding cluster centroids. The distance between Cluster E (above average) and Cluster F (very low) profiles was the largest intercluster distance ($d = 6.57$) in this study. Therefore a switch in styles from E to F was considered to be most "serious", and was weighted most heavily. Kappa ranges from -1.0 (maximum disagreement) to +1.0 (complete agreement) and can be interpreted in a

Table 17
Similarity in Teaching Style Compared to Similarity in
Course Level

	Same Style	Different Style	Total
Same Level	12	3	15
Different Level	6	14	20
Total	18	17	35

Note: Table entries are numbers of classes.

manner similar to a reliability coefficient. For pairs of courses at different levels taught within the same year, the obtained k was .193, indicating very little stability and supporting the chi-square result. For pairs of courses at the same level and within the same year ($n = 15$), teaching styles in the two courses were found to be significantly related, $\chi^2(16) = 42.05$, $p < .001$, and the kappa coefficient indicated a very high level of agreement ($k = .772$).

For instructors in Sample B, analyses similar to the above were conducted for pairs of courses at the same and at different levels taught in two different years. The time span for these pairs of courses was between 6 and 8 years. For pairs at different levels ($n = 17$), the relationship between teaching style classifications was not statistically significant, $\chi^2(12) = 15.42$, and the kappa coefficient indicated a very low level of agreement ($k = .155$). For pairs of courses at the same level ($n = 18$), the between-course stability of teaching styles was found to be clearly significant, $\chi^2(25) = 44.57$, $p < .01$, and the kappa coefficient showed moderate agreement ($k = .376$).

To summarize, the results obtained from between-course comparison of teaching style classifications indicate that the highest level of agreement is found for instructors teaching two courses at the same level in the same year. Less agreement is obtained for instructors teaching two courses at the same level in different years. Very little agreement is obtained for instructors teaching two courses

at different levels in the same year or in different years. Thus, instructors do appear to behave differently in different courses, but on the other hand they show some consistency in behaviour when teaching similar courses separated by as much as eight years.

Stability of Profiles

The stability of teaching styles across occasions can also be examined in terms of TBI profiles. For example, the two profiles for each instructor can be compared in terms of the distance between each profile and the mean score profile for a given teaching style. This assesses the degree to which the similarity of the instructor to each of the six styles or clusters generalizes across classes and/or across years. The correlations of these distances across classes are presented in Table 18. There does not appear to be a consistent pattern in the size of these correlations across the four categories of class combinations (same vs. different year X same vs. different level). It is clear, however, that the correlations are higher for Clusters D, E, and F than for the other clusters. These three clusters had profile elevations that were considerably above or below average. It appears that the elevation component of the distance measure was more consistent across years and/or classes than was the shape component.

A second way of comparing instructors' TBI profiles across occasions is in terms of shape, scatter, and

Table 18

Correlation Coefficients Assessing Stability of ProfileDistance to Cluster Centroids

		Same Year		Different Years	
		Same	Different	Same	Different
		Level	Level	Level	Level
		(n=15)	(n=20)	(n=18)	(n=17)
Distance Computed to this Cluster	A	.27	.33	.22	.15
	B	.04	.05	.30	.09
	C	.02	.11	.02	.08
	D	.50	.49	.48	.13
	E	.42	.45	.48	.40
	F	.31	.44	.49	.49

elevation. A within-person correlation computed between two profiles assesses similarity of shape. A corresponding covariance measure assesses similarity of both shape and scatter, whereas a distance measure (e.g., Euclidean distance) assesses similarity of shape, scatter, and elevation. For each of the 70 pairs of instructor profiles reported above, these three measures were computed. Table 19 presents the means of these coefficients for the four groups of classes representing all possible combinations of same vs. different year X same vs. different course level. Euclidean distances did not display a clear trend across these four categories. Covariances, on the other hand, showed a tendency to decrease as courses or years differed. Correlations were in the same decreasing order as covariances, but the differences among the first three groups of courses (i.e., same year and/or same level) were negligible. These results seem to indicate that changes in profile shape might occur primarily with the most extreme changes in teaching situation. Profile shape and scatter in combination may be more sensitive to subtle changes in the teaching situation.

Stability of Individual Differences in Scores

A third way of assessing stability of teaching behaviours across situations is to compare individual differences in these behaviours in different situations. This can be accomplished by correlating teaching behaviour

Table 19

Mean Values of Profile Similarity Measures

	Same Year		Different Years	
	Same	Different	Same	Different
	Level (n=15)	Level (n=20)	Level (n=18)	Level (n=17)
Correlation	.55	.54	.50	.30
Covariance	.38	.31	.27	.13
Euclidean distance	2.69	2.61	2.58	2.68

scale scores for all instructors at two different points in time, or in two different courses. Stability correlations for each of the TBI scale scores are presented in Table 20. As with the distance measures, there was no clear trend for greater stability of scale scores to be found in more similar teaching situations. There were, however, differences in the mean stability of the various TBI scales. Expressiveness seemed to generalize extremely well across course levels and across years, whereas Use of Class Time was stable within years but not across years.

Summary

The results of Study 3 suggest that teaching style, as defined by classification into one of the six clusters, is quite stable over courses at the same level taught in the same year. Even over a time period as long as eight years, there is still moderate stability in teaching style classification, provided that the courses are at the same level. However, for courses at different levels, taught either in the same or in different years, the stability of teaching style drops off dramatically. Going beyond categorical data and examining the similarity of each instructor to the six "types", it was observed that an instructor's tendency to be above or below the average is stable across courses and across years. However, the degree to which an instructor conformed to the Speech Giver or Structured but Boring or Entertaining but Unstructured

Table 20

Correlation Coefficients Assessing Stability of TBI Scale Scores

	Same Year		Different Years	
	Same Level (n=15)	Different Level (n=20)	Same Level (n=18)	Different Level (n=17)
Disclosure	.49	.57	.56	.30
Expositional Style	.60	.66	.50	.51
Expressiveness	.76	.81	.85	.71
Organization	.48	.43	.68	.31
Rapport	.68	.29	.48	.71
Relaxed Manner	.27	.43	.37	.24
Speaking Ability	.33	.60	.52	.05
Use of Class Time	.49	.35	.01	.14

styles showed little or no stability across courses or years. Although the within-instructor correlations suggested that there is some stability in profile shape across years and/or course levels ($r = .50$ to $.55$), apparently this is insufficient to ensure that the instructor's profile remains in a similar position relative to these three teaching styles. TBI scale scores, on the other hand, were generally stable across course levels and years.

The sample sizes employed in this study were small, and some classes and instructors appeared in more than one comparison. The results must, therefore, be considered tentative. There is some indication that the stability of behavioural scales found in this study approximates the stability of global student ratings reported in the previous literature. This stability, however, refers only to stability of individual differences, or an individual's tendency to be above or below the average on a given measure. The stability of patterns of behaviour, or teaching styles, is lower. An individual's relative strengths and weaknesses may shift considerably from one situation to another. This is not an uncommon finding when the component scores contributing to a pattern are, themselves moderately or highly correlated, as were the TBI scale scores in this study. In such cases, the reliability of the difference between two scores is often very low indeed. It is possible that more stability in teaching

styles would be observed with behavioural scales that are intercorrelated to a lesser extent.

DISCUSSION

Utility of Method

The results of this research are encouraging with respect to the potential role of empirical classification approaches in the description, and ultimately the understanding, of teacher behaviour. First, the use of student responses to a behavioural rating inventory produced reliable scores which were directly interpretable in terms of discrete, observable classroom teaching behaviours. Very little inference on the part of students was required in rating specific instructor behaviours, and similarly, very little inference on the part of the researcher was required in the interpretation of cluster profiles. One possible benefit which can be foreseen from this approach is that the description of naturally occurring patterns of teaching behaviour can proceed unhampered by the problems of trait definition or construct interpretation that otherwise tend to be all too common in the early stages of inquiry in psychology.

Second, the empirical methodology employed in this research successfully identified teaching styles which could be replicated in various ways. For example, the styles were stable across different starting points within the analysis. Furthermore, instructors were found to be grouped similarly regardless of whether specific behaviours or behavioural scale measures were used. Perhaps most importantly, the six

teaching styles were replicated in split halves of the sample, indicating that there is some generalizability of the styles across different groups of instructors. The relationships found between teaching style on the one hand and course level, subject area, and global effectiveness ratings on the other hand confirm the meaningfulness of the empirically identified styles.

The Six Teaching Styles

Through cluster analysis, six patterns of university lecturing behaviour were identified in this study. Members of the first cluster, Speech Givers, are characterized by their relaxed and confident manner, their speaking ability, and their lack of emphasis on explaining grading criteria or reminding students of deadlines. Structured but Boring instructors and Entertaining but Unstructured instructors are somewhat like two sides of the same coin. The former emphasize the organization and structure of the subject matter, but are dull in their presentation of it, whereas the latter are exciting and enthusiastic but show very little organization in their approach to the course content. Ineffective Lecturers are below average in almost all areas investigated, whereas Effective Lecturers are above average in almost all areas. Finally, Mumlbers are extremely hesitant when speaking in front of the class, are disorganized and unenthusiastic, do not explain the material well, and so on. In short, this final group appears to be

composed of not simply ineffective, but very poor lecturers. We can take some comfort in the indications provided by this research that Effective Lecturers represent a substantial proportion of the teaching population at the university where the research was conducted, and that Mumblerers are a much rarer breed.

Firm conclusions regarding the relative and absolute prevalence of these six teaching styles cannot, of course, be drawn from the data presently available. The effect of self-selection of instructors into research studies of teaching effectiveness is unknown. It is most likely, however, that the more successful instructors will tend to be over-represented in any volunteer sample. Almost one-half of the instructors participating in Study 1, and all of those in Study 2, were volunteers who responded to a general notice received in the mail. Many of the participants in Study 1 were purposefully "recruited" to provide equal representation of instructors receiving high, medium, and low overall student ratings. This latter approach would tend to over-represent extreme groups and under-represent average instructors.

The impact of overall profile elevation, or in other words global teaching effectiveness, on the identification of the six teaching styles cannot be denied. Three of the six styles are noted more for being above or below average than for being particularly strong or weak in any one specific area. The other three styles have clear patterns

of strengths and weaknesses. Future research should be directed toward a further investigation of this effect. In order to determine the meaningfulness of distinctions made on shape alone versus those made on both shape and elevation.

It is instructive to compare the six teaching styles identified in the present research to those discussed by Brown, Bakhtar, and Youngman (1984). The Oral Lecturer described by Brown et al. may be similar to the Structured but Boring lecturer identified here. Both types are above average on structure and show little variation from the traditional oral presentation approach. Two other styles identified by Brown et al., Amorphous and Exemplary, were described as being generally ineffective and effective, respectively, and thus might be viewed as similar to the Ineffective Lecturers and Effective Lecturers found in the present research. Also, the Brown et al. Self-Doubter category, whose members were described as having difficulty in keeping to the main topic of the lecture and in structuring the course material, might be considered analogous to the Entertaining but Unstructured group in the present study, although the degree to which Brown's Self-Doubter group would be considered enthusiastic is unclear. The remaining style discussed by Brown and his colleagues, Information Providers, did not have a clear parallel among the six styles identified by the present investigator. Similarly, the Speech Givers and Mumbler

categories of the present research did not have identifiable counterparts in the Brown et al. study.

These linkages are tentative at best, as differences in the number and type of behaviours assessed make direct comparisons difficult. Two major differences in methodology between the present research and that conducted by Brown and his colleagues should also be noted. With regard to instructor populations, Brown et al. studied lecturers in the British university system, whereas the present study used lecturers in a Canadian university. It is possible that there are fundamental differences between teaching practices in British vs. Canadian systems that could contribute to differences in teaching style classification results. A second and perhaps more important methodological difference is that Brown et al. employed self-report measures of teacher behaviour, whereas the present investigator used student behavioural ratings based on classroom observation. The Brown et al. approach is based on the perhaps questionable assumption that instructors can provide generally accurate reports of their own classroom behaviour. Despite these methodological differences, a considerable degree of consistency may be noted in the teaching style taxonomies proposed by Brown et al. and by the present author, and this is particularly encouraging.

Stability of Teaching Style

The results of Study 3 showed that teaching style was stable across two courses taught within the same year provided that the two courses were at the same level. Even over as many as eight years there was considerable stability in the teaching style an instructor displayed in teaching two courses at the same level. In contrast, results for pairs of courses at different levels, taught either in the same or in different years, indicated that an instructor's style was not the same across courses. These findings are consistent with conclusions drawn by Murray, Rushton, and Paunonen (1986) that similarity of course plays a major role in determining the stability of global student ratings of instructor effectiveness. In the present research, it was possible to make only relatively gross distinctions among course levels (e.g., Introductory, Junior Honours). Sample size did not permit a separate examination of instances in which an instructor taught the exact same course on two occasions (as opposed to two courses in the same category). It is possible that even greater stability would be observed in situations where the same instructor teaches, for example, two sections of the same course in one year, or the same course in two different years.

In the present study, teaching style was described solely in terms of observable teaching behaviours, and as such, it was found to be stable across occasions. This result gives us some assurance that the stability observed

in past studies of global student ratings (e.g., Marsh, 1982; Marsh & Hocevar, 1984; Murray, Rushton, & Paunonen, 1986) is attributable to consistency in actual classroom behaviours of the instructor.

In examining the results of Study 3, it is clear that the stability of overall differences among instructors was greater than the stability of differences among behaviours within a teacher. This is evidenced by consistently high values for measures of relative stability (e.g., correlations of TBI scales across teachers on two occasions or the stability of distances to those clusters showing non-zero elevations) as compared to somewhat lower values for measures of absolute stability in style (e.g., correlations of profiles within a teacher across two occasions or stability of distances to those clusters showing primarily shape differences). In other words, an effective instructor is likely to be effective in all classes, provided that the classes are at the same level. Similarly, and perhaps unfortunately, an ineffective instructor is likely to be consistently ineffective across classes. However, an instructor who is "task-oriented" in one class may not be so in another. Future research should be addressed to this issue. For instance, further attention could be given to understanding what causes this apparent lack of stability in teaching style. It might, in fact, represent real differences in behaviour. Alternatively, it could be symptomatic of measurement procedures which are

prone to a certain amount of error -- for example, in relation to student judgments, or assessment of difference scores between two behaviour scales, or allocation of instructors to teaching style categories. Another possible direction for future research would be to investigate differences in overall effectiveness between instructors who use the same style consistently across courses vs. those who are less consistent in style, and thus might be viewed as more flexible or adaptable.

Relationships between Teaching Style and Other Variables

A number of educationally relevant variables were examined in relation to instructor teaching styles in this study. There was some evidence that the frequency of occurrence of the six teaching styles differed across academic disciplines and across course levels. Arts instructors in Study 1 were found to be over-represented in the Entertaining but Unstructured and Effective Lecturers clusters. Science instructors in Study 1 were more likely than instructors in other fields to fall in the Ineffective Lecturers category, and were less likely to be Speech Givers. The significant relationship found between teaching style and subject area in Study 1 was not replicated in Study 2. The difference in results cannot be fully explained with available data, although it is possible that it reflects differences in the volunteer vs. recruit composition of the two samples.

With regard to course levels, the results indicated an increasing tendency toward a less enthusiastic and more structured teaching style as the level of the course increased. In introductory and general interest courses, instructors were more likely to be Speech Givers or Entertaining but Unstructured, exhibiting a TBI score pattern of Expressiveness exceeding Organization. In honours courses the reverse was true, with the Structured but Boring teaching style being more common. Study 1 results indicated that lecturers in general interest courses were much more likely to be Mumlbers than were lecturers at other levels, a finding which might reflect departments' purposeful assignment of poor instructors to these low-priority courses. Although slight differences in results were noted between Study 1 and Study 2, the general findings with regard to the relationship between teaching style and course level were consistent.

The six teaching styles were also significantly different in terms of student ratings of overall teaching effectiveness. Generally this effect was accounted for by differences between teaching styles with uniformly high or uniformly low profiles across all teaching behaviour scales. Instructors who were above average on all behavioural scales were rated by students as being effective, whereas instructors who were consistently below average on all scales were rated as being ineffective. This finding is consistent with previous research by Murray (1983a, b) and

others showing that student ratings of overall instructor quality can be predicted from frequency ratings of specific classroom/teaching behaviours. Similar to the results reported above in relation to stability of teaching styles, it can be seen that profile elevation played a major role and profile shape a relatively minor role in determining student perceptions of overall effectiveness.

Relationships between teaching style and instructor attitudes or self-ratings were generally not statistically significant in this study. This finding is consistent with the general lack of significant relationships between teacher attitudes and overall teacher effectiveness (Feldman, 1986). There are several possible explanations for the failure to find links between instructor attitudes and instructor behaviours. First, the measures utilized in this study may not have been adequate for assessing the relevant attitudes. These measures have not been investigated with regard to construct validity, reliability, possible response biases, or potential restriction of range in volunteer samples. Second, it may be that instructor goals and attitudes are related to their perceptions of their own behaviour, and to the behaviour patterns which they intend to display, but are not related to the specific behaviours that they do display in the classroom, as accurately reported by students. Research evidence indicating that instructors are inaccurate at describing their own classroom behaviour (Aubrecht, Hanna, & Hoyt,

1986; Braskamp, Caulley, & Costin, 1979) suggests that intended behaviours and actual behaviours may in fact differ. Third, it is possible that the behavioural teaching style exhibited by an instructor is in fact unrelated to his or her attitudes toward education and teaching. This situation might exist when instructors lack the behavioural repertoire needed to display their target teaching style (McLean, 1979) or when conflicting roles or values restrict them from implementing their preferred style.

Whatever the reason or reasons, it must be acknowledged that the lack of significant relationships found between teaching styles and teacher self-report variables in this study casts some doubt on the meaningfulness or external validity of the typology of six teaching styles. On the other hand, it is fair to note that counterbalancing positive evidence for the meaningfulness of the taxonomy was found in the replication of the clusters across split samples, across scale and item data, and across various analytical methods, as well as in the evidence for stability of teaching styles across courses at the same level; and in the significant relationships found between teaching style on the one hand and course level and overall effectiveness on the other. Obviously future research should examine in greater depth relationships between teaching styles and other educational variables postulated to be relevant.

Future Research

The present research was designed to be primarily descriptive in nature. Clearly, the results obtained are more exploratory than definitive, but they do offer encouragement for the usefulness of this approach in future research endeavours. The goal of adequately describing lecturing styles through empirical classification has probably not yet been completely attained. The present description of identified teaching styles could be refined considerably by applying the typology to a larger sample for purposes of collecting "norms" on these behaviour patterns. Further information on the frequency of occurrence of various teaching styles in university faculty populations would be helpful in designing teacher improvement programs and in designing subsequent research studies. Beyond this, but still with the general goal of describing teacher behaviour, methodologies similar to those utilized here could be applied to the study of classroom behaviour in situations other than the traditional lecture format, such as laboratory courses, small group discussions, and tutorials (e.g., Baumgart, 1976).

Following the development of a fully validated descriptive typology, long-term research needs in this area will include the investigation of causal relationships, of aptitude-treatment interactions, and of differential presage-process-product relationships across styles. Causes of teaching behaviour are only vaguely understood at

present. Recent research by Erdle, Murray, and Rushton (1985) suggests that instructor personality traits directly affect classroom teaching behaviours, which in turn determine student instructional ratings. However, the direction of causation among these variables, plus the role of course characteristics and student perceptions of these characteristics, are obviously in need of further research.

Through the use of an empirical teaching style approach to describing educational treatments, we may hope to achieve more fruitful results in the investigation of aptitude-treatment interactions at the postsecondary level. Such an approach will allow researchers to reliably categorize instructors into qualitatively different groups, an essential requirement for detecting differential effects on particular groups of students. Equally important, the ability to describe the "treatment" in terms of discrete, observable behaviours will facilitate the interpretation of findings and the design of subsequent experimental studies for rigorous hypothesis testing.

It seems reasonable to speculate that relationships among presage, process, and product variables may differ across different teaching styles. For example, efficient use of class time might be expected to have positive effects on student learning for instructors who show a Structured but Boring style, whereas this same factor might be expected to have negative effects for instructors with a more free-wheeling Entertaining but Unstructured style, in that

efficient time management is probably inconsistent with this style.

In summary, there is some cause for encouragement in the present findings. The approach of describing lecturer behaviour in terms of patterns of behaviours, or teaching styles, seems to have merit. It is worthy of consideration in the design of future research studies in postsecondary teaching effectiveness.

REFERENCES

- Abbott, J. E. (1970). A theoretical model of college teaching style. Dissertation Abstracts International, 31, 6357-A.
- Aitkin, M., Anderson, D., & Hinde, J. (1981). Statistical modelling of data on teaching styles (with discussion). Journal of the Royal Statistical Society, A, 144, 419-461.
- Aitkin, M., Bennett, S. N., & Hesketh, J. (1981). Teaching styles and pupil progress: A re-analysis. British Journal of Educational Psychology, 51, 170-186.
- Andrews, J. D. (1980). The verbal structure of teacher questions: Its impact on class discussion. POD Quarterly, 2, 130-163.
- Aubrecht, J. D., Hanna, G. S., & Hoyt, D. P. (1986). A comparison of high school student ratings of teaching effectiveness with teacher self-ratings: Factor analytic and multitrait-multimethod analyses. Educational and Psychological Measurement, 46, 223-231.
- Axelrod, J. (1973). The university teacher as artist. San Francisco: Jossey-Bass.
- Baumgart, N. L. (1976). Verbal interaction in university tutorials. Higher Education, 5, 301-317.

- Bennett, S. N. (1976). Teaching styles and pupil progress. London: Open Books.
- Bennett, S. N. & Entwistle, N. J. (1976). Rite and wrong. A reply to 'A chapter of errors'. Educational Research, 19, 217-222.
- Bonnici, C. A. (1978). The two personalities of teaching. English Journal, 67(6), 48-50.
- Braskamp, L. A., Caulley, D., & Costin, F. (1979). Student ratings and instructor self-ratings and their relationship to student achievement. American Educational Research Journal, 16, 295-306.
- Brophy, J. (1986). Teacher influences on student achievement. American Psychologist, 41, 1069-1077.
- Brown, G. A., Bakhtar, M., & Youngman, M. B. (1984). Toward a typology of lecturing styles. British Journal of Educational Psychology, 54, 93-100.
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 70, 213-220.
- Costin, F. (1971). Empirical test of the "teacher-centered" versus "student-centered" dichotomy. Journal of Educational Psychology, 62, 410-412.
- Cronbach, L. J., & Gleser, G. (1953). Assessing similarity between profiles. Psychological Bulletin, 6, 456-473.
- Cronbach, L. J., & Snow, R. E. (1977). Aptitude and instructional methods. New York: Irving Publishers.

- Grookes, F. W. (1977). An investigation of the interaction of educational cognitive style, teaching style, and instructional area for selected community college instructors. Dissertation Abstracts International, 37, 7504-A.
- Donald, J. (1983). Structures of knowledge as predictors of learning. In H. G. Murray (Moderator), Research on college teaching. Symposium conducted at the meeting of the Canadian Psychological Association, Winnipeg.
- Dunkin, M. J. (1983). A review of research on lecturing. Higher Educational Research and Development, 2(1), 63-78.
- Dunkin, M. J. (1986). Research on teaching in higher education. In Wittrock, M. C. (Ed.), Handbook of Research on Teaching (3rd ed.). New York: Macmillan.
- Dunkin, M., & Biddle, B. (1974). The study of teaching. New York: Holt, Rinehart & Winston.
- Ebel, R. L. (1951). Estimation of the reliability of ratings. Psychometrika, 16, 407-424.
- Erdle, S., & Murray, H. G. (1986). Interfaculty differences in classroom teaching behaviors and their relationship to student instructional ratings. Research in Higher Education, 24, 115-127.
- Erdle, S., Murray, H. G., & Rushton, J. P. (1985). Personality, classroom behavior, and student ratings of college teaching effectiveness: A path analysis. Journal of Educational Psychology, 77, 394-407.

- Feldman, K. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review & synthesis. Research in Higher Education, 24, 139-213.
- Fischer, B. B., & Fischer, L. (1979). Styles in teaching and learning. Educational Leadership, 36, 245-251, 254.
- Elanders, N. A. (1970). Analyzing teaching behavior. Reading, Mass.: Addison-Wesley.
- Gagne, R. (1977). Conditions of learning. (3rd ed.). Toronto: Holt, Rinehart and Winston.
- Gall, M. D., Ward, B. A., Berliner, D. C., Cohen, L. S., Winne, P. H., Elashoff, J. D., & Stanton, G. S. (1978). Effects of questioning techniques in recitation on student learning. American Educational Research Journal, 15, 175-199.
- Good, T. L., & Weinstein, R. S. (1986). Schools make a difference: Evidence, criticisms, and new directions. American Psychologist, 41, 1090-1097.
- Cray, J., & Satterly, D. (1976). A chapter of errors. Educational Research, 19, 45-56.
- Hamachek, D. E. (1972). Personality styles and teacher behavior. The Educational Forum, 36, 313-322.
- Hambleton, R. K. (Ed.). (1983). Applications of item response theory. Vancouver: Educational Research Institute of British Columbia.

- Hegarty, E. H. (1978). Levels of scientific enquiry in university science laboratory classes: Implications for curriculum deliberations. Research in Science Education, 8, 45-57.
- Hope, A. C. (1968). A simplified Monte Carlo significance test procedure. Journal of the Royal Statistical Society, B, 30, 582-598.
- Klare, G. (1974-1975). Assessing readability. Reading Research Quarterly, 10, 62-102.
- Mann, R. D., Arnold, S. M., Binder, J. L., Cytrynbaum, S., Newman, B. M., Ringwald, B. E., Ringwald, J. W., & Rosenwein, R. (1970). The college classroom: Conflict, change, and learning. Toronto: John Wiley & Sons, Inc.
- Marsh, H. W. (1981). The use of path analysis to estimate teacher and course effects in student ratings of instructional effectiveness. Applied Psychological Measurement, 6, 47-60.
- Marsh, H. W., & Hocevar, D. (1984). The factorial invariance of student evaluations of college teaching. American Educational Research Journal, 21, 341-366.
- McLean, D. (1979). The effects of midsemester feedback upon weekly evaluations of university instructors. Unpublished master's thesis, University of Western Ontario.

- Merlino, A. (1977). A comparison of the effectiveness of three levels of teacher questioning on the outcomes of instruction in a college biology course. Dissertation Abstracts International, 37, 5551-A.
- Miller, B. J. (1979). Differences in teaching styles between college disciplines on student evaluations of instruction when low-inference items are used and halo effect is controlled. Dissertation Abstracts International, 40, 3975-A.
- Murray, H. G. (1977). How do good teachers teach? An observational study of the classroom behaviors of Social Science lecturers receiving low, medium, and high teacher ratings. Ontario Universities Program for Instructional Development Newsletter, No. 14.
- Murray, H. G. (1980). Evaluating university teaching: A review of research. Toronto: Ontario Confederation of University Faculty Associations.
- Murray, H. G. (1983a). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. Journal of Educational Psychology, 75, 138-149.
- Murray, H. G. (1983b). Low-inference classroom teaching behaviors in relation to six measures of college teaching effectiveness. Paper presented at the Conference on the Evaluation and Improvement of University Teaching, Montebello, Quebec.

- Murray, H. G., Rushton, J. P., & Paunonen, S. V. (1986). Teacher personality traits and student instructional ratings in six types of college courses. Manuscript submitted for publication.
- Newport, J. F. (1980). Describing teaching styles in operational terms. School Science and Mathematics, 80, 486-490.
- Popham, W. J. (1978). Criterion-referenced measurement. Englewood Cliffs, NJ: Prentice-Hall.
- Punj, G., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. Journal of Marketing Research, 20, 134-148.
- Raines, R. H. (1976). A comparative analysis of learning styles and teaching styles of mathematics students and instructors. Unpublished doctoral dissertation, Nova University. (ERIC Document Reproduction Service No. ED 162 709).
- Rogosa, D., Floden, R., & Willett, J. B. (1984). Assessing the stability of teacher behavior. Journal of Educational Psychology, 76, 1000-1027.
- Rosenshine, B., & Furst, N. F. (1971). Research on teacher performance criteria. In B. O. Smith (Ed.), Research in teacher education: A symposium. Englewood Cliffs, NJ: Prentice-Hall.
- Sarason, S. B. (1982). The culture of the school and the problem of change (2nd ed.). Boston: Allyn & Bacon.

- Shavelson, R. & Atwood, N. (1977). Generalizability of measures of teaching process. In Borich, G. D. The appraisal of teaching: Concepts and process. Don Mills Ont.: Addison-Wesley Publishing Co.
- Skinner, H. A. (1979). Dimensions and clusters: A hybrid approach to classification. Applied Psychological Measurement, 3, 327-341.
- Skinner, H. A. (1981). Toward the integration of classification theory and methods. Journal of Abnormal Psychology, 90, 68-87.
- Smith, P. L. (1979). The generalizability of student ratings of courses: Asking the right questions. Journal of Educational Measurement, 16, 77-87.
- Sneath, P. H. E. & Sokal, R. R. (1973). Numerical taxonomy. San Francisco: Freeman.
- Tobias, S. (1979). Anxiety research in educational psychology. Journal of Educational Psychology, 71, 573-582.
- Wishart, D. (1978). CLUSTAN User Manual (3rd ed.). Edinburgh: Edinburgh University.
- Wishart, D. (1982). CLUSTAN User Manual Supplement (to 3rd ed.). Edinburgh: Edinburgh University.
- Wyett, J. L. (1967). A pilot study to analyze cognitive style and teaching style with reference to selected strata of the defined educational sciences. Dissertation Abstracts International, 28, 3579-A.

APPENDIX A

Teacher Behaviors Inventory

TEACHER BEHAVIORS INVENTORY
(Revised October 1978)

Instructor: _____

Course: _____

Observer: _____

Instructions to observer:

In this inventory you are to report your observations of the classroom teaching behaviors of the instructor named above. In other words, you are to report specific, concrete things that the instructor does (or fails to do) in his or her classroom teaching. It is important that your responses be "descriptive" rather than "evaluative". Also, your reports should be based solely upon your own personal observations.

Each section of the inventory begins with a definition of the general category of teaching to be reported in that section. Following the category definition is a list of specific teaching behaviors falling within that category. Your task is to estimate the frequency of occurrence of each of these behaviors for the instructor in question, using the frequency rating scale shown below:

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

There are 100 behaviors listed on this inventory. Please mark your ratings on the two computer cards provided, using columns 1 to 50 of the green computer card for questionnaire items G1 to G50, and using columns 1 to 50 of the red computer card for questionnaire items R1 to R50. Note that you are to mark an A to indicate "Never", B for "Rarely", C for "Sometimes", D for "often", and E for "Always".

It is absolutely essential that you mark your answers with an ordinary lead pencil only. Also, be sure to press firmly, fill in the entire oval-shaped answer space, mark only one space per question, and erase completely any errors or stray marks.

Before beginning the questionnaire, please print the name of the instructor being rated on the back of both computer cards. Do not mark your own name or ID number anywhere on the answer card.

Note: (1) Frequency estimates for a given instructor should be based on a minimum of three separate one-hour classroom observation periods.

(2) The data you provide in this inventory will be kept strictly confidential and used solely for purposes of a research project on university teaching conducted by Dr. H. G. Murray of the Department of Psychology, University of Western Ontario.

Mannerisms movements, gestures, and other nonverbal behaviors of instructor in classroom situation

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- G1 moves back and forth in front of class
- G2 gestures with hands or arms
- G3 plays with chalk or pointer
- G4 exhibits facial gestures or expressions
- G5 avoids eye contact with students
- G6 rocks or plays on heels
- G7 walks up aisles beside students
- G8 gestures with head or body
- G9 leans on desk or lectern
- G10 smiles or laughs while teaching
- G11 exhibits distracting mannerisms

Explanation: ways in which instructor explains concepts or principles; methods used to explicate or clarify subject matter

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- G12 gives several examples or illustrations of each concept
- G13 uses concrete examples to explain abstract principles
- G14 fails to define new or unfamiliar terms
- G15 uses audiovisual aids (films, tapes, maps, slides, etc.) to illustrate concepts
- G16 writes key terms on blackboard or overhead screen
- G17 dwells excessively on obvious points
- G18 suggests ways of memorizing complicated ideas
- G19 repeats difficult ideas several times
- G20 stresses most important points by pausing, speaking slowly, raising voice, etc.
- G21 uses graphs or diagrams to facilitate explanation
- G22 gives everyday, real-life examples to illustrate concepts or principles
- G23 explains subject matter in informal, colloquial language

Organization: ways in which instructor attempts to organize subject matter; methods used to create structure and coherence

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- G24 reviews topics covered in previous lecture at beginning of each class
- G25 uses headings and subheadings to organize lectures
- G26 explains how each topic fits into the course as a whole
- G27 digresses from major theme of lecture
- G28 relates lectures to text, readings, or labs
- G29 puts outline of lecture on blackboard or overhead screen
- G30 lectures follow a logical sequence
- G31 covers very little material in class sessions
- G32 gives preliminary overview of lecture at beginning of class

- G33 periodically summarizes points already made
- G34 organizes lecture by means of a list of points or series of questions
- G35 clearly indicates transition from one topic to another
- G36 previews topics to be covered in future lectures
- G37 suggests organizational schemes designed to aid students in learning material

Interaction verbal interactions between teacher and students in classroom

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- G38 encourages students to ask questions or make comments during lectures
- G39 praises students for good ideas
- G40 asks questions of individual students
- G41 asks questions of class as a whole
- G42 encourages students to participate in classroom discussion
- G43 answers students' questions thoroughly
- G44 criticizes students when they make errors
- G45 asks if students understand before proceeding to next topic
- G46 expects students to answer questions directed to class
- G47 encourages students to think independently
- G48 fails to take the initiative in classroom interactions
- G49 sticks to the point in answering students' questions
- G50 incorporates students' ideas into lecture

Speech qualities of instructor's voice; manner of talking or speaking

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- R1 speaks in a dramatic or expressive way
- R2 stutters, mumbles, or slurs words
- R3 voice lacks proper modulation (speaks in monotone)
- R4 speaks softly
- R5 speaks loudly
- R6 pauses in mid-sentence
- R7 says "um" or "ah"
- R8 speaks slowly
- R9 speaks rapidly
- R10 does not speak clearly
- R11 voice fades in mid-sentence

Interest ways in which instructor attempts to generate student interest in subject matter

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- R12 tells jokes or humorous anecdotes
- R13 shows strong interest in subject matter
- R14 relates subject matter to student interests or student activities
- R15 describes personal experiences relevant to subject matter
- R16 reads lecture from prepared notes or text, thus reducing spontaneity

- R17 uses a variety of different activities, media, or formats (e.g., guest lecturers, panel discussions, etc.)
- R18 relates subject matter to current events
- R19 points out practical applications of concepts or principles
- R20 states own point of view on controversial issues
- R21 asks rhetorical questions
- R22 presents challenging, thought-provoking ideas
- R23 focuses on controversial issues within the subject matter of the course

Disclosure: instructor's explicitness concerning teaching objectives, course requirements, and grading criteria

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- R24 advises students as to how to prepare for tests or exams
- R25 makes students aware of overall objectives of course
- R26 tells students exactly what is expected of them on tests, essays, or assignments
- R27 provides sample exam questions
- R28 states objectives of each lecture
- R29 suggests supplementary readings on topics covered in lectures
- R30 reminds students of test dates or assignment deadlines
- R31 tells student which topics are most important for exam purposes

Rapport: instructor's attitudes and feelings toward students, interpersonal relations between teacher and students

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- R32 knows individual students by name
- R33 sensitive to students' feelings
- R34 tolerant of other points of view
- R35 concerned that students understand and learn subject matter
- R36 available for consultation outside of class
- R37 interested in students' ideas
- R38 fair and impartial
- R39 talks with students before or after class
- R40 flexible regarding deadlines and requirements
- R41 offers to help students with problems

Affect: emotional behaviors or affective states exhibited by teacher in classroom situation

A	B	C	D	E
Never	Rarely	Sometimes	Often	Always

- R42 concerned (vs. aloof)
- R43 relaxed (vs. nervous)
- R44 confident (vs. apprehensive)
- R45 happy (vs. sad)
- R46 casual (vs. formal)
- R47 agreeable (vs. argumentative)
- R48 expressive (vs. bland)
- R49 friendly (vs. reserved)
- R50 critical (vs. approving)

APPENDIX B

Student Rating Form I

Instructor: _____

Course: _____

Teaching Behaviours Study: Student Rating Form I

In this questionnaire you are asked to evaluate the effectiveness of your course and instructor. It is important that you be as candid and as objective as possible in making your ratings. You should attempt to rate your course and instructor as they compare to other university courses and instructors. DO NOT mark your name or ID number on the questionnaire. Please respond to each statement by writing the appropriate number, from the scale below, in the space provided.

1	2	3	4	5	6	7
strongly disagree	disagree	slightly disagree	neutral	slightly agree	agree	strongly agree

1. The work load in this course was relatively heavy. _____
2. This instructor has led me to think about things in a different way. _____
3. Compared to other courses, I feel I have learned a lot from this course. _____
4. Compared to other instructors, I feel I have learned a lot from this instructor (independent of course readings and assignments). _____
5. I found the material covered in this course interesting. _____
6. The material covered in this course was relatively difficult. _____
7. A lot of material was covered in this course. _____
8. This course has improved my ability to analyze information or solve problems. _____
9. This instructor has increased my interest in this area. _____
10. The course has increased my factual knowledge of this subject. _____
11. This instructor's lectures increased my understanding of this field. _____
12. I would take another course in this area, taught by this instructor (if not restricted by degree requirements). _____
13. I would take another course in this area, taught by a different instructor (if not restricted by degree requirements). _____
14. Overall, the instructor was an effective teacher. _____
15. Overall, the course was valuable. _____

Approximately what percentage of this instructor's lectures in this course did you attend? (check one)

- | | |
|---------------------|--------------------|
| less than 50% _____ | approx. 80% _____ |
| approx. 50% _____ | approx. 90% _____ |
| approx. 60% _____ | all lectures _____ |
| approx. 70% _____ | |

APPENDIX C

Instructor Self-Rating Questionnaire

Code: _____

TEACHING BEHAVIOURS STUDY:

INSTRUCTOR SELF-REPORT QUESTIONNAIRE

This questionnaire asks you for information about your courses, your opinions about your classroom teaching, and your attitudes toward teaching in general. Please feel-free to be candid in your responses, and add any comments that might clarify your views. Return the completed questionnaire to Denise McLean (Psychology Department) in the envelope provided.

INSTRUCTOR INFORMATION

Department: _____

Academic Rank: _____

Sex: Male _____ Female _____

Age: _____

No. of previous years teaching experience: _____

Teaching load this term: _____ hours per week (average).

_____ different courses taught

_____ students enrolled in all classes taught

PART I -- COURSE INFORMATION

Course _____ Class Size _____ Class Times _____

1. What is the usual method of presentation in this course? (check one)
- | | |
|--------------------------|------------------------------|
| lecture _____ | seminar _____ |
| lecture/laboratory _____ | other (please specify) _____ |
| lecture/discussion _____ | _____ |

2. What is the grading format? (please indicate percentage of final grade)
- | | |
|-----------------------------------|------------------------------|
| multiple choice exams _____ | seminar presentations _____ |
| problems/short answer exams _____ | class participation _____ |
| essay exams _____ | other (please specify) _____ |
| student papers/assignments _____ | _____ |

3. What percentage of the course content is available to students only through instructor lectures? _____

4. Which additional sources of material are available to students in this course? (check all that apply)

- | | |
|---|------------------------------|
| textbook _____ | class discussions _____ |
| study guide _____ | student presentations _____ |
| readings assigned by instructor _____ | tutorials _____ |
| guest speakers _____ | lab exercises, etc. _____ |
| commercially prepared films, etc. _____ | other (please specify) _____ |

5. What percentage of students do you estimate enroll in this course solely because it is a requirement of their program? _____

6. How many times have you taught this course in the past? _____

7. Approximately how many hours per week do you spend preparing for this course? _____

8. How similar is the subject matter of this course to your specific field of research interest? (circle your response)

Not at all similar							Extremely similar
1	2	3	4	5	6	7	

9. How much do you enjoy teaching this particular class?

Do not enjoy it at all						Enjoy it very much
1	2	3	4	5	6	7

Comments:

Course _____ Class Size _____ Class Times _____

1. What is the usual method of presentation in this course? (check one)
 lecture _____ seminar _____
 lecture/laboratory _____ other (please specify) _____
 lecture/discussion _____

2. What is the grading format? (please indicate percentage of final grade)
 multiple choice exams _____ seminar presentations _____
 problems/short answer exams _____ class participation _____
 essay exams _____ other (please specify) _____
 student papers/assignments _____

3. What percentage of the course content is available to students only through instructor lectures? _____

4. Which additional sources of material are available to students in this course? (check all that apply)

textbook _____	class discussions _____
study guide _____	student presentations _____
readings assigned by instructor _____	tutorials _____
guest speakers _____	lab exercises, etc. _____
commercially prepared films, etc. _____	other (please specify) _____

5. What percentage of students do you estimate enroll in this course solely because it is a requirement of their program? _____

6. How many times have you taught this course in the past? _____

7. Approximately how many hours per week do you spend preparing for this course? _____

8. How similar is the subject matter of this course to your specific field of research interest? (circle your response)

Not at all							Extremely
similar							similar
1	2	3	4	5	6	7	

9. How much do you enjoy teaching this particular class?

Do not enjoy						Enjoy it
it at all						very much
1	2	3	4	5	6	7

Comments:

Comparison of the Two Courses

1. Overall, how similar do you think the two courses are to one another? Consider type of student, course content, course structure, etc. (circle your response)

Extremely dissimilar								Extremely similar
1	2	3	4	5	6	7		7

Comments:

2. How similar do you think your classroom teaching styles are, in the two courses?

Extremely dissimilar								Extremely similar
1	2	3	4	5	6	7		7

Comments:

3. Overall, which students do you think enjoy their class more?

students enjoy class much more								students enjoy class much more
1	2	3	4	5	6	7		7

Comments:

General comments:

PART II -- CLASSROOM TEACHING STYLE

1. Please rate your classroom teaching, relative to that of other university instructors, on each of the following dimensions. Use the following scale and write your responses in the spaces provided. (If you see your teaching differently in different courses, add the necessary columns -- please, identify which column belongs to which course.)

Much below average	1	2	3	Average 4	5	6	Much above average 7
DISCLOSURE:				Explicitness and helpfulness regarding course requirements and grading criteria			_____
EXPOSITIONAL STYLE:				Methods used to explain, clarify, and/or generate student interest in subject matter			_____
ORGANIZATION:				Ways in which instructor attempts to organize subject matter; methods used to create structure and coherence in course content			_____
USE OF CLASS TIME:				Efficient utilization of class time; appropriate pacing of lecture; covering a large quantity of material			_____
SPEAKING ABILITY:				Qualities of instructor's voice; tendency to speak loudly and clearly			_____
EXPRESSIVENESS:				Use of movement, gesture, and vocal inflection to capture student attention during class			_____
RAPPORT:				Indications of positive instructor attitudes and feelings toward students; good interpersonal relations between teacher and students			_____
RELAXED MANNER:				Indications that the instructor is comfortable and at ease in front of the class			_____

Comments:

2. The various aspects of teaching can be characterized in many different ways. Listed below are six different roles that most instructors display from time to time. Try to indicate the approximate percentage of classroom time that you give to each of the roles. (Again, if you see yourself differently in different courses, you can add a second or third column and give more than one distribution of time.) The total might not equal 100.

- EXPERT: transmitting subject matter knowledge to students _____
- FORMAL AUTHORITY: setting goals, rules, deadlines for assignments, clearly defined standards of excellence _____
- SOCIALIZING AGENT: encouraging promising students to enter the instructor's field or discipline _____
- FACILITATOR: assisting students in achieving their own goals for learning, removing barriers to success (e.g., explaining how to use the library, find field experience, etc.) _____
- EGO IDEAL: serving as a model for students (e.g., indicating what a good historian, chemist, sociologist, etc. would do) _____
- PERSON: revealing experiences and feelings which indicate an existence outside of the classroom (e.g., using personal experiences, dreams, etc. to clarify an explanation) _____

3. (a) Which methods have you used to assess/improve your teaching effectiveness? (check all that apply)

- inferences from student academic performance _____
- informal discussion with students _____
- informal observation of student reaction _____
- review of formal student ratings _____
- use of supplementary written surveys of student opinion _____
- reading of articles related to teaching _____
- participation in earlier research projects _____
- consultation with colleagues _____
- consultation with Educational Development Office _____
- participation in formal teacher training program _____
- attendance at workshops on teaching _____
- substantial change in teaching style _____
- substantial reorganization/redesign of course _____
- other (please specify) _____

(b) How satisfied are you with the results of those methods (listed above) which you have used?

- Extremely
dissatisfied

Extremely
satisfied

1 2 3 4 5 6 7

4. Given below are brief descriptions of six classroom teaching styles which have been identified in earlier research. Please read each description and indicate how similar your teaching is to each of these. You may find that in one paragraph some characteristics are true of you and some are not; try to judge the overall similarity of your style to the description. Rate each description independently; do not rate them relative to each other.

Extremely dissimilar		Neither similar nor dissimilar				Extremely similar
1	2	3	4	5	6	7

- A. This style of lecturing is marked by an extremely focused and businesslike use of class time. The instructor almost always sticks close to the main topic during lectures and covers a lot of material. Speaking style can be described as dynamic, expressive, and confident. Very little class time is used for reiterating course requirements or helping students to prepare for examinations.
- B. Instructors using this style devote relatively little class time to "structuring activities" such as clarifying course requirements and organizing the sequence of topics for students. They are often less dynamic in their teaching style and probably spend little effort to "get to know" the students.
- C. This approach to lecturing represents mainly an attempt to capture and maintain student attention, primarily through the use of a dynamic and expressive teaching style. Strict adherence to the main topic, elaboration of grading criteria, and a highly structured presentation of the course material are apt to be considered less important.
- D. Lecturers using this classroom style give considerable attention to explaining how each topic fits with the others and how the lectures, and more broadly the subject matter field, are organized. They are perhaps not as dynamic in their presentation as other instructors, but they use class time more efficiently.
- E. A lecturer in this category is seen as having good rapport with the students and as trying to assist them in preparing for exams and in meeting the course requirements. Topics are usually presented in a highly structured manner. Students find the instructor to be a good speaker, capable of explaining material in a clear and interesting way.
- F. This lecturing style is characterized primarily by a tendency to speak quietly and sometimes hesitantly in front of a class. Very little class time is spent on "structuring activities" such as elaborating course requirements and grading criteria or emphasizing the organization of the lectures and the subject matter. Students may perceive that the instructor does not "relate to them". They may also think that class time is used inefficiently and that little material is covered.

PART III -- ATTITUDES TOWARD TEACHING

1. The faculty role is often described as consisting of teaching, research, and service.

(a) What percentage of your working time do you currently allocate to each of these? (total should be 100)

Teaching _____ Research _____ Service _____ Other _____
(specify) _____

(b) How important is each of these roles to you personally at this point in your career? (circle your responses)

	Not at all important						Extremely important
Teaching	1	2	3	4	5	6	7
Research	1	2	3	4	5	6	7
Service	1	2	3	4	5	6	7
Other (as above)	1	2	3	4	5	6	7

(c) How personally satisfied are you with your current performance in these areas? (circle your responses)

	Extremely dissatisfied						Extremely satisfied
Teaching	1	2	3	4	5	6	7
Research	1	2	3	4	5	6	7
Service	1	2	3	4	5	6	7
Other (as above)	1	2	3	4	5	6	7

2. How do you rate your teaching relative to other instructors at your rank?

Much below average			Average			Much above average	
1	2	3	4	5	6	7	

Comments:

4. This final series of items gives you an opportunity to describe how you prefer to teach. There are no right or wrong answers, so be as candid as possible in your responses. Read each of the 9 sets and, within each set, rate the four responses according to how well they describe your personal reactions or feelings. Judge the four responses in each set independently, not relative to each other.

Strongly
disagree

1

2

Neither agree
nor disagree

3

4

5

6

Strongly
agree

- A. I would describe the most effective teachers as:
- liking students and having a sincere desire to understand them.
 - considering each person as an individual and letting students work to their individual abilities and interests.
 - controlling their class and requiring everyone to meet some minimum requirement.
 - precise about what is to be done, how it is to be done, when it is to be done, etc.
- B. Generally speaking, I feel that the students who are doing very poorly in university
- shouldn't be in school at this stage in their lives.
 - have probably not done well in the past and are so far behind they'll probably never catch up.
 - would do OK if the faculty had more time to work with them.
 - could probably do well if they received the type of instruction they need to learn the material.
- C. I think that evaluation of student performance:
- often is unrelated to the supposed content and purposes of the course.
 - often doesn't really distinguish between those who are doing well and those not doing well.
 - often creates jealousies and hard feelings among students.
 - often isn't very helpful to students in evaluating their progress and focusing their study activities.
- D. In my undergraduate teaching,
- I am pleasant, friendly, and take a personal interest in the students.
 - I provide specific and detailed information about assignments and requirements.
 - I give students the opportunity to decide what they want to study and how they want to do things.
 - I set high standards and require the students do the work necessary to achieve them.
- E. If there were a small number of students in my class who appeared uninterested, I would
- consider that it is the student's problem since basically it is their responsibility to learn the material.
 - call on them more in class to try to force more application to their studies.
 - vary my approach and type of assignments to try to spark their interest.
 - put them with some of the better and more enthused students on a group project.

Strongly disagree			Neither agree nor disagree			Strongly agree
1	2	3	4	5	6	7

- F. In conducting an effective class, I think it is important that
 - a. students are allowed to set their own goals and study those topics of most interest to them. _____
 - b. the course is well organized and the topics follow one another so they can be easily interrelated. _____
 - c. frequent information is given the students on how their work compares to others', and the best work is fairly recognized. _____
 - d. the students share their ideas and enjoy one another. _____

- G. If a student fails a course, it is most likely because
 - a. the student didn't ask for help when he or she ran into difficulty. _____
 - b. the student shouldn't have been enrolled in the class in the first place. _____
 - c. the student didn't study enough. _____
 - d. the teacher didn't provide enough guidance or help. _____

- H. In my opinion, one of the most important responsibilities of a teacher is to
 - a. organize course content and programs in a logical and meaningful order. _____
 - b. motivate students to do the best work of anyone in the class. _____
 - c. get the students to cooperate with and help one another. _____
 - d. encourage the students to make their own decisions as to what they can accomplish. _____

- I. I would describe the worst teachers I've known as
 - a. more interested in the subject matter than in the students. _____
 - b. vague about assignments, so that the students were never sure what was expected of them. _____
 - c. too restrictive and not letting the students come to their own conclusions or do things on their own. _____
 - d. too easy, so that the students lost respect for them. _____

Comments: