

Improving the Instructional Effectiveness of University Teaching Assistants: Report I

Ronald W. Marx, Jack Martin, John F. Ellis, and John Hasell
simon fraser university

Soixante-sept assistants-professeurs venant des facultés d'art, de sciences et des études interdisciplinaires de l'Université Simon Fraser ont participé à une étude expérimentale d'un programme prévu pour initier les assistants aux principes de base et aux techniques de la pédagogie. Trois groupes d'assistants-professeurs ont reçu la formation en question tandis qu'un autre groupe, approximativement assorti au premier selon l'expérience d'enseignement et le département auquel les assistants appartenaient, était utilisé comme contrôle. Les mesures des attitudes et des perceptions des assistants-professeurs et des étudiants, bien que ne montrant pas un résultat efficace du traitement, ont confirmé de façon cohérente, la faveur envers les groupes ayant reçu une formation pédagogique.

C'est la première d'une série d'expériences actuellement en cours de réalisation à l'Université Simon Fraser et le document détermine les grandes lignes de la stratégie générale de ces études et traite des implications à considérer pour réussir des programmes de recherche.

The earliest research on teaching at the university level was concerned with class size (McKeachie, 1967). From the pioneering work of Edmondson and Mulder (1924) up to the present day, numerous studies have attempted to determine whether small classes are really more effective for teaching than are large classes. Generally, these studies have shown that large classes are simply not as effective as small classes for retention, critical thinking, and attitude change (McKeachie, 1970). Typical explanations proffered in support of these findings have been that the small class probably involves more student-to-student interaction, more discussion, and more individual projects, all of which translate, in accord with one or another theory, into greater student achievement and satisfaction. Since a second major finding from this body of research has been that students and faculty members consistently prefer small classes (McKeachie, 1967), it is not surprising that universities have attempted to provide small instructional settings for students. In many universities this has meant that graduate and undergraduate students have played increasing roles as instructors. On many campuses, undergraduate students engage in at least as much instructional contact with student teaching assistants as they do with professors or other faculty members.

Given the extensive instructional responsibilities of graduate and, to a lesser extent, undergraduate students in their roles as seminar or tutorial leaders, laboratory supervisors, and project consultants, evidence on their effectiveness and training requirements is extremely sparse. Projects aimed at training teaching assistants in pedagogical skills are few, and those in existence are typically far from well-developed (see Marx, Ellis & Martin, in press, for a report on programs in Canadian universities). Published empirical studies of the instructional effectiveness of graduate teaching assistants, and for that matter anyone else engaged in higher education, are virtually nonexistent. In the 1975 *Review of Research in Education*, McKeachie and Kulik were able to find only one research study in this area. This study (Macomber & Siegel, Note 1) indicated that graduate student teaching assistants leading small sections were at least as effective as professors lecturing to large groups.

Educational research and practice have come more and more to view instruction (teaching) as a specialized skill-oriented activity in which questions of "what to do" and "how to do it" hold similar importance to "the content being covered" (Gage, 1976). But university instruction appears still to be guided by the dictum "if you know your material, you'll teach it effectively." Consequently, the improvement of instructional effectiveness receives little more than lip-service and very few university faculties have developed programs aimed directly at improving teaching performance. The widespread practice of collecting student evaluations of the instruction they receive might be cited as evidence of a desire to improve teaching. But, in fact, these procedures are often used to make decisions about tenure and promotion (Murray, 1977). In any case, the global feedback received through student ratings does not seem to help instructors improve their performance (Kulik & McKeachie, 1975; Wittrock & Lumsdaine, 1977).

A similar development reflects the point of view that good teaching is idiosyncratic and, thus, cannot be the subject of scientific inquiry. A corollary of this position is that specific elements of good teaching cannot be taught. This point of view is clearly articulated by Bonham (1977), who states that "the essential success ingredients that go into good teaching can hardly be analyzed scientifically . . . for good teaching remains an art, not a science" (p. i). This comment by Bonham is part of an editor's remarks relating to a series of articles in *Change*, which attempt to document improved teaching within disciplines in university departments.

Simon Fraser University has attempted to improve the performance of its teaching assistants and to demonstrate by empirical means that improvements have been effected. The Dean of Graduate Studies and the Faculty of Education have jointly developed a program aimed at the training of teaching assistants in basic pedagogical skills and, at the same time, have developed procedures for assessing the effects of the training. In a series of evaluation studies, the teaching performances of TAs receiving pedagogical

instruction through the Teacher Assistant Training Programme (TAT) are being compared to the performances of TAs who have not received such training. These studies make use of a wide variety of data-collection methods, and consider a uniquely representative sample of criterion variables in attempting to ascertain whether TAT participants have a greater sense of satisfaction and competence in their instructional roles than do a comparable group of TAs who have not had this training; whether TATs are perceived by their students as better teachers than non-TATs; whether the actual classroom teaching skills of TATs differ in desirable directions from those of non-TATs; and whether the students of TATs achieve more than do their counterparts taught by non-TATs.

What follows is the first in a series of reports which will document the Simon Fraser TAT program. This study reports data collected during the first semester of the evaluation, and contains evidence regarding the TAs' sense of satisfaction and competence in their instructional roles and student perceptions of TAs' behavior.

METHOD

Training. The Teaching Assistant Training Programme consisted of 15 hours of instruction distributed over five evening sessions of three hours each. The objectives of the program were to teach the participants how to incorporate principles of learning in the design of instruction; to define, state, and use behavioral objectives in planning and delivering instruction; to develop valid and reliable measures of learning; to plan and conduct effective discussion groups or laboratory sections; and to be sensitive to ethical and moral dilemmas in teaching.

The first session covered five basic principles of learning. The class started off with a general discussion of "good" and "bad" teachers, and the roles of skilled behavior and personality in effective teaching. Following this, discussion focused on five principles to be considered when designing and delivering instruction. First, instruction should be meaningful to the learner. Second, activity and relevant practice needs to be provided. Third, students should receive prompt and informative feedback concerning the quality of their responses and assignments. Fourth, variability among students should be acknowledged and instruction and assignments modified accordingly. Fifth, teachers need to consider differences in student motivation — making instruction engaging and challenging. Both inductive and deductive teaching approaches were used in this session and an attempt was made on the part of the instructor to model desirable teacher behavior.

The second session dealt with planning instruction and stating behavioral objectives. The first part of the class was devoted to a presentation of behavioral objectives, including a discussion of the difference between course outlines and behaviorally stated objectives for student learning. In the second half of the lesson the participants wrote three behavioral objec-

tives and discussed them in small groups of three or four people. During this small group activity, each group generated three advantages and three disadvantages to the use of behavioral objectives. These formed the basis of the discussion which followed in the large group.

The third meeting dealt with principles and techniques of evaluation. Concepts central to the process of evaluation were described, discussed, and illustrated. These included: reliability and validity, sampling and comparison, measurement and evaluation, objectivity and subjectivity. Various types of test items were examined and discussed and the advantages and disadvantages of each were cited. The relationship of valuation to both instructional objectives and instructional procedures was stressed.

The fourth session considered techniques and problems in leading discussion groups and working in laboratories. The first part of the lesson was a didactic presentation which distinguished between specific teaching skills and more general strategies of teaching. The participants discussed the skills required, for example, to enhance participation in discussion groups, questioning techniques, and how to deal with obstreperous students. In the second part of the meeting, the participants broke into smaller groups and discussed the role of the TA. The session concluded with a large group discussion based on the roles defined in the smaller groups.

The last meeting started off with a visit from the Dean of Graduate Studies, who led a discussion on the general problems of graduate studies and the relationship between graduate student obligations and teaching assistant responsibilities. The second half of the session was devoted to discussion of some ethical dilemmas teaching assistants may find themselves confronted with. For example, what does the TA do when the professor appears incompetent, or when students court the TA with favors?

Subjects. During the first phase of the TAT project a total of 67 teaching assistants participated in the study, of whom 60 were graduate students and 7 were undergraduates. (Some departments employ undergraduate TAs when their respective graduate programs are nonexistent or unable to provide a large enough workforce.) The TAs were selected by their respective departments with the requirement that they had no previous teaching experience and were sufficiently competent speakers of English to participate in and learn from the seminars. Unfortunately, neither of these two criteria was adhered to strictly. Nineteen of the participants had previous teaching experience with a minimum defining criterion of at least one year of teaching at any level of schooling. Also, there were a number of students (10%–15%) who had difficulty speaking English.

Instruments. Two paper and pencil inventories were used to assess the effectiveness of the TAT program. The Teaching Assistant Inventory (TAI) assessed the TA's perception of and attitudes toward the TAT program and the concepts employed in the program. The Student Inventory (SI) assessed student perceptions of the TAs.

The TAI was designed to test seven hypotheses. The 28-item inventory

had 4 items on each of seven subtests relating to TAs' attitudes toward: their own teaching, their students, their academic discipline, their own learning, their desire to improve their teaching, the Faculty of Education, and the TAT program. Internal consistency coefficients for the seven subscales ranged from .53 to .72. One item, relating to TAs' attitudes toward their students, was dropped from the second subtest because it did not contribute substantially to the subtest score.

The SI was designed to test 10 hypotheses concerning students' perceptions of the teaching of their TAs and 4 hypotheses concerning student attitudes. The 10 hypotheses concerning perceptions focused on: clarity of objectives, correspondence of objectives and evaluation, application of principles of learning, use of instructional materials, ability to lead tutorial groups, TA satisfaction with teaching, TA attitude toward students, TA attitude toward the academic discipline of the course, TA attitude toward learning, and TA willingness to improve teaching. The 4 student attitudes were related to: satisfaction with their TA, attitude toward the academic discipline of the course, attitude toward the course, and attitude toward their own learning. In addition, 2 cluster scores were formed, 1 from items contributing to the former 10 subtests and 1 from items contributing to the latter 4 subtests. Internal consistency coefficients ranged from .51 to .81 for the subtests, and were .93 and .81 for the 2 cluster scores, respectively. The subtest assessing student perception of the TAs' correspondence of objectives and evaluation had very poor internal consistency (.31). Consequently, that subtest was dropped from further analyses. Most of the hypotheses tested by the TAI and SI were derived from the stated objectives and instructional content of the TAT program. Remaining hypotheses, such as attitude toward the Faculty of Education, were included to provide information relevant to the administration of the TAT program at Simon Fraser University. Since the items constructed to test these hypotheses appeared to have face validity with respect to the foregoing bases, and met acceptable levels of reliability, more sophisticated tests of instrument validity were not undertaken.

Design. The original design called for three groups. Group A was to receive instruction on Wednesday evenings during the first half of the fall semester, 1976. Twelve of the students in this group had scheduling conflicts and a new group (A') was formed to meet on Thursday evenings. Group B met during the second half of the semester and received the same instruction as groups A and A'. Group C was a control group which did not receive instruction.

Table 1 summarizes the design of the study showing the number of TAs in each group, the number of TAs with and without previous teaching experience, and the three faculties in the university from which the students came. Three departments, Economics and Commerce, Psychology, and Biological Science agreed to have their TAs participate in group B, and the rest of the participants were assigned to A or A' according to their

Table 1 / Number in Sample by Group, Prior Teaching Experience, and Faculty

Faculty	Group									
	A		A'		B		C		Total	
	E	NE	E	NE	E	NE	E	NE	E	NE
Arts ^a	2	9	0	1	4	7	2	4	8	21
Interdisciplinary Studies ^b	1	3	1	5	0	1	1	5	3	14
Science ^c	2	4	3	2	2	2	2	4	9	12
Total	5	16	4	8	6	10	5	13	20	47

Note: E = prior teaching experience; NE = no prior teaching experience.

a. Departments of Archaeology, Economics and Commerce, History, Political Science, Psychology, and Sociology and Anthropology.

b. Departments of Communications Studies, Computing Science, and Kinesiology.

c. Departments of Biological Science, Chemistry, and Physics.

respective schedules. Demographic data were collected from all of the TAs at the university, and subjects in group C were randomly selected from this pool, stratified by previous teaching experience and department.

The TAI was administered to all participants in groups A, A', and B at the end of the last period of instruction. A random sample of one-third of these participants, stratified by faculty, in groups A, A', and B, were selected to obtain data on the Student Inventory. In addition, all of the participants in group C contributed data to the SI. The SI was sent to each TA in this smaller sample, who in turn administered it to their students. Four TAs did not return SI data: one from group A, one from group A', and two from group C. In addition, the SI data from one TA in group A' contained considerable missing data and were excluded from analysis.

RESULTS

Two separate sets of 3×2 factorial analyses of variance were performed on the seven subtests on the TAI. For the first set of analyses, the three-level factor was groups (A, A', and B) and the two-level factor was previous teaching experience (present and not present). The second set of analyses used groups (A, A', and B) as the three-level factor and discipline (science area or arts area) as the two-level factor. Since the second set of analyses showed no significant main effects for the discipline factor, and since main effects for the group factor were consistent with those revealed by the first set of analyses, only the results of the group-by-previous-teaching-experience analyses are reported. Table 2 presents the means and standard deviations for the six cells in each analysis of variance, and the *F* statistics associated with the two main effects and the interaction. While it is a shortcoming of this study that TAI forms were not administered to the control group C,

Table 2 / Summary of Analyses of Variance Results for the TAI Subtests: Group and Previous Teaching Experience Factors

Attitude toward	Group						F Group (df = 2 / 43)	F Experience (df = 1 / 43)	F G × E (df = 2 / 43)
	A / NE (n = 16)	A / E (n = 5)	A' / NE (n = 8)	A' / E (n = 4)	B / NE (n = 10)	B / E (n = 6)			
Teaching	3.98 .54	3.25 1.41	3.94 .45	3.56 .69	4.06 .21	3.92 .46	<1	4.71*	<1
Students	4.23 .60	3.60 .44	4.04 .21	4.17 .79	3.97 .37	4.11 .98	<1	<1	2.08
Academic discipline	4.45 .46	4.45 .67	4.16 .44	5.00 .00	4.18 .58	4.38 .61	1.08	3.36†	2.26
Learning	4.36 .54	4.45 .67	4.44 .37	4.69 .32	4.15 .43	4.71 .29	<1	4.20*	<1
Improved teaching	3.89 .56	3.65 .65	4.00 .44	3.75 .58	3.79 .64	3.46 .68	<1	2.30	<1
Faculty of Education	3.70 .76	3.72 .71	3.88 .61	4.12 .32	3.78 .34	3.71 1.37	<1	<1	<1
TAT Program	3.81 .92	3.70 1.05	4.44 .50	4.44 .52	3.70 .52	3.83 1.45	2.68†	<1	<1

Note: Upper number is mean; lower number is standard deviation.

E = prior teaching experience; NE = no prior teaching experience.

†p < 0.10; *p < 0.05.

TAI results for the experimental groups are included in this report because they assist in the interpretation of subsequent analyses on the SI scales.

The only significant main effect for groups was in the attitude toward the TAT program subtest, and this was only marginally significant. Two orthogonal contrasts, weighted to adjust for unequal cell size, were performed on the group main effect. One compared groups A and A' with B to determine if there was an effect for the timing of the TAT program with respect to the academic semester. (Recall that groups A and A' participated in the first half of the semester and group B participated in the second half.) The second orthogonal contrast compared groups A and A'. The contrasts indicated that the significant effect on the attitude toward the TAT program subtest resulted from a significant difference between groups A and A' ($t = 2.14, p < .05$).

There were three significant main effects for previous teaching experience. The first, on the attitude toward teaching subtest, resulted from higher scores from TAs with no previous experience. The two remaining significant effects, on attitude toward the TAs' own academic discipline and toward the TAs' own learning, resulted from significantly higher scores from TAs with previous teaching experience.

Intercorrelations among the seven TAI subtests were computed separately for each group. There were no significant differences among groups on these intercorrelations, so the data were pooled across all 49 participants in groups A, A', and B. The resulting intercorrelation matrix on the TAI subtests showed that most of the subtests measured relatively independent attitudes. The only exception involved attitudes toward the Faculty of Education and the TAT program ($r = .74, p < .01$). Excluding this correlation, the intercorrelations ranged from $-.11$ to $.48$, with a median of $.21$.

The SI items were grouped into the 13 subtests and two cluster scores referred to above. Mean subtest scores were computed across all students taught by the respective TAs. These scores were used for all SI analyses.

Adopting the same analysis strategies employed with the TAI scores, the SI scores were analyzed for group (A, A', B, C), previous teaching, and discipline effects, through two separate sets of 4×2 factorial analyses of variance. However, due to the sampling design for the SI, only one TA with previous teaching experience was selected for each of groups A and B, and only one TA from a discipline in the arts area was selected from group A'. Additionally, one of the two TAs with previous teaching experience in group A' had considerable missing data on the SI protocols, and his data were not included in the analyses. Consequently, three of the eight cells in the group by previous teaching experience design, and one of the eight cells in the group by discipline design had only one subject per cell. This invalidated the testing of the interaction effects in the factorial analyses of variance, so the variance due to the interaction was pooled with residual variance to form the error term for the main effect F ratios.

Few *F* tests in the group by previous teaching experience analyses were significant. It is noteworthy that the presence or absence of previous teaching experience in the TA was virtually irrelevant to student perceptions of the TA's teaching. Indeed, the only *F* statistic for the previous teaching variable that approached significance was for the student attitude toward the academic discipline of the course ($F_{1,25} = 3.28, p < .10$) and this favored TAs without previous experience.

In the analyses of variance results of the group by discipline design, significant *F* ratios for the discipline effect were observed on the subscales measuring student perceptions of tutorial leadership, student perceptions of the TA's attitude toward his students, and student attitudes toward the course discipline. Students in the arts areas perceived their TAs as offering superior tutorial leadership and as possessing more positive attitudes to their students. Science area students showed more positive attitude to the discipline that they were studying in the course.

To assist in the interpretation of the analysis of variance results for the group effect, orthogonal contrasts, weighted for unequal cell size, were also used to test group differences on the SI subtests. Table 3 reports the *t* statistics resulting from these contrasts.

The first contrast compares the experimental groups with the control

Table 3 / Tests from Orthogonal Contrasts Between Groups on Student Inventory Subtests

Subtest	Contrast											
	A A' B C				A A' B C				A A' B C			
	1	1	1	-3	1	1	-2	0	1	-1	0	0
Objectives	1.50				0.12				-0.69			
Principles of learning	1.88†				0.55				-2.37*			
Materials	0.00				2.04†				-1.14			
Tutorial Leadership	0.49 (9.1)				1.10 (8.5)				-1.63 (8.0)			
RA satisfaction	1.13 (14.3)				0.16 (7.2)				-4.05* (4.2)			
Attitude, students	1.55				0.60				-2.81**			
Attitude, discipline	-0.08				2.16*				-1.52			
Attitude, learning	1.23				0.37				-1.69			
Improve teaching	1.00				1.55				-1.35			
Total: Teaching	1.17				1.73†				-2.44*			
Satisfaction with TA	0.94				1.38				-2.05†			
Attitude, discipline	1.25				0.32				-0.75			
Attitude, course	1.30				0.40				-2.09*			
Attitude, own learning	0.04				0.47				-2.15*			
Total: Student	0.60				0.27				-3.25**			

Note: Two-tailed tests, *df* = 26 (except where separate variance estimates were employed; in these cases *df* are given in brackets following each *t* value).

†*p* < .10; **p* < .05; ***p* < .01.

group. The second contrast compares the two experimental groups participating at the beginning of the semester with the group receiving instruction at the end of the semester. The third contrast compares the two groups participating at the beginning of the semester.

All of the mean differences, except one, for the first contrast were in a direction which favored the experimental groups, although only one approached significance when subjected to *t* tests. Also, the mean differences for all of the second contrasts favored experimental groups A and A' over B. The most consistent and significant effects occurred in the third contrast, which demonstrated the superiority of group A' over A, and indeed, over groups B and C as well. Tests for homogeneity of variance (*F* max, Bartlett-Box, and Cochran's *C*) showed severe problems with 2 of the 15 variables on the SI. Although these statistics are notoriously problematic with small sample sizes (see Guilford, 1965), the contrasts were calculated with separate variance estimates for each cell to ameliorate the problems of heterogeneity of variance. The 2 variables for which separate variance estimates for each cell were employed are identified in table 3.

Intercorrelations among the 13 SI subtests were computed for the subsample which completed the instrument. These ranged from $-.04$ to $.85$ with a median of $.43$, indicating that student attitudes as measured by the SI tended to be more globular than TA attitudes as determined by the TAI.

DISCUSSION

The statistical tests provided little evidence of a strong treatment effect for the TAT instructional program. Only one of the first-order contrasts on the SI variables approached significance. This variable, students' perception of their TA's use of appropriate principles of learning, favors the TAT groups as a whole. These outcomes are not really surprising when one considers that the instructional program was of short duration (15 hours of contact) and attempted only to introduce student TAs to the basic principles and techniques of teaching.

Nonetheless, the absence of strong statistically significant outcomes on single SI variables should not be interpreted as evidence that TAT training was ineffective. Indeed, a ranking of group means across the three TAT groups and control group over all the SI variables indicates a general trend in the data which strongly favors group A' and places group B slightly ahead of groups A and C. Comparing this overall trend in the SI data to the TAI results tabulated in table 2, it comes as no surprise to find that group A' evidenced significantly more positive attitudes to the TAT program than did groups A and B.

In light of these disclosures, it appears that the TAT program was more effective for one of the three TAT groups. Positing post hoc speculations which attempt to account for this group-program interaction effect can be a risky business. For example, when the dynamic interactive nature of

instructional exchanges is considered, the question is probably moot as to whether the more positive attitudes held by TATs in group A' caused them to benefit more from the TAT sessions, or whether something in the TAT sessions caused these more positive attitudes. Further, to maintain that, since the instructional activities were held constant across all TAT groups, the relative success of group A' must be attributable to some distinctive characteristic possessed by A' and not by A and B fails to recognize that the moment-to-moment reactions of a student group during instruction can produce subtle changes in instructor styles, while formal strategies, activities, and content remain essentially unaltered (McLeish & Martin, 1975). What is known, however, is that group A' was smaller than groups A and B ($n = 12$, compared to n 's of 21 and 16 respectively) and that a greater proportion of students in A' were from the faculties of science and interdisciplinary studies, as opposed to groups A and B.

A final point of interest arising from the SI and TAI results concerns the familiar, although somewhat sobering, discovery that the enthusiasm of an instructor for his own discipline is not always transmitted to the students in his classroom. This phenomenon is rather graphically illustrated by the fact that TAs with previous teaching experience held more positive attitudes toward their own academic disciplines than did inexperienced TAs (see table 2), but the latter group appeared to be more successful in promoting positive attitudes to the course discipline on the part of their students. A corollary to this interesting incongruence in perceptions between TAs and students is that while students in science courses had more positive attitudes towards the course discipline than did their counterparts in the arts area, they held less positive perceptions of their TAs. Presumably students are capable of discriminating course content from the instructional styles of their TAs.

In closing this first report, it should be remembered that any program which attempts to improve the teaching effectiveness of instructors at any level must be adjudicated from the framework of dependent variables of student achievement as well as those of student and instructor attitudes and perceptions. The data reported here are from the first semester (Autumn 1976) during which the TAT program was evaluated. The experimental groups of this study have subsequently (Spring 1977) been compared to more recent TAT and control groups, and achievement as well as attitudinal and perceptual data were collected across all these groups. In the second evaluation study, control groups completed both SI and TAI questionnaires (see Martin, Marx, Hasell, & Ellis, 1978).

To summarize, the results of this study appear to show that the TAT program is influencing participants in desirable directions, although perhaps less dramatically than might have been hoped. The interaction between the instructional activities and different sets of participants seems to have favored one group more than the other two. Group size and com-

position are factors which appear to warrant further experimental attention. In this particular study, positive attitudes toward the TAT program did not seem to be differentially associated with previous teaching experience or academic discipline.

NOTE

This study was supported by funds provided by Simon Fraser University. The authors would like to thank Dr. Jon Wheatley, Dean of Graduate Studies, for his support, and Wayne Materi for his assistance with data analysis.

REFERENCE NOTE

1. Macomber, G. G., & Siegel, L. *Final report of the experimental study in instructional procedures*. Oxford, O.: Miami University, 1960.

REFERENCES

- Bonham, G. W. About this special report. *Change: The Magazine of Higher Learning*, 1977, 9(1), i.
- Bondson, J. B., & Mulder, F. J. Size of class as a factor in university instruction. *Journal of Educational Research*, 1924, 9, 1-12.
- Gage, N. L. (Ed.). *The psychology of teaching methods: The seventy-fifth yearbook of the National Society for the Study of Education*. Chicago: University of Chicago Press, 1976.
- Guilford, J. P. *Fundamental statistics in psychology and education* (4th ed.). New York: McGraw-Hill, 1965.
- Kulik, J. A., & McKeachie, W. J. The evaluation of teachers in higher education. In F. N. Kerlinger (Ed.), *Review of research in education 3*. Itasca, Ill.: Peacock, 1975.
- Martin, J.; Marx, R. W.; Hasell, J.; & Ellis, J. F. Improving the instructional effectiveness of university teaching assistants: Report II. *Canadian Journal of Education*, 1978, 3, 13-26.
- Marx, R. W.; Ellis, J. F.; & Martin, J. The training of teaching assistants in Canadian universities: A survey and case study. *Canadian Journal of Higher Education*, in press.
- McKeachie, W. J. Research in teaching: The gap between theory and practice. In C. B. Lee (Ed.), *Improving college teaching*. Washington, D.C.: American Council on Education, 1967.
- McKeachie, W. J. *Research on college teaching: A review*. Washington, D.C.: ERIC Clearinghouse on Higher Education, 1970.
- McKeachie, W. J., & Kulik, J. A. Effective college teaching. In F. N. Kerlinger (Ed.), *Review of research in education 3*. Itasca, Ill.: Peacock, 1975.
- McLeish, J., & Martin, J. Verbal behavior: A review and experimental analysis. *Journal of General Psychology*, 1975, 93, 3-66.
- Murray, H. G. Limitations of student ratings of college teaching. *Canadian Association of University Teachers Bulletin*, 1977, 25(3), 28.
- Wittrock, M. C., & Lumsdaine, A. A. Instructional psychology. *Annual Review of Psychology*, 1977, 28, 417-459.

Ron Marx and Jack Martin are assistant professors, John Ellis a full professor, and John Hasell a master's candidate in the Faculty of Education at Simon Fraser University, Burnaby, B.C., V5A 1S6.