Improving the Instructional Effectiveness of University Teaching Assistants: Report II

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

Ce document présente les résultats de la seconde phase d'une évaluation expérimentale du programme de formation des assistants-professeurs à l'Université Simon Fraser. Quatre groupes d'assistants-professeurs qui ont participé à ce programme au cours de l'année 1976/77 sont évalués les uns par rapport aux autres, par rapport à un groupe de contrôle et par rapport à un groupe ayant participé antérieurement (1975). L'évaluation est faite en fonction de variables telles que les attitudes et les perceptions des assistants-professeurs, les attitudes et les perceptions des étudiants, les résultats relatifs au rendement de l'étudiant. Les résultats confirment bien l'efficacité du programme de formation des assistants-professeurs en améliorant la qualité de l'enseignement donné par les assistants-professeurs. Le document traite d'un certain nombre de questions relatives à l'offre de tels programmes.

For most people, the concept evoked when the word "university" is read or heard is undergraduate education. Although universities fulfil many other roles, it is undergraduate education which is most visible. Within the last decade, the competence with which Canadian universities fulfil the highly critical role of undergraduate education has come under increasing attack. From the earlier pressures of student cries for "relevance," to more recent pressures toward fiscal restraint imposed by provincial legislatures, universities are feeling the need to demonstrate evidence of accountability and teaching effectiveness.

The responses by academics to these pressures have taken a number of forms. The most visible has been the institutionalization of efforts to improve instruction. Many examples can be cited of this development (Marx, Ellis, & Martin, 1978). The University of New Brunswick holds an annual seminar on teaching, which focuses on topics such as interpersonal communication, evaluation of teaching and learning, and the instructional use of media. McGill University, through the Centre for Learning and Development, has prepared a package of instructional modules entitled "Teaching and Learning," which is designed to improve the effectiveness of "graduate students who are teaching, directing seminars, grading papers, and leading laboratory sessions" (Notar, 1977, p.14). Similar efforts at centres such as the University of Manitoba's In-service Program for University Teachers have been underway since the early 1970s. Perhaps the most

extensive institutionalized efforts to improve university instruction are carried out by the Ontario Universities Program for Instructional Development, whose 1977/78 budget is $500,000 (*Minister increases program funding*, 1977). This organization funds a wide variety of programs in areas such as curriculum development, faculty evaluation, and teaching skills workshops.

Aside from the more extensive institutional attempts to improve instruction, most efforts have remained ad hoc. Usually the responsibility for such efforts falls on the various university faculties and departments, often with the admonition that evidence of teaching effectiveness must be presented for tenure and promotion decisions. While all of these movements to improve the effectiveness of undergraduate instruction are laudatory, they have – except in isolated cases such as McGill's program cited above or Toronto's course offered in New College (*University of Toronto offers course*, 1976) – ignored teaching assistants, a group who collectively deliver much of the undergraduate instruction in today's universities.

At Simon Fraser University, the Dean of Graduate Studies and the Faculty of Education have collaborated in the development of a program to train teaching assistants. The Teacher Assistant Training Programme, which has been operating for four years, has trained over 150 teaching assistants in many departments across the University. Each student is given a stipend for participating, and in most cases, this stipend is used to reduce the teaching load of the student by 20% to 25%.

During the first two years the program was evaluated primarily on the basis of opinions of participants. Generally, these opinions supported the program as a valuable experience which resulted not only in enhanced pedagogical skills but in increased interdisciplinary communication through contacts made in the seminars. During the 1976/77 academic year, however, a formal, large-scale evaluation was undertaken to assess the impact of the program on teaching assistants' opinions and attitudes, undergraduate students' attitudes and perceptions, and undergraduate student achievement.

Teaching assistant attitudes were sampled immediately upon completion of the program, and undergraduate student attitudes and perceptions of their teaching assistants were sampled at the end of the semester during which the program took place. These data are presented in Report I (Marx, Martin, Ellis, & Hasell, 1978). The results indicated that the program's effects were moderated by the particular group of teaching assistants attending the seminars. In particular, one group of TAs had far better attitudes about teaching and learning than the other two (trained) groups or the control group. There were no systematic effects related to previous teaching experience, a result which suggests that the program was sufficiently effective to overcome whatever learning occurred in the course of previous teaching experience. There were some differential effects on under-

graduate student attitudes and perceptions as a function of whether the teaching assistant was in a science or non-science discipline, indicating either that the program had different effects depending on the discipline of the teaching assistant, or that undergraduate students in sciences and non-sciences held different attitudes and perceptions regardless of the level of training of their TA.

This paper reports additional data concerning the effects of the program. First, it includes data from another group of participants. Second, it reports data parallel to that reported in the first paper (concerning teaching assistant attitudes and undergraduate student attitudes and perceptions) but collected a semester after instruction. Third, data are reported from a subsample of the program's participants, comparing the achievement of their students with that of students taught by control teaching assistants teaching the same courses. Finally, detailed interviews were conducted with a subsample of participants, focusing on the quality of the instruction in the program and the satisfaction experienced by participants.

METHOD

Training

The Teaching Assistant Training Programme has been described in detail by Marx, Martin, Ellis, & Hasell (1978). Briefly, the program consisted of 15 hours of instruction distributed over five evening sessions of three hours each. The respective sessions addressed (1) the incorporation of learning principles in instructional design, (2) the use of behavioral objectives in preparing and conducting instruction, (3) the development of valid and reliable measures of student learning, (4) instructional skills appropriate to the conduct of effective discussion groups or laboratory classes, and (5) ethical and moral issues in teaching. In short, the instructional component of the program consisted of a modified, condensed course in educational psychology. The instructional activities included brief lectures, large group discussions, and small group problem-solving sessions.

Subjects

The data in this report were collected during the second phase of the TAT evaluation project (January–April, 1977). A total of 80 teaching assistants participated in the study. Seven of the 80 were undergraduates. All the undergraduate TAs were from the Department of Computing Science, which does not offer graduate training. The TAs who participated in the TAT instruction were selected by their respective departments with the stipulation that they had no previous teaching experience and that they possessed a sufficient mastery of oral English to enable them to contribute to and benefit from the seminars. Neither of these criteria was adhered to strictly. Consequently, 28 of the participants had previous teaching experience with a minimum defining criterion of at least one year of teaching

at any level of schooling prior to becoming a TA, and approximately 10%–15% had difficulty speaking English.

Instruments and Data Sources

The dependent variables examined to determine the effectiveness of the TAT program were perceptions and attitudes of TAs toward a variety of instructional dimensions, student perceptions of TAs, TA attitudes toward the TAT program, and student achievement outcomes.

The *Teaching Assistant Inventory* (TAI), a 28-item paper-and-pencil questionnaire, was designed to test seven hypotheses. Four items on each of seven subtests related to TA's 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 .29 to .83. One item relating to TAs' attitudes toward their students was dropped from the second subtest because in a previous investigation (Marx, Martin, Ellis, & Hasell, 1978) it did not contribute substantially to the subtest score. Internal consistency coefficients were low on the first. second, and fifth subtests in the TAI; however, to permit comparisons between the two reports, these data were included in the analyses.

A second paper-and-pencil questionnaire, the *Student Inventory* (SI) was designed to test 10 hypotheses concerning students' perceptions of the teaching of their TAs and 4 hypotheses concerning their own attitudes. The 10 hypotheses concerning perceptions focused on: clarity of objectives, correspondence between 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 his/her own teaching. The 4 hypotheses concerning attitudes related to student satisfaction with their TA, attitude toward the academic discipline of the course, attitude toward the course, and attitude toward their own learning. In addition, two cluster scores were formed, one from items contributing to the former 10 subtests and one from items contributing to the latter 4 subtests. The total number of items on the SI was 51; the number of items on specific subtests ranged from two to nine. Internal consistency coefficients ranged from .44 to .87, and were .63 and .78 for the two cluster scores, respectively.

Precise information vis-à-vis TA attitudes toward the various elements of the TAT program was obtained through the *Teaching Assistant Personal Interview* (TAPI). The TAPI included 16 "yes-no" questions concerning orientation to the TAT program, time demands of the program, program content, program instruction, etc., and 14 open-ended questions which probed reasons behind the responses given to the "yes-no" items. The TAPI was administered in an attempt to solicit specific reactions and suggestions which would have utility in improving the program in future semesters.

All interviews were conducted in the two weeks following the conclusion of the TAT sessions, by a research assistant (who had not attended the sessions) employing a standard interview format.

Finally, end-of-term student achievement data were obtained from eight multisection courses which employed both TATs and non-TATs.

Design

The 80 TAs in the study came from six distinct groups. Four of these groups (A, A′, B, D) were participants in the current TAT program, one (C) was a control group which did not receive instruction, and one (E) was a group made up of TAs who had participated in an earlier version of the TAT program (in September–October 1975) prior to the commencement in September 1976 of the current evaluation project.

Of the four TAT groups, A and A′ received instruction during the first half of the fall semester, 1976. Group B participated in the program during the second half of that same semester, and group D attended the TAT sessions during the first half of the spring semester, 1977. Groups A, A′, and B included only those 1976 TAT participants who had TA assignments in the spring 1977 semester. All the data reported in this study, with the exception of the TAPI results, were collected at the end of the 1977 spring semester. Consequently, groups A, A′, and B are not precisely the same groups as those discussed in the first report of the TAT Program (Marx, Martin, Ellis & Hasell, 1978), owing to attrition from those original groups.

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 discipline areas from which the TAs came. Demographic data were collected from all the TAs at the University, and subjects in group C were randomly selected from this pool, stratified by previous

Table 1 / Number in Sample by Group, Previous Teaching Experience, and Discipline

	Group													
	A		A′		B		D		C		E		Total	
Discipline	E	NE	E	NE	E	NE	E	NE	E	NE	E	NE	E	NE
Arts[a]	1	7	0	1	5	6	4	5	3	5	3	6	16	30
Science[b]	2	2	3	6	2	3	1	4	4	5	0	2	12	22
Total	3	9	3	7	7	9	5	9	7	10	3	8	28	52

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

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

b. Departments of Biological Science, Chemistry, Computing Science, Kinesiology, and Physics.

teaching experience and department. Group C participants received the same semester wage as did the TAT participant; however, the TAs in group C carried a full teaching load as compared to TAT participants who had their teaching loads reduced by 20% to 25%.

The TAI was administered to all the TA groups, with the items pertaining to TA attitudes toward the TAT program being omitted from the forms given to TAs in group C. The SI was administered by TAs to the students in one of the tutorial or laboratory classes taught by each TA in the sample. All TAs were requested to read a standardized set of administration instructions to their students, and to limit their remarks to that set of instructions. The TAPI was conducted with each of the TAT participants in groups A, A′, B, and D within two weeks following the TAT instructional periods in which they were involved.

The collection of achievement data did not adhere to the overall design of the study since non-TATs in the eight multisection courses from which data were obtained did not coincide with TAs in group C. Therefore, the control TAs for achievement comparisons had different demographic characteristics than did TAs in group C. The most critical difference was that they had more experience as TAs in the University.

RESULTS

Two separate sets of factorial analyses of variance were performed on the seven subtests on the TAI. TAI protocols were returned by 73 of the 80 subjects, with missing data distributed roughly proportionately across groups. For the first set of analyses, one factor was group, with six levels (A, A′, B, D, E, and C), and the second factor was previous teaching experience, with two levels. (Since the control group did not respond to the attitude toward the TAT program subtest, this particular analysis of variance was a 5×2 design.) There was only one significant F ratio in all of these analyses. This was a previous teaching experience effect on the attitude toward students' subtest ($F_{1,60} = 4.37, p < .05$), favoring TAs without previous teaching experience ($M = 4.10, SD = .44$) over the group with previous experience ($M = 3.84, SD = .62$).

The second set of analyses contrasted groups (six levels, except for the attitude toward the TAT program subtest) and academic discipline (science and arts). Since for this design there was one cell with only one subject (A′ — arts), the interaction term was pooled with residual variance for the error term. Only one F ratio was significant in these analyses. This was an academic-discipline effect on the attitude-toward-improving-teaching subtest ($F_{1,61} = 8.19, p < .01$), and favored arts TAs ($M = 3.85$, $SD = .45$) over science TAs ($M = 3.54, SD = .49$).

To explore further the group differences on the TAI, a set of a priori orthogonal contrasts were conducted, weighted for unequal cell size, comparing all of the TAT groups with the control group, and then comparing the TAT groups with each other. An additional contrast, not orthogonal

to the others, compared groups A, A′, and D with B, to ascertain if there was an effect for the timing of the program. (Groups A, A′, and D participated in the first half of the semester, group B participated in the second half.) The means, standard deviations, and t values of all six groups on the various TAI subtests (for which there was at least one significant t test on the a priori contrasts) are reported in table 2.

On three of the TAI subtests, significant heterogeneity of variance was encountered, as measured by three tests (Bartlett-Box, Cochran's C, F Max). Although these tests are problematic when individual variances are based on small sample sizes (Guilford, 1965), the t tests for the orthogonal contrasts on these subscales were based on separate variance estimates.

The data shown in table 2 indicate that the TAs who participated in the beginning of the semester had more favorable attitudes toward their own learning, their academic discipline, and the Faculty of Education one semester after completing the program. These results are essentially the same as those from the non-orthogonal contrast, including group D, whose responses came a half-semester after completion.

Scores on the seven TAI subtests were intercorrelated to ascertain the degree to which they measured differentiable attitudes. The 21 intercorrelations ranged from .00 to .64, with a median of .31. The two highest intercorrelations were between attitudes toward the Faculty of Education and the TAT program ($r = .64$, $p < .01$), and attitudes toward the Faculty of Education and the TAs' own academic disciplines ($r = .57$, $p < .01$). These were followed in magnitude by correlations between attitudes toward teaching and students ($r = .48$, $p < .01$) and between the TAs' own academic discipline and their own learning ($r = .44$, $p < .01$). The remaining intercorrelations did not exceed .40.

The TAPI was administered to a randomly selected subsample of approximately two-thirds of the participants in groups A, A′, B, and D. This resulted in a sample of 38 TAs ($n_A = 11$, $n_{A'} = 8$, $n_B = 9$, $n_D = 10$). Thirteen items on the TAPI can be used to judge TA satisfaction with the instructional and logistic aspects of the program. The responses to 12 of these 13 items were strongly in favor of the program, with chi-square at $p < .05$. The only item which was not strongly favorable asked the TAs if three hours a night, once a week, was a reasonable amount of time to spend on the program. Most of the TAs who responded negatively would have preferred to spend fewer hours per session over a longer period. Also the students from the Faculty of Science were generally satisfied with the 3-hour meetings, while the non-science students generally stated preference for shorter meetings.

One item on the TAPI asked the TAs if their respective departments were doing anything to assist them in their TA training. Only six of the 38 TAs responded affirmatively, and only two of these considered their department's efforts to be related to the TAT program.

The four groups were compared on their responses to an item asking if

Table 2 / Descriptive Statistics and Significant A Priori Contrasts for Groups on the TAI

Attitude toward		Group A (n = 11)	A′ (n = 9)	B (n = 15)	D (n = 14)	E (n = 9)	C (n = 15)	Significant A Priori Contrasts[a]													
Teaching	M	3.52	3.94	3.62	3.75[b]	3.86	3.77	A	A′	B	D	E	C	t = 1.95†							
	SD	.58	.36	.46	.60	.36	.43	1	−1	0	0	0	0								
Academic discipline	M	4.45	4.44	4.06	4.29	4.11	4.28	A	A′	B	D	E	C	t = 1.89†	A	A′	B	D	E	C	t = 1.79†
	SD	.65	.53	.75	.54	.55	.51	1	1	−2	0	0	0		1	1	−3	1	0	0	
Learning	M	4.48	4.39	4.16	4.40	4.33	4.35	A	A′	B	D	E	C	t = 1.88†	A	A′	B	D	E	C	t = 2.00*
	SD	.47	.44	.55	.31	.38	.39	1	1	−2	0	0	0		1	1	−3	1	0	0	
Faculty of Education	M	4.08	3.78	3.53	3.54	3.53	3.43	A	A′	B	D	E	C	t = 1.87†							
	SD	.55	.40	.85	.72	.38	.52	1	1	−2	0	0	0								

a. Two-tailed tests, df = 67, except on the teaching subtest, df = 66.
b. n = 13.
†$p < .10$; *$p < .05$.

the program's content was what they expected. The chi-square associated with the eight-cell contingency table categorizing responses as yes or no by groups was 10.17, $df = 3, p < .05$. This resulted primarily from TAs in groups B and D responding negatively. However, when they were probed further, their general agreement was that their expectations were initially low, and the content of the program was much better than expected.

The SI was sent to 77 TAs who had teaching assignments involving responsibility for one or more groups of students. The remaining participants were either no longer teaching or were working as consultants on a drop-in basis (e.g. assisting undergraduate computing science students with errors in their computer programs). Data were returned by 68 of these 77 students, and, as with the TAI, missing data were roughly proportionate across groups. The data used for analysis were the mean subtest scores computed across the students responding to the SI within each TA's respective tutorial group.

Two separate sets of two-factor (6×2) analyses of variance were computed. As with the TAI, these were grouped by previous teaching experience, and grouped by academic discipline (science or arts). The interaction term was pooled with the error term in the latter analysis owing to one cell having only one subject. There were no significant effects for previous teaching experience over all of the SI subtests. However, there were three significant differences on the SI as a function of academic discipline. These are reported in table 3. All three results favored the science TAs over the arts TAs.

One significant omnibus F ratio for the group effect was on student perceptions of the willingness of TAs to improve their teaching ($F_{5,56} = 2.60, p < .05$). To further explore the group main effect, a set of a priori

Table 3 / Significant Analysis of Variance Effects for the SI Comparing Arts and Science TAs

Subtest		Group: Science ($n = 28$)	Group: Arts ($n = 41$)	F ($df = 1/62$)
Objectives and evaluation	M	3.50	3.37	3.41†
	SD	.36	.40	
Instructional materials	M	3.09	2.84	6.22*
	SD	.51	.59	
Attitude, discipline	M	3.79	3.64	2.88†
	SD	.35	.38	

†$p < .10$; *$p < .05$.

orthogonal contrasts, weighted for unequal cell size, were performed. The results of these analyses on contrasts for which there was at least one significant t test are reported in table 4, along with the means and standard deviations for the six groups. There were no significant differences on any SI variable for contrasts comparing groups A, A′, and B with D, or groups A and A′ with B. Similarly, a contrast not orthogonal to the others comparing A, A′, and D against B to test for effects due to timing of the course yielded no statistically significant t tests.

As table 4 shows, the students of TAT participants saw their respective TAs as being better tutorial leaders and as being more willing to improve their teaching than did the students of control TAs. Responses to the specific items which made up this particular subscale indicated that the students of TAT participants perceived their TA as attempting to elicit critical feedback from students about her/his teaching; and, upon receipt of such feedback, was open to exploring alternative instructional strategies and formats. Also, the students of TAs who participated in TAT during 1976/77 thought their TAs were more satisfied with their own teaching than did the students of TAs who had participated in 1975/76. Finally, for the two TAT groups which received instruction at the same time, group A participants were seen as exhibiting more effective teaching than the TAs in group A′.

The 14 independent SI subtests, excluding the two cluster scores, were intercorrelated to assess the students' ability to differentiate the areas of perception and attitude measured by the instrument. These intercorrelations ranged from –.04 to .82, with a median of .46. The subtests which were most highly correlated, all at $p < .01$ (student attitude toward the academic discipline of the course and attitude toward the course itself, $r = .81$; perception of quality of TA leadership and satisfaction with TA, $r = .81$; perception of TA satisfaction with teaching and perception of TA attitude toward his/her own academic discipline, $r = .82$), indicated that the students had difficulty differentiating their attitudes about and satisfaction with the course from similar views of the TA. Similarly, the intercorrelation between the two cluster scores was .78, suggesting that student perceptions of the quality of a course and their attitude toward the course are nearly synonymous.

The 14 SI subtests were intercorrelated with the 7 TAI subtests. The correlations ranged from –.31 to .26, with a median of .05. Thus, while the two measures were essentially unrelated, the two highest positive correlations are worth noting even though they were only .26 ($p < .05$). These correlations were between the TAs' attitudes toward their own disciplines and their respective students' perceptions of these attitudes, and between the TAs' willingness to improve their teaching and their respective students' perceptions of this willingness.

The student achievement data from the eight multisection courses were analyzed separately for each of the courses. The data used were the final

Table 4 / Descriptive Statistics and Significant A Priori Contrasts for Groups on the SI

Subtest		Group A (n = 10)	A′ (n = 8)	B (n = 14)	D (n = 13)	E (n = 10)	C (n = 14)	Significant A Priori Contrasts[a]
Objectives and Evaluation	M	3.70	3.36	3.43	3.32	3.37	3.40	A A′ B D E C
								t = 1.91†
	SD	.34	.42	.41	.34	.21	.48	1 −1 0 0 0 0
Tutorial Leadership	M	3.99	3.81	3.88	3.81	3.79	3.67	A A′ B D E C
								t = 1.72†
	SD	.27	.26	.45	.38	.19	.46	1 1 1 1 1 −5
TA Satisfaction	M	3.89	3.66	3.74	3.72	3.54	3.67	A A′ B D E C
								t = 2.02*
	SD	.26	.45	.29	.25	.27	.31	1 1 1 1 −4 0
Attitude, Students	M	4.17	3.88	4.06	4.06	4.06	3.94	A A′ B D E C
								t = 2.13†
	SD	.25	.32	.26	.18	.16	.44	1 −1 0 0 0 0
Attitude, Learning	M	3.94	3.56	3.71	3.57	3.60	3.63	A A′ B D E C
								t = 1.99†
	SD	.49	.31	.37	.26	.10	.34	1 −1 0 0 0 0
Improve Teaching	M	3.45	3.08	3.38	3.32	3.44	3.09	A A′ B D E C
								t = 2.47*
	SD	.31	.43	.20	.43	.31	.26	1 1 1 1 1 −5
Total: Teaching	M	3.73	3.50	3.60	3.54	3.54	3.50	A A′ B D E C
								t = 2.03*
	SD	.19	.27	.25	.25	.10	.29	1 −1 0 0 0 0

A A′ B D E C
1 −1 0 0 0 0

t = 2.39*

a. Two-tailed tests, df = 63 (except on the Attitude, Learning subtest, where separate variance estimates are used, df = 15). †$p < .10$; *$p < .05$.

course grades as submitted to us by the respective professors. In most cases part of the grade was the result of a TA marking of her/his own students, thus introducing some confounding into the data. However, there was always a component of objectively scored tests or of balanced marking of student papers across all TAs teaching in the course. In all cases, a one-way analysis of variance was computed within each course over all tutorial groups. This was followed by a weighted means analysis using the error term from the analysis of variance which compared groups taught by TATs with groups taught by non-TATs. Only one of the eight weighted means analyses showed significant differences, this favoring the TAT over his two untrained colleagues ($t = 2.45, p < .05$).

DISCUSSION

The findings of this study, considering the results on SI, TAI, TAPI, and student achievement variables, support the effectiveness of the TAT program in improving the quality of TA instruction. Perhaps most encouraging are the SI results which indicate that not only are TAT participants perceived to be better tutorial leaders by their students, but that they are also perceived as being more willing to improve their teaching than are non-TAT control subjects. If student perceptions have any predictive validity, these findings should argue well for longer-term as well as relatively short-term effects of the TAT program. The overall treatment effects disclosed by the study are particularly noteworthy given the relatively short duration (15 hours) of the instruction. While it is not necessarily the case that a more extensive training program would accentuate such effects, it seems reasonable to argue that such programs can yield effects which easily rationalize the costs they may incur.

The potential efficiency of TA instructional programs is further supported by the student achievement data collected as part of the present evaluation project. While in only one of the eight multisection classes surveyed were significant differences found between the achievement of students taught by TATs and the achievement of students taught by non-TATs (the one case favoring the TAT group), it must be remembered that these groups of TAs differed greatly in the number of TA semesters they had taught previously. From demographic data collected from the TAs in the achievement data sample, it was found that the mean number of semesters taught by TAT participants prior to September, 1976, was .6, whereas a comparable figure for the non-TAT tutorial leaders was 3.57. (The levels of the previous teaching experience factor used in the analyses reported in the results section of this paper were determined by the criterion "one full year of instructional experience *prior to* becoming a TA." Consequently, even for the TAI and SI data collected under the main design of the study, the non-TAT sample had more actual TA experience than did the TAT samples.) It would thus seem that where it matters most – i.e. the actual learning and achievement of undergraduate students – the TAT program

matches or even surpasses the informal acquisition of teaching competence as a result of experience in leading tutorials. In short, by the most rigorous measure of teaching effectiveness, participation in the TAT program is at least equivalent in instructional value to an average of approximately three full semesters in the "school of hard knocks." The idea that it is possible to directly teach individuals to teach appears not to be as "intuitively fallacious" as is sometimes implied (Bonham, 1976).

The findings of this evaluation are also informative with respect to some logistic decisions which arise in the offering of such programs. From table 2 it is quite clear that instructional programs of this variety are better offered at the beginning, rather than the middle or end, of the semester during which participating TAs have actual tutorial responsibilities. More generally, it appears that if a program has any validity it should occur before or early in a participants' "instructional career." It may well be that the less-than-optimal practices acquired through experiential trial and error tend to resist alteration and perhaps impede the learning of more reasonable approaches to instruction.

A second decision, concerning the differentiation of instructional programs for TAs according to discipline, obviously deserves further investigation. On several of the variables examined in this study, science and arts TAs differed considerably (see table 3 for some examples). While the interaction between different curriculum areas and various instructional techniques has not, to date, been definitively addressed (Rosenshine, 1976), it seems likely that programs like TAT should pay more attention to discipline differences. The open-ended responses to the TAPI contained several suggestions which might be useful in this regard. One of these, which has already been incorporated into the design of future TAT programs, calls for a more detailed examination of techniques appropriate for conducting laboratory classes. In the future, science TAs with laboratory assignments will be able to study such techniques, while discussion-oriented tutorial leaders will pursue a more detailed examination of techniques appropriate to their situations.

Finally, this second report offers some useful, although slightly inferential, information vis-à-vis student ratings of instructors and courses. Generally, these findings tend to support recent conclusions by McKeachie and Kulik (1975) and Wittrock and Lumsdaine (1977). Intercorrelations among SI variables indicated that students in this study did not discriminate between the perceived quality of instruction in a course and the content of the course. While recognizing that the interpretation of such correlations can be hazardous, it seems unlikely that TAs in all "high-interest" courses were equally competent, or that low student interest was always associated with poorly conducted tutorials and laboratories. The fact that student satisfaction with TA and student perception of TA leadership quality were so highly correlated lends credence to the Wittrock and Lumsdaine view that student ratings of teacher effectiveness are based

more on globular TA characteristics than they are upon specific instructionally related skills and abilities.

To summarize, data concerning teaching attitudes and undergraduate student attitudes, perceptions, and achievement positively support the effectiveness of the Simon Fraser TAT Programme in improving the instructional effectiveness of university teaching assistants. It appears that such programs are possible with little strain on the university purse-strings, and that the cost-effectiveness ratio of programs like TAT is likely to be favorable. Further experimentation is required to determine which of a variety of instructional strategies (small group discussions, demonstrations, individual counselling, lectures, microteaching, etc.) are most effective in promoting instructional competence across the wide range of teaching roles which TAs perform.

NOTE

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

REFERENCES

Bonham, G. W. About this special report. *Change: The Magazine of Higher Learning*, 1976, *9*(1), i.

Guilford, J. P. *Fundamental statistics in psychology and education* (4th ed.). New York: McGraw-Hill, 1965.

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*, 1978, in press.

Marx, R. W.; Martin, J.; Ellis, J. F.; & Hasell, J. Improving the instructional effectiveness of university teaching assistants: Report I. *Canadian Journal of Education*, 1978, *3*, 1–12.

McKeachie, W. J., & Kulik, J. A. Effective college teaching. In F. N. Kerlinger (Ed.), *Review of research in education, vol. 3*. Itasca, Ill.: Peacock, 1975.

Minister increases program funding. *Ontario Universities Program for Instructional Development Newsletter*. March 1977, *15*, 1.

Notar, A. Teaching and learning: An individualized course for instructors in higher education. *Canadian Society for the Study of Education News*, 1977, *3*(7), 14–15.

Rosenshine, B. Recent research on teaching behaviors and student achievement. *Journal of Teacher Education*, 1976, *27*, 61–65.

University of Toronto offers course in "theory of tutoring" for teaching assistants, science tutors, and laboratory supervisors. *Ontario Universities Program for Instructional Development Newsletter*, October 1976, *12*, 7.

Wittrock, M. C., & Lumsdaine, A. A. Instructional psychology. *Annual Review of Psychology*, 1977, *28*, 417–459.

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