Jurnal Pendidik dan Pendidikan, Jil. 21, 115–128, 2006

EFFECTIVENESS OF INTERACTION ANALYSIS FEEDBACK ON THE VERBAL BEHAVIOUR OF PRIMARY SCHOOL MATHEMATICS TEACHERS

See Kin Hai[*] and Lim Siew Bee
Department of Science and Mathematics Education, Education Faculty,
Universiti Brunei Darussalam

Abstract: This study attempted to investigate the relative effectiveness of interaction analysis feedback on the verbal behaviour of teachers teaching mathematics in Primary 5 classes of four randomly selected primary schools in Brunei-Muara district. It also attempted to investigate the effects of the feedback system on pupils' attitude towards mathematics and their academic achievement in mathematics. The sample used for the study consisted of eight primary school teachers teaching mathematics subject. These eight teachers were preselected and were divided into feedback (experimental) and non-feedback (control) groups. One teacher from each of the four schools was selected as the experimental group and the other teacher from the same school was picked as the control group. One hundred and fifty two pupils of average ability from four urban schools were involved. To examine the extent to which the feedback system was effective, a modified Flanders Interaction Analysis Categories System (FIACS) was used to record classroom communications and the results provided as a feedback to the experimental group. Results showed that the feedback groups accepted students feelings more, praised students more, used students' ideas and initiated more student talk in the classroom. Effects of the feedback were encouraging with higher student academic achievement and more favourable attitudes after teachers were given feedback.

Abstrak: Kajian ini cuba menyiasat tentang keberkesanan maklum balas analisis interaksi terhadap kelakuan verbal guru-guru yang mengajar matematik Darjah 5 dalam empat buah sekolah rendah dalam kawasan bandar di daerah Brunei-Muara. Kajian ini juga menyiasat tentang kesan sistem maklum balas ini terhadap kelakuan dan pencapaian murid-murid dalam subjek matematik. Sampel yang digunakan terdiri daripada lapan orang guru yang mengajar matematik. Mereka dibahagikan kepada empat orang dalam kumpulan kawalan (yang tidak menerima maklum balas) dan empat orang dalam kumpulan eksperimen (yang menerima maklum balas). Seramai 152 orang murid dalam tahap kebolehan sederhana dipilih untuk kajian ini. Instrumen *Flanders Interaction Analysis Categories System* (FIACS) yang telah diubahsuaikan dan diguna untuk mencatat komunikasi di dalam bilik darjah. Catatan ini digunakan sebagai maklum balas kepada kumpulan eksperimen. Hasil kajian ini menunjukkan bahawa guru-guru dalam kumpulan eksperimen menerima perasaan murid, memuji, menggunakan idea-idea murid dan memulakan komunikasi murid. Murid-murid dalam kumpulan eksperimen juga mencapai keputusan dan kelakuan yang lebih baik daripada kumpulan kawalan.

[*] Corresponding author: khsee@shbie.ubd.edu.bn

INTRODUCTION

The use of direct classroom observation over the last century has resulted in the accumulation of an impressive body of information about the nature of effective teaching (Brophy & Good, 1986; Good & Brophy, 2000). Types of classroom communication have a significant impact on student outcomes (Wang, Haertel, & Walberg, 1993). Results of studies indicate that teachers' classroom verbal behaviour affect students' achievement (Good & Brophy, 2000). In fact, students' opportunity to participate actively in the classroom communication contributes to one of the most important predictors of student achievement (Berliner & Biddle, 1995). However, students' opportunity to participate in the classroom communication may vary with different verbal behaviours of teachers, with their achievement and attitude (Allington, 1991; Good & Weinstein, 1986), and gender (Sadker & Sadker, 1994; *Houston Chronicle,* 2001).

This study intends to complement the studies done in other countries which have been devoted to use FIACS to help teachers develop and control their teaching behaviour, to explain the variations which occur in the chain of classroom events and their relationships with classroom communications and educational outcomes. The study intends to answer the following questions:

i. Do teachers who receive feedback show greater change in the 10 categories of FIACS (i.e., accepts pupils' feeling, praises pupils, accepts or uses pupils ideas, asks questions, lectures, giving directions, criticizing, pupils' response, pupil initiates talk, silence or confusion) than those who did not receive feedback?

ii. Are there any significant differences in pupil achievement and attitudes towards mathematics between the experimental and control groups?

Amidon and Powell (1967), Campbell and Barnes (1972), Kantowaski (1977), and Gorard (2000) used FIACS in their studies and discovered that teachers who were perceived as effective engaged largely in accepting students' feeling and ideas, used more praise and encouragement in their classroom communication.

Flanders (1970) investigated the effects of FIACS feedback on the verbal behaviours of teachers found that teachers who received feedback differed significantly in their use of certain verbal behaviours from those who did not receive feedback. Teachers who received feedback were found to use more praise, accept and clarify student ideas more, use more indirect talk, use more positive reinforcement after teacher-initiated student talk, use less corrective feedback, criticise students less, ask more questions, use less lecture method, give

fewer directions and less teacher-initiated talk. Studies by Kline and Sorge (1974), Younger, Warrington, and Williams (1999) have shown that even teachers who were not trained in the mechanics of interaction analysis will change their classroom verbal behaviours as a result of feedback from the interaction analysis. Findings from Swann and Graddol (1988), and Younger and Warrington (1996) have implied that teachers' classroom verbal behaviours could affect significantly primary pupils' achievement in mathematics and their attitude towards the subject.

METHOD

Sample

The subjects involved in this study were eight primary school teachers teaching mathematics in Primary 5 classes of four randomly selected primary schools in Brunei-Muara district. The teachers were selected on the basis of having been trained and with at least three years of teaching experience. Two teachers from each school were chosen with one as the experimental group and the other as control group. C_1 represents the teacher together with the class pupils from school 1 as the control group and E_1 represents the respective experimental group. Hence, the control groups are represented by C_1, C_2, C_3 and C_4 and the experimental groups are represented by E_1, E_2, E_3 and E_4. A total of 152 students (83 females and 69 males) of average ability from four urban schools participated in this study.

Instrument

FIACS adapted by Hopkins and Moore (1993) was used in this study to record teachers' verbal behaviour in classroom communication. Table 1 showed the FIACS used consisted of 10 categories: seven of these categories reflected teacher talk (category 1 to 7), two reflected pupil talk (category 8 and 9) and the last category indicated silence or confusion. The category system attempts to include as many type of verbal communication as possible between teacher and pupils in the classroom. A questionnaire consisting 20 items of Likert scale was used to measure students' attitude towards mathematics (Mathematical Attitude Scale, MAS). Each item was evaluated on a 5-point scale. The score ranges from 1 (strongly disagree) to 5 (strongly agree). However, for negative items the scoring was reversed. Test retest reliability coefficient (Pearson's correlation) of 0.87 ($p < 0.05$, 2-tailed) was obtained by giving the same questionnaire with a lapse of two weeks to 152 students at the end of second school term. Reliability analysis for this instrument produced a Cronbach alpha of 0.83. Achievement in mathematics was measured using common tests (pre- and post-test for Units 9, 10

and 11) developed by the teachers with the help of the researcher. Difficulty index for the test items range from 0.3 to 0.7. The Pearson correlation of 0.86 ($p < 0.05$) and Cronbach alpha of 0.79 was considered high for the test reliability. Content or face validity of the items was determined by the teachers taking part in the experiment to check the items in the test together with the researcher.

Table 1. Flanders Interaction Analysis System

Teacher/Student/Other Behaviors Observed			Tallies	Anecdotal notes
Teacher talk	Indirect influence	1. **Accepts Feeling:** Accepting and clarifying the feeling tone of students in a nonthreatening manner. Feelings may be positive or negative. Predicting or recalling feelings is included.		
		2. **Praises or Encourages:** Praising or encouraging student action or behavior. Jokes that release tension, but not at the expense of another individual; nodding head, saying "um hm?" or "go on" are included.		
		3. **Accepts or Uses Ideas:** Clarifying, building, or developing ideas suggested by a student. As more of the teacher's own ideas come into play, shift to Category 5.		
		4. **Asks Questions:** Asking a question about content or procedure with the intent that a student answers.		
	Direct influence	5. **Lectures:** Giving facts or opinion about content or procedures; expressing the teacher's own ideas, asking rhetorical questions.		
		6. **Gives Directions:** Giving directions, commands or orders with which a student is expected to comply.		
		7. **Criticisms or Justifies Authority:** Making statements intended to change student behavior from unacceptable to acceptable pattern; bawling out someone; stating why the teacher is doing what he/she is doing; extreme self-reference.		
Student talk		8. **Responds:** Talk by student in response to teacher. Teacher initiating the contact or solicits student statement.		
		9. **Initiates:** Talk by students, which they initiate. If "calling on" students is only to indicate who may talk next, observer must decide whether student wanted to talk. If so, use this category.		
Silence		10. **Silence or Confusion:** Pauses, short periods of silence, and periods of confusion in which communication cannot be understood by the observer.		

Procedure and Statistical Analysis

The pre- and post-tests for mathematics achievement were administered to students in the control and experimental groups at the start and end of the second school term. Tests ta-9 and TA-9, ta-10 and TA-10, ta-11 and TA-11 were pairs of parallel tests that were administered to all pupils as pre- and post-tests. Performance in three selected topics of the Primary 5 mathematics were measured by the three post-tests, namely TA-9, TA-10 and TA-11. The pre-test was adjusted for comparison purposes. MAS was administered at the end of second term. While teachers in the control group in each school were given 15 minutes briefing to explain that observations of their teaching would be made at certain times during the second term without any adverse reports made about them, their experimental counterparts were given a longer period of induction.

Informal sessions were conducted for the experimental group to explain to them the various categories, the different behavioural effects that might affect student reception, attitudes and achievement. However, it was not possible to give immediate feedback on the same day to the teachers. As the data were collected from observing classroom dynamics, analysis of the data were not done at the end of the lesson, thus, the delay in providing feedback was inevitable. Anyway, the six teachers in the experimental group received the FIACS feedback form and had a discussion with the researcher between one to three days after the observation.

To reduce extraneous factor in the data, all classes were observed simultaneously by two observers at the beginning of the second week of second term. The two observers sat at the back of the classroom in a position which allowed them to hear and see all participants (teacher and pupils). Each lesson lasted 30–60 minutes. The observers began coding the interactions that occurred approximately five minutes after the lesson began. At an interval of 15 seconds, the observers would decide which of the categories of FIACS best represented the interacting events just completed and then coded and recorded them into the observation form as shown in Table 1. Inter-observer agreement was computed based on Scott's formula and also by comparing their tallies on an item-by-item basis. Scott's formula $\pi = \frac{P_o - P_e}{100 - P_e}$ where $P_e = \sum_{i=1}^{k} P_i^2$ (P_i = proportion of tallies falling into each category; P_o= proportion of agreement; k = number of categories; P_e = proportion of agreement expected by chance; π = amount that two observers exceeded chance agreement divided by the perfect agreement that exceeded chance). Scott's coefficient obtained from pairs of observers for each observation made in this study was found to be satisfactory as indicated in Table 2.

Table 2. Scott's coefficient for pairs of observers using FIACS

Class	Observation number		
	1	2	3
C1	0.856	0.879	0.901
C2	0.926	0.921	0.875
C3	0.868	0.931	0.896
C4	0.917	0.873	0.852
E1	0.916	0.837	0.938
E2	0.962	0.917	0.912
E3	0.927	0.932	0.894
E4	0.835	0.938	0.868

Table 3. Inter-observer reliability for pairs of observers using FIACS

Class	Observation number		
	1	2	3
C1	0.823	0.861	0.890
C2	0.902	0.911	0.866
C3	0.676	0.895	0.876
C4	0.905	0.831	0.817
E1	0.876	0.810	0.878
E2	0.954	0.903	0.883
E3	0.917	0.921	0.835
E4	0.811	0.909	0.854

As for item-by-item analysis (Bailey, 1975), the inter-observer reliability = A/(A+D), where A means total agreement of the two observers and D means total disagreement among the two observers. Table 3 shows the inter-observer reliability for all observations that was never below 0.82 for most of the cases except the class C_3. Statistical Package for Social Science (SPSS) program was used to analyse the data.

RESULTS AND DISCUSSION

Graphical Representation of the Teaching Behaviours of the Experimental and the Control Groups

Different types of variables were plotted according to the three observations done. For the experimental group (E), the first observation was done without any feedback given while for the other two, observations were taken after feedback had been given to the teachers. For the control groups (C), three observations were taken without any feedback given to the teachers. The three observations were plotted side-by-side to allow for comparison. The behaviour categories and ratio variables were plotted graphically below.

For the first category (Fig. 1), there was a marked difference before and after feedback was given to the experimental group as compared with the control group where the percentage increase was minimal (0.2%–0.3%).

Figure 2 showed a marked increase in this reinforcing behaviour for the E group after feedback was given. For the C group, the increase was not distinct. The result suggested that the experimental teachers tend to involve pupils more in the instructional process through positive reinforcement with an increase use of praise.

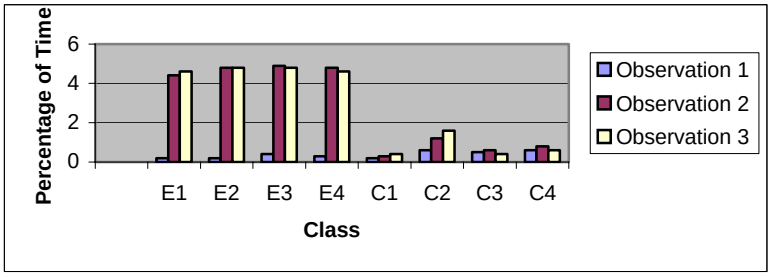

Figure 1. Percentage of time teacher accepts pupils' feeling (Category 1)

Figure 2. Percentage of time teacher praises or encourages pupils (Category 2)

A marked distinction was shown between the E and C groups in the use of this behaviour (Fig. 3). This behaviour was particularly important in influencing pupils' academic achievement. E group showed a marked increased of more than 3.9 percent after feedback was given.

Figure 4 showed an increase in the use of asking questions for all the groups (E and C groups). It is also clearly shown that teachers in the E group spent more time in asking questions in the classroom than teachers in the C group.

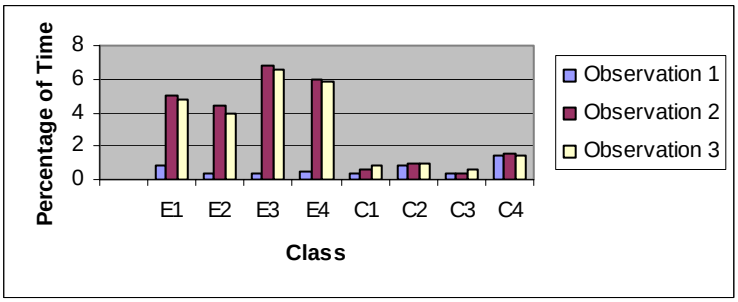

Figure 3. Percentage of time teacher uses pupils' ideas (Category 3)

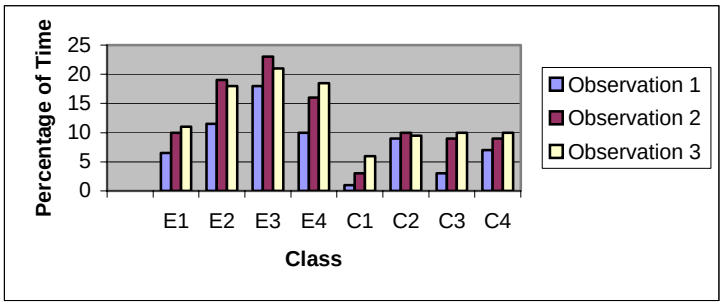

Figure 4. Percentage of time teacher asks questions (Category 4)

From Figure 5, it was distinct that there was no increase in the usage of lecture in this category in the E group. There was a tendency for the use of lecture method to decrease after the feedback and they tended to concentrate on other reinforcement behaviours. This characteristic was not shown in all the control subgroups.

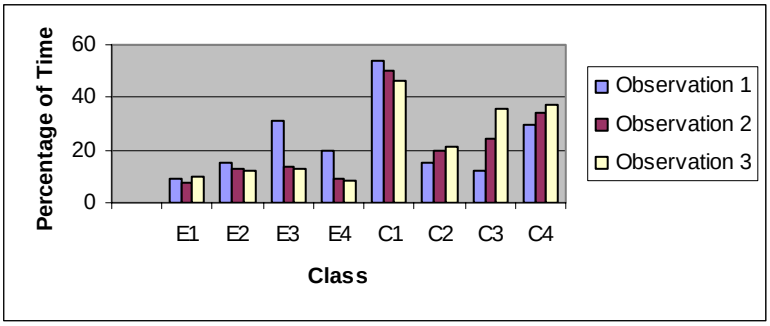

Figure 5. Percentage of time teacher lectures (Category 5)

As shown in Figure 6, the greatest difference between the C and E groups was a definite decline in the use of this category after feedback was given to the E group. However, the C group did not show much change in their behaviour in this category.

Figure 7 showed a minimal behaviour in Category 7. There was not much difference between E and C groups except for E_1 and C_2. This low percentage was due to the teachers' awareness that they were being observed.

From Figure 8, it was clear that there was an overall increase for the E group except for E_3, where this verbal behaviour was initially high and remained almost the same. However, the general trend was a decrease from higher to lower amount of pupil talk. Feedback seemed to be effective to improve this behaviour.

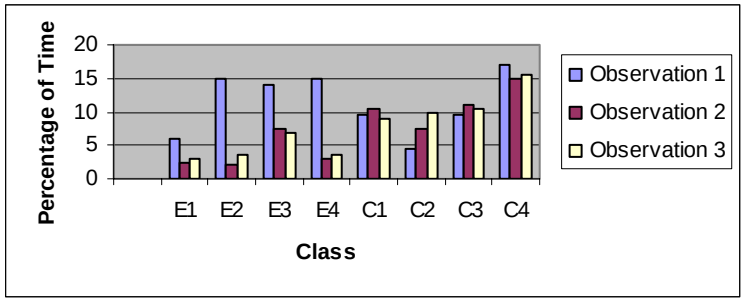

Figure 6. Percentage of time teacher giving directions, orders or commands (Category 6)

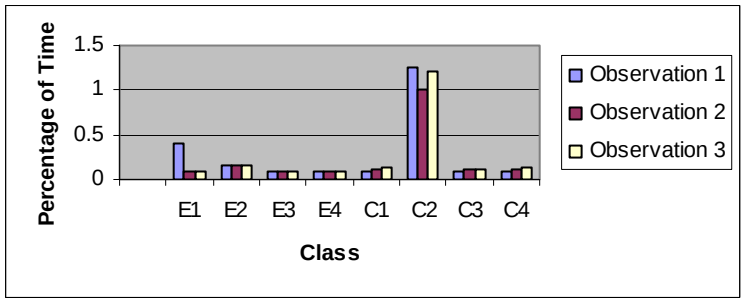

Figure 7. Percentage of time teacher criticises pupil behaviour (Category 7)

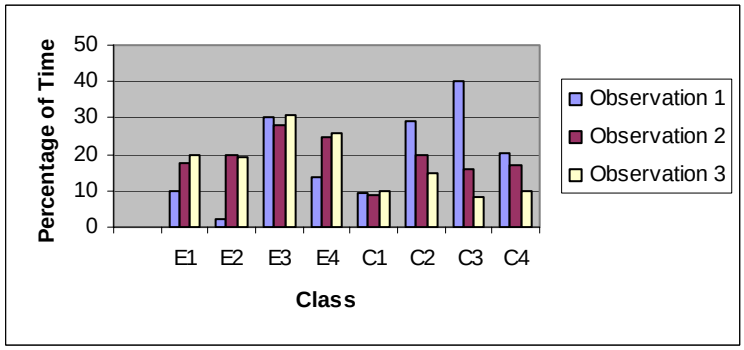

Figure 8. Percentage of time pupil responded to teacher's talk (Category 8)

From Figure 9, it was obvious that pupils in the E group initiated more talk than the C group. The increase was more than 6 percent for the E group. It seemed that feedback was effective in encouraging this behaviour.

As seen in Figure 10, it was clear that the amount of silence in the E group was reduced after feedback was given when compared with the C group. Feedback

appeared to be effective in maintaining this behaviour if the initial score was low as in E_3.

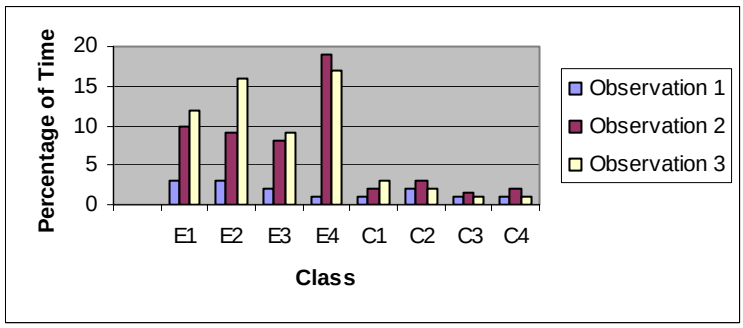

Figure 9. Percentage of time pupil initiated talk (Category 9)

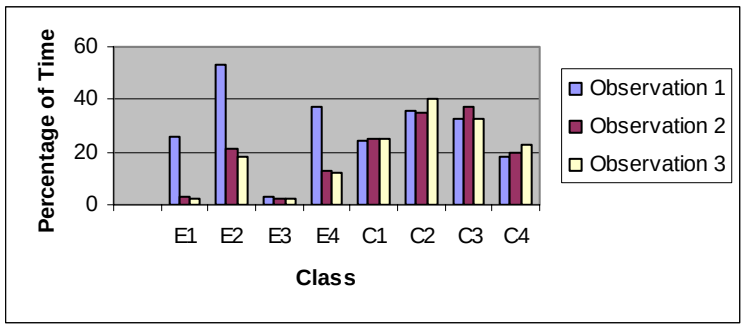

Figure 10. Percentage of time the classroom was in silence (Category 10)

T-tests for the Group Mean Differences in the Verbal Categories between the E and C Groups

Table 4 showed the results of the t-tests for comparison between the E and the C groups on each of the verbal category. It was found that categories 1, 2, 3, 4, 5, 6, 9 and 10 were significantly different at 0.05 level between the E and the C groups. However, categories 7 and 8 were found to be not significantly different at 0.05 level between the two groups.

Table 4. Results of t-tests for group means of FIACS variables for the E and C groups

Independent variables	Experimental group		Control group		
	X_1	SD	X_2	SD	t-test
Category					
1	3.232	2.1897	0.6500	0.1664	4.0730**
2	3.408	2.3067	0.6250	0.5416	4.1570**
3	3.416	2.5507	0.8330	0.4701	3.4510**
4	15.1980	5.2130	7.2500	3.2160	4.4890**
5	13.5880	6.8430	31.5150	13.6340	–4.0697**
6	6.8920	5.0010	10.9290	3.2040	–2.3377*
7	0.0350	0.0519	0.2940	0.4990	–1.7870
8	20.3400	8.8060	17.0000	9.8280	0.8770
9	17.4300	15.9500	1.1330	0.8940	3.5358**
10	16.0860	16.2100	29.7700	6.9480	–2.6880*

*significant at 0.05 level
**significant at 0.01 level

Results of ANCOVA on Pupils' Achievement in Mathematics Tests

The E and C groups were compared on three tests, i.e., Units 9, 10 and 11. The ANCOVA of the sum of scores of the three post-tests, TA-9, TA-10 and TA-11 adjusted using the sum of scores of pre-tests as covariates are shown in Table 5. Adjusted sum of square (SS) within groups for Unit 9 showed no significant difference at 0.05 level for the C and E groups. This result indicated that feedback did not bring about better achievement in mathematics for the E group at the earlier stage of the experiment. However, results of F = 28.76 and F = 15.25 for Units 10 and 11, respectively indicated that FIACS feedback was effective in enhancing the achievement of E groups at a later stage of experiment.

Table 5. ANCOVA for the sum of scores of each achievement test of the E and C groups

Test	Source of variation	df	Adjusted SS	Adjusted MS	F ratio
Unit 9	Among means	1	866.38	866.38	
	Within groups	310	85788.06	276.77	3.13
Unit 10	Among means	1	4765.92	4765.92	
	Within groups	310	51371.28	165.71	28.76**
Unit 11	Among means	1	4877.36	4877.36	
	Within groups	310	99158.81	319.87	15.25**

** significant at 0.01 level
MS: mean of scores

Results of ANCOVA on the Attitude Tests

Results of the ANCOVA on MAS (Table 6) showed that the adjusted SS and adjusted mean of the E group was significantly higher than the C group. This

finding seemed to imply that feedback was effective in promoting a healthy attitude towards the learning on mathematics.

Table 6. ANCOVA and means on the attitude towards mathematics for the E and C groups

Source of variation	df	Adjusted SS	Adjusted MS	F ratio
Among means	1	1232.723	1232.732	
Within groups	310	86023.320	277.490	4.4424*

From Table 5: df 1/310
F at 0.05 level = 3.84
F at 0.10 level = 6.63
* significant at 0.05 level

Group	Mean MAS (pre-test)	Mean MAS (post-test)	Adjusted post-test
Experimental	41.3	53.50	53.35
Control	40.9	45.70	45.85

CONCLUSION

In general, it was found that teachers given feedback in FIACS differed significantly from those teachers who were not provided with feedback on most of the categories of verbal behaviours. The experimental group (feedback group) were found to use more of accepting pupils' feeling, praised pupils more, used more pupils' ideas, asked more questions, used less lecture, initiated more pupils' talk, gave less directions and had less silence or confusion in the classroom. This indicated that there were less disciplinary problems than the control group. However, the E and C groups did not differ very much on criticising behaviour and pupil responding to teacher-talk behaviour. This finding suggested that FIACS feedback was effective in bringing about a change in teachers' verbal behaviour after feedback was given. This was further supported by the t-tests performed on the mean gain of the E and C groups before and after feedback was given. As for the academic achievement in mathematics, it was found that as the experimental time progressed, the E group performed significantly better than the C group. This had indicated that feedback in FIACS was effective in enhancing pupils' performance in mathematics. Attitude towards mathematics for the E group was also significantly higher than the C group indicating that feedback was effective in upgrading pupils' attitude towards mathematics.

The use of technology in providing instant feedback to teachers should be used in further research. A larger and wider sample covering bigger area in Brunei Darussalam should be done to find out if the findings hold. Other classroom communication instruments to measure classroom verbal and non-verbal

behaviours of pupils and teachers may be used to discover other factors for effective teaching that might be useful for practising teachers.

REFERENCES

Allington, R. (1991). Children who find learning to read difficult: School responses to diversity. In E. Hiebert (ed.). *Literacy for a diverse society.* New York: Teachers College Press, 237–252.

Amidon, E. J., and Powell, E. (1967). *Interaction analysis as a feedback system in teacher preparation.* Philadelphia: Temple University.

Bailey, L. G. (1975). An observation method in the foreign language classroom: A closer look at interaction analysis. *Foreign Language Annals, 9.*

Berliner, D., and Biddle, B. (1995). Tempus educare. In P. Peterson and H. Walberg, (eds.). *Research on teaching: Concepts, findings, and implications*. Berkeley, CA: McCutchan, 769–818.

Brophy, J., and Good, T. L. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (ed.). *Handbook for research on teaching*. New York: Macmillan, 328–375.

Campbell, J. R., and Barnes, C. W. (1972). Interaction analysis: A break through? In Sperry, Len (ed.). *Learning performance and individual differences.* Glenview, IL: Scott, Foresman and Company.

Flanders, N. A. (1970). Intent, action and feedback: A preparation for teaching. *Journal of Teacher Education, 14,* 251–260.

Good, T., and Brophy, J. (2000). *Looking in classrooms* (8th ed.). New York: Longman.

Good, T., and Weinstein, R. (1986). Schools make a difference: Evidence, criticism, and new directions. *American Psychologist, 41,* 100–1907.

Gorard, S. (2000). One of us cannot be wrong: The paradox of achievement gaps. *British Journal of Sociology of Education, 21,* 391–403.

Hopkins, W. S., and Moore, K. D. (1993). *Clinical supervision: A practical guide to student teacher supervision.* Madison, WI: Brown & Benchmark Publishers.

Houston Chronicle. (2001). Reading skills still not getting better. *100(176),* 1.

Kantowaski, M. G. (1977). Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education, 8,* 163–179.

Kline, C. E., and Dennis, H. Sorge. (1974). How effective is interaction analysis feedback on the verbal behaviour of teachers? *Educational Leadership, 32(1),* 55–62.

Sadker, D., and Sadker, M. (1994). *Failing at fairness: How America's schools cheat girls.* New York: Scribner.

Swann, J., and Graddol, D. (1988). Gender inequalities in classroom talk. *English in Education, 22,* 48–65.

Wang, M., Haertel, G., and Walberg, H. (1993). Towards a knowledge base for school learning. *Review of Educational Research, 63,* 249–294.

Younger, M., and Warrington, M. (1996). Differential achievement of girls and boys: Some observations from the perspectives of one school. *British Journal of Sociology of Education, 17,* 299–314.

Younger, M., Warrington, M., and Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education, 20,* 325–344.