Chapter Four

ANALYSIS OF RESULTS - TWO

In this chapter, an attempt is made to answer the following research questions:

1. What is the best linear combination of independent variables to predict mathematics achievement? The term 'best' is used in a statistical sense to indicate the intention of developing a subset of independent variables that is more useful in predicting mathematics achievement and of eliminating those variables that do not provide substantial additional prediction given this basic set.
2. What is the relative importance of each set of independent variables in predicting mathematics achievement?
3. Which of the variables in each of the five sets of independent variables represent that set for predicting mathematics achievement?

Considering the nature of these questions, it is decided to make use of multiple regression technique. The result of applying multiple regression to a data set is an equation that represents a best fit line between a dependent variable and several independent variables.

The stepwise regression analysis is used in order to determine the contribution of each variable taken singly and also in possible combinations. It is more appropriate especially when the intention is to derive a prediction equation with minimum number of independent variables and with optimal efficiency.

In stepwise regression, the order of entry is controlled by the sample data. That variable, which has the highest correlation with the dependent variable, is selected in the first step. At each subsequent step, the variable that adds most to the prediction equation, in terms of increasing R^2, is entered. The process will be stopped when adding the most useful remaining variable produces no statistically significant increase in the multiple correlation (Quraishi, 1979).

The problems of multicollinearity and singularity are handled well by stepwise regression. With stepwise entry, if two variables are highly correlated with each other, the first to enter takes with it both its unique variance and the variance it shares with the other; so that, the second variable rarely has enough influence remaining to enter the equation.

The difficulty in using the stepwise procedure for exploratory research, where it is less controversial compared to explanatory analysis, lies in the variability of beta weights (and thus, the contribution of variables) over samples from the same population that could produce a misleading subset of variables, and order of variables within the set, if decisions were to be based on a single sample (Tabachnick & Fidell, 1983). This problem seems to be more serious with small samples and with more number of independent variables. So the 'case-to-variable ratio' should be much high in stepwise method compared to other methods. Tabachnick & Fidell (1983) have suggested the ratio to be 40 to 1 for stepwise method while it may be 20 to 1 for other methods. The logic for this stems from the fact that stepwise technique is notorious for capitalizing on chance. In the present study, the case-to-variable ratio is 53 to 1 (1841/35) which is much more than the requirement. The dependent variable class mean mathematics achievement (MA-C) will not be regressed as this requirement is lacking due to the small sample size (56).

In the following sections, the results of the analyses are presented. The analysis of mathematics achievement - student level (MA) in the effective sample and in the restricted sample are respectively presented in the first and second sections. Concluding remarks are made in the third section.

The following are the major objectives of each step-wise regression:

1. To determine which of the independent variables (IVs) under consideration are essential for making a meaningful prediction of the dependent variable (DV).
2. To understand the order of priority.
3. To determine the efficiency of prediction by having multiple R and multiple R^2.
4. To determine how much does a particular IV add to multiple R^2 after IVs with higher priority have contributed their share to the prediction equation. It is done by analysing 'R^2 change' which is the squared semipartial correlation (S_r^2). [R^2 change = R^2 of n^{th} step - R^2 of $(n-1)^{th}$ step.]
5. To have the prediction equation. For this, the unstandardised regression coefficients (B values) and the value of intercept are to be determined. For comparative purposes, standardised regression coefficients (β values) are also helpful.
6. To locate the suppressor variables, if any. Suppressor

variable is an IV which is virtually uncorrelated with the DV, but is useful in predicting the DV and in increasing the multiple R by virtue of its correlations with other IVs. A suppressor variable "Suppresses" some variance in the other IVs that is irrelevant to the prediction of the DV. Either one of the following conditions signals the presence of a suppressor variable:

a) the absolute value of the simple correlation between the IV and DV is substantially smaller than the beta weight for that IV, or

b) the simple correlation and beta weight have opposite signs (Tabachnick & Fidell, 1983).

4.1 : Analysis of Mathematics Achievement - Student Level - Effective Sample

In this section, results of the stepwise regressions that are carried out with different groups of independent variables, are presented. Each sub- section presents one regression analysis. In the last sub-section, the results are holistically analysed.

There are five sets of independent variables in the study - student variables, instructional variables, teacher variables, class variables, and school variables. Instructional variables and teacher variables are clubbed with class variables in the analsysis because of two reasons:

1. The number of variables in these sets is limited
2. The correlations of these variables with mathematics achievement are comparatively low.

4.1.1 : Stepwise Regression :Student Entry Characteristics

Out of the eight student variables, six (except tuition and UTG) are used in this analsysis. Along with the five measures of student entry characteristics, the variable SES is also included. SES, being a background variable, is included by considering it as a control variable. Do all these six variables make unique constributions? If not, which of these variables are mostly relevant? What is the order of priority? How significantly these selected variables predict achievement? Based on the stepwise regression, it is possible to answer these questions. The analysis done with the help of computer by using SPSS package, shows the following results.

Quite naturally, because of the highest correlation with mathematics achievement (MA), Knowledge of Basic

Operations in Mathematics (KBOM) is entered in step one. R is .75 and R^2 is .56. Because of the large sample size, adjusted R^2 is not very different from R^2. F is highly significant. The variable CEC is entered in the step two. Multiple R has increased to .81 and R^2 is .65. So, the 'R^2 change' or S_r^2 which shows the unique contribution of CEC after considering the influence of KBOM, is .09. After the entry of ASC and ACA in steps three and four respectively, the regression stops. The result is shown in Table 4.1.

The following observations can be derived from the table:

1. The variables KBOM, CEC, ASC and ACA represent the predictive capacity of student entry characteristics.
2. The equation has high predictive capacity which is evident from the multiple R which is .83. These four variables could explain 69% of the variance in the mathematics achievement.
3. A specific point has to be noted with respect to the order of the variables. The correlations of KBOM and CEC with MA are almost the same. So for a different sample drawn from the same population, KBOM need not be the first variable to enter; it may be CEC. However, we can very well conclude that the variables of cognitive preparedness are the most important ones

Table 4.1: <u>Stepwise Regression of Student Entry Characteristics on Mathematics Achievement</u>

(Student Level)

No	Variables in the Equation	CORRELATIONS MA	KBOM	CEC	ASC	ACA	Step No:4 B	β	s_r^2
1	KBOM	.7464					1.4823	.3741	.5571
2	CEC	.7462	.70				.4545	.3280	.0977
3	ASC	.63	.51	.57			4.2087	.2186	.0360
4	ACA	.46	.41	.43	.41		1.5703	.0747	.0042
							intercept= −14.7487		

Variables not in the Equation	MA	KBOM	CEC	ASC	ACA
SES	.33	.31	.31	.32	.23
ACM	.47	.44	.45	.47	.58

	Steps 1	2	3	4
multiple R	.7464	.8092	.8311	.8337
R^2	.5571	.6548	.6908	.6950
Adjusted R^2	.5568	.6544	.6902	.6943

among student entry characteristics. Also, it is clear that each of the variables CEC and KBOM has a unique contribution to make in the prediction of mathematics achievement.

4. Even after the inclusion of KBOM and CEC, ASC has unique contribution to make. As observed by Bloom (1976), academic self-concept is highly relevant than other affective variables.
5. Though ACA could enter the equation, its contribution is marginal. So, the predictive capacity of school-related affect and subject-related affect is already used by cognitive variables and ASC.
6. After the inclusion of cognitive and affective variables, SES becomes insignificant. So the effects of SES seem to mediate through cognitive and affective characteristics of children.
7. There are no suppressor variables.

4.1.2 : Stepwise Regression : ACA,ACM, SES, TN and UTG

What are the influences of student variables other than the cognitive ones? It is for this purpose, this stepwise regression is carried out. The most important predictors in the previous regression - KBOM, CEC and ASC - are removed and the variables of tuition (TN) and utilisation of text-book and guide (UTG) are added. What is the nature of the prediction equation with these five student variables - ACA, ACM, SES, TN and UTG?

In this case, ACM becomes the most powerful predictor with an R of .47. The second one is tuition which brings significant change in R^2 - from .22 to .30. All the variables are entered with ACA, SES and UTG taking third, fourth and fifth positions. Total multiple R is .605 and R^2 is .367. The results are shown in Table 4.2.

The following are some of the relevant points which can be derived from the table:

1. If we consider student variables other than the cognitive ones, affective characteristics - mathematics (ACM) becomes the most important predictor. Influence of this variable will be nullified if the cognitive variables are included.
2. The influence of tuition (TN) needs special attention. In the final regression equation with five variables (results of which are presented in the Table 4.2), tuition is the variable which has most predictive power which is evident from the highest β coefficient. The β coefficient of tuition is higher than that of ACM

Table 4.2 : Stepwise Regression of ACM, TN, ACA, SES and UTG on MA

Variables in the Equation	CORRELATIONS						Step 5		S_r^2
	MA	ACM	TN	ACA	SES	UTG	B	β	
1. ACM	.47						4.415	.228	.224
2. TN	.41	.30					10.216	.238	.079
3. ACA	.46	.58	.27				4.521	.215	.036
4. SES	.33	.25	.26	.23			0.755	.148	.022
5. UTG	.18	.14	.08	.14	.12		1.520	.078	.006
Variables not in the Equation	: NIL						Intercept = - 20.102		

	Steps				
	1	2	3	4	5
multiple R	.473	.550	.582	.601	.605
R^2	.224	.303	.339	.361	.367
Adjusted R^2	.223	.302	.338	.359	.365

though ACM is the first-entered variable. An analysis of the trend of β coefficients in other steps will clarify this phenomenon. In the second step, β coefficient of ACM is more than that of TN. But in the third step, β of TN is more. The reason can be attributed to the entry of ACA. When ACA, which is highly correlated with ACM, is entered, the explanatory power of ACM has decreased largely. The amount of decrease of β for ACM when ACA is entered is much more than the decrease of β for TN. It is due to the similarity of ACM and ACA. Presence of ACA suppresses the effect of ACM.

3. The effect of SES is not completely mediated through the affective variables or tuition. The importance of cognitive variables in the process of mediation is clearly demonstrated.
4. Ommission of cognitive variables has affected the efficiency of prediction. R^2 has come down from .695 to .367. So the variables of prerequisites are essential to make the prediction efficient.
5. There are no suppressor variables.

4.1.3 : Stepwise Regression : Quality of instruction - interview, teacher variables and teacher ratings

Having done two regressions with student variables, now we turn to the analysis of other sets of variables. Because of the limited number of variables in the categories of instructional variables and teacher variables, they are clubbed with the teacher ratings of the class. The dependent variable is, of course, mathematics achievement - student level (MA).

Out of the ten variables (QII, TE, TI, EHM, FTM, CA, CM, PN, SR, QCE), six are entered in the equation. Total multiple R is .337 and R^2 is .113. These variables predict much less compared to the student variables. The results of the regression are displayed in Table 4.3.

In this regression, the presence of a suppressor variable is observed : participation as rated by teachers. While the sign of the correlation coefficient of participation (PN) with the dependent variable (MA) is positive, the sign of β is negative. It implies that PN has entered the equation not because of its direct effect on MA, but because of its effect on other independent variables

Table 4.3 : Stepwise Regression of Quality of Instruction, teacher variables and teacher ratings of the class on MA

Variables in the Equation	CORRELATIONS *							Step 6		S_r^2
	MA	QCE	EHM	PN	FTM	TI	SR	B	β	
1. QCE	.27							4.9738	.2821	.0737
2. EHM	.20	.27						3.1309	.1530	.1615
3. PN	.14	.73	.28					-4.0544	-.1630	.0141
4. FTM	.10	.23	.09	.29				2.6128	.0729	.0035
5. TI	.13	.18	.08	$\bar{0}.05$	$\bar{0}.12$			2.6393	.0575	.0029
6. SR	.20	.63	.04	.38	.06	.15		1.7227	.0686	.0027
Variables not in the Equation								Intercept = -4.5339		
TE	.01	.20	$\bar{0}.08$	.35	.16	$\bar{0}.12$	.20			
CA	.20	.59	.14	.27	.01	.30	.26			
CM	.25	.79	.18	.42	.07	.23	.54			
QII	.14	.38	$\bar{0}.08$	.20	.12	.10	.44			

	Steps					
multiple R	.2715	.3003	.3229	.3284	.3327	.3367
R^2	.0737	.0902	.1043	.1078	.1107	.1134
Adjusted R^2	.0732	.0892	.1028	.1059	.1083	.1105

* All correlations are calculated at student level i.e., N = 1841.

- in this case, QCE and EHM. PN controls the unnecessary variation of these variables.

Quality of classroom environment (QCE) - the rating given by the teacher - seems to represent other ratings especially that of class ability and class motivation. Quality of instruction - Interview (QII) has become insignificant once the other variables are entered. The effect of quality of instruction is represented by other variables. Teacher interest (TI) and the teacher-rating of 'study regularity' (SR) have unique contributions to make even after other four variables are entered.

Next regression analyses the influences of other class variables which is presented in the next subsection.

4.1.4 : Stepwise Regression : Student body Characteristics of the class

In the previous regression we have noticed the predict -ive influences of teacher ratings of the class on student learning. We have two types of variables in the set of class variables - teacher ratings and the objective measures of class. Which type of variables are more significant in prediction? For answering this question, one stepwise regression has been carried out with the objective measures of the student body characteristics of the class. Six variables are incorporated in the analysis - CEC-C, KBOM-C, ACA-C, ACM-C, SES-C and non-absenteeism. Out of these, four are entered in the equation to produce a multiple correlation of .45. The results are presented in the Table 4.4.

Among class variables, ability measures of the student body are the most important predictors. Both the measures are entered in the equation, one at the first step and the other at the fourth step. At the second step, SES of the class has entered which implies that SES contains information other than that explained by the cognitive measure of CEC. Non-absenteeism (NA) also has much predictive value. The influence of motivational variables are carried by other variables - partly by cognitive measures, partly by SES, and the rest by NA. So the cognitive measures, SES and NA represent the student body characteristics of the class. These objective measures are much more important in predicting student learning than teacher ratings.

No suppressor variable is found in the analysis.

Table 4.4 : Stepwise Regression of Student body Characteristics of the Class on MA

Variables in the Equation	CORRELATIONS *					Step 4		s_r^2
	MA	CEC-C	SES-C	NA	KBOM-C	B	β	
1. CEC-C	.42					4.351	.2224	.1780
2. SES-C	.33	.57				3.251	.1110	.0121
3. NA	.22	.31	.16			2.657	.1166	.0098
4. KBOM-C	.39	.83	.62	.17		3.207	.1200	.0039
Variables not in the Equation						Intercept = -5.170		
ACA-C	.29	.52	.58	.41	.40			
ACM-C	.33	.67	.52	.26	.62			

	Steps			
	1	2	3	4
Multiple R	.4219	.4360	.4472	.4514
R^2	.1780	.1901	.1999	.2038
Adjusted R^2	.1775	.1892	.1986	.2020

* All corrections are calculated at student level i.e. N = 1841. The correlations of class variables that are presented in the last chapter are done at class level (N=56) and hence, may be different in values.

Class variables seem to explain 20% of the variance in student learning. But this analysis does not make it clear whether class variables add anything significantly if student variables are already entered in the equation.

4.1.5 : Stepwise Regression : Student body Characteristics of the School

Review of literature makes it clear that variations in school quality constitute differential context for learning and result in differential output. But, it is not sure, whether difference in school quality arises from institutional factors or student body characteristics or both, leading to academic success. If the answer is both, which of them is more influential in terms of predictive capacity? For answering this question, two regressions are carried out - one with student body characteristics of the school and the other with variables related to school per se.

We have five meassures of student body characteristics - CEC-S, KBOM-S, ACA-S, ACM-S and SES-S. Do all these five make unique contributions to the prediction equation or are some of them redundant? How meaningfully they predict mathematics achievement? The results of the stepwise regression are presented in Table 4.5.

Table 4.5 : Stepwise Regression of Student body Characteristics of the School on MA

Variables in the Equation	CORRELATIONS *					Step 4		s_r^2
	MA	CEC-S	ACA-S	ACM-S	KBOM-S	B	β	
1. CEC-S	.39					9.7532	.4499	.1551
2. ACA-S	.34	.42				6.0948	.1979	.0365
3. ACM-S	.24	.34	.24			3.3053	.0959	.0083
4. KBOM-S	.34	.93	.39	.31		-6.0589	-0.1843	.0044
Variables not in the Equation						Intercept = -3.2178		
SES-S	.32	.73	.49	.39	.82			

	Step			
	1	2	3	4
Multiple R	.3939	.4377	.4471	.4520
R^2	.1551	.1916	.1999	.2043
Adjusted R^2	.1547	.1907	.1986	.2026

* All correlations are calculated at student level i.e., N = 1841

Student body characteristics of the school can explain 20 per cent of the variance of mathematics achievement. The influence is almost the same as that of student body charact -eristics of the class. Here also mean CEC entered first. School mean KBOM also has entered, but the nature of entry is totally different. Here, the role of KBOM-S is that of a suppressor variable. It enhances the predictive capacity of other independent variables. SES is not entered but the affective variables are entered. It is illuminating to find that SES of the student body does not enter the equation if cognitive and affective variables are already entered. It implies that the influence of SES is mediated through cognitive and affective variables. This interpretation is possible because SES is theoretically prior to cognitive and affective entry measures.

Now we turn to the analysis of other school variables.

4.1.6 : Stepwise Regression: School Variables other than Student body Characteristics

Four variables are included in this analysis - School locality (SL), School type (ST), Past achieveent of the school (PAS) and Psycho-social environment (PSE). Out of these, only the last two are entered in the equation. SL and ST do not add anyhting significantly to the prediction made by PAS and PSE. Multiple R is .44 and R^2 is .19. So the efficiency of prediction is one percent less than that made by student body characteristics of the school. No suppressor variable is found in this analysis. The results of the analysis are given in Table 4.6.

4.1.7: Stepwise Regression: School locality and School type

In the previous regression, the effects of school locality (SL) and school type (ST) are carried by PAS and PSE. But, how much do locality and type predict? Another regression is carried out with these two variables alone. The results are displayed in Table 4.7. School locality has entered first making an R^2 of .034. When ST has entered the second step, R^2 has increased to .089. So the 'R^2 change' made by ST (S_r^2) is .055 which is, interestingly, higher than that of SL. Here, ST acts as a suppressor variable and controls the unnecessary variation made by SL. The β of SL has not decreased but increased in the second step (from .1855 to .2519). The presence of ST has made SL much more strong in prediction. It is the interaction effect which made the difference.

Table 4.6 : Step wise Regression of SL, ST, PAS and PSE on MA

Variables in the Equation	CORRELATIONS *			Step 2		S_r^2
	MA	PAS	PSE	B	β	
1. PAS	.42			8.4010	.2882	.1771
2. PSE	.39	.75		3.3135	.1762	.0135
Variables not in the Equation				Intercept = 2.7996		
SL	.19	.53	.05			
ST	.17	.24	.43			

	Step 1	Step 2
Multiple R	.4209	.4366
R^2	.1771	.1906
Adjusted R^2	.1767	.1897

* All correlations are calculated at student level i.e., N = 1841

Table 4.7 : Stepwise Regression of School Locality (SL) and School Type (ST) on MA

Variables in the Equation	CORRELATIONS			Step 2		s_r^2
	MA	SL	ST	B	β	
SL	.19			11.3001	.2519	.0344
ST	.17	-0.27		10.1143	.2429	.0546
Variables not in the Equation : NIL				Intercept = -4.2685		

	Steps	
	1	2
Multiple R	.1855	.2984
R^2	.0344	.0890
Adjusted R^2	.0339	.0881

In the regression 4.1.6 it can be seen that R^2 is .1906 for the variables of PAS and PSE. In that equation, SL and ST are not entered which implies that the predictive influences of SL and ST are carried by the other two. In the present regression (4.1.7), we note that R^2 is .0890 for SL and ST. So, nine percent of the variation in the dependent variable, which is explained by PAS and PSE together, is also explained by the variables SL and ST together. As such, the unique contribution of PAS and PSE together seems to be 10 per cent of the variation in the dependent variable.

Thus far, we were analysing mathematics achievement (Student level) with different groups of independent variables. In the next section, the dependent variable will be regressed with all the independent variables.

4.1.8 : Stepwise Regression: All Independent Variables

In the study, the major dependent variable is mathematics achievement - student level. There are 35 independent variables categorised into five sets. Seven regression analyses are described with different groups of independent variables. If all the independent variables are considered together in a stepwise regression, what will be the multiple R^2? How many variables are required for optimizing R^2? Two kinds of criteria may be used for obtaining an optimal R^2 with minimum number of variables: statistical significance and meaningfulness (Kerlinger & Pedhazur, 1973. page 286). After having the criterion of statistical significance by the method of stepwise regression, the results will be analysed by using the criterion of meaningfulness. The results of the regression are displayed in the Table 4.8.

Though there are a total of 35 independent variables, only 28 are included in the analysis. Quality of instruction -observation, and attention and participation of students are excluded because they apply only to the restricted sample. Teacher ratings of the class except that of quality of classroom environment (QCE) are not included because of two reasons : (a) objective measures of studentbody characteristics of the class are more meaningful in prediction which is clear from the correlations, and the regressions with class variables, and (b) QCE seems to represent most of them.

Out of the 28 variables, 14 are entered in the prediction equation. Total multiple R is .855 and R^2 is .732. F is highly significant. As far as prediction is

Table 4.8 : Stepwise Regression of all Independent Variables on MA

Variables in the Equation	CORRELATIONS *														Step 14		
	MA	1	2	3	4	5	6	7	8	9	10	11	12	13	B	β	S_r^2
1.KBOM	.7464														1.354	.342	.557
2.CEC	.7462	.70													.424	.306	.098
3.ASC	.63	.51	.57												4.038	.210	.036
4.TN	.41	.35	.29	.25											5.678	.133	.015
5.PSE	.39	.36	.39	.20	.05										1.846	.098	.009
6.KBOM-S	.34	.41	.41	.23	.18	.58									$\bar{3}.915$	$\bar{0}.119$	.005
7.ACA-S	.34	.30	.31	.17	.10	.46	.39								2.169	.070	.004
8.FTM	.10	.08	.06	.02	.18	.20	.34	$\bar{0}.10$							1.437	.040	.002
9.ACA	.46	.41	.43	.41	.27	.16	.15	.26	$\bar{0}.06$						1.072	.051	.001
10.PAS	.42	.39	.41	.27	.14	.75	.75	.64	.19	.20					2.758	.095	.001
11.KBOM-C	.39	.50	.46	.24	.19	.56	.75	.41	.17	.17	.66				$\bar{1}.519$	$\bar{0}.057$	.001
12.UTG	.18	.14	.16	.16	.08	.05	.06	.10	$\bar{0}.003$	.14	.07	.11			.596	.031	.001
13.ACA-C	.29	.26	.30	.20	.10	.44	.31	.65	$\bar{0}.13$	.32	.52	.40	.13		$\bar{1}.136$	$\bar{0}.045$	.001
14.QII	.14	.11	.11	.08	.10	.22	.16	.10	.12	.03	.13	.22	.05	.14	.646	.026	.001

Intercept =-25.390

* All correlations are calculated at student level i.e., N = 1841

Table 4.8 (Continued)

Variables not in the Equation: SES, ACM, TE, TI, QCE, NA, CEC-C, ACM-C, SES-C, SL, ST, CEC-S, ACM-S, SES-S.

	Steps													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Multiple R	.746	.809	.831	.840	.846	.848	.851	.852	.853	.853	.854	.854	.855	.855
R^2	.557	.655	.691	.706	.715	.720	.724	.726	.727	.728	.729	.730	.731	.732
Adjusted R^2	.557	.654	.690	.705	.714	.719	.723	.724	.726	.727	.728	.729	.729	.730
Standard Error	13.81	12.19	11.54	11.26	11.09	10.99	10.92	10.89	10.86	10.84	10.82	10.81	10.79	10.78
F *	2312.94	1743.05	1367.70	1102.10	921.01	785.72	685.74	605.42	542.08	490.56	448.33	412.50	382.29	355.91

* All F's are highly significant at .0000 level.

concerned, R^2 seems to be quite high. As the sample size is high,adjusted R^2 is very close to R^2 : .73. Out of the 14 variables entered, three - KBOM-S, KBOM-C and ACA-C act as suppressor variables which is evident from the negative signs of Beta weights though their correlations with the dependent variable are positive. These variables help in the process of prediction by controlling the unnecessary variations of other independent variables.

As far as increments in R^2 are concerned, there are notable increments in the first few steps. After step 6, increments become very marginal. At the step 6, R^2 is .720. At the step 14, it has increased only upto .732. It implies that only one percent of additional variance can be accounted for by the last eight variables. So,as far as the criterion of meaningfulness is considered, the analysis could have stopped at the step six. These six variables seem to represent the prediction capacity of all independent variables.

Out of these six major variables, first four are student variables. Psycho-social environment (PSE) of the school is the fifth one. The sixth one is also a variable of the school,but it acts as a suppressor variable. None of the variables from the sets of instructional variables, teacher variables or class variables has occupied a position in the first six variables. These variables do not have much unique contributions to make when the effect of other variables are partialled out.

Student variables occupy important positions in the equation. Out of the eight student variables, six are entered in the equation. The left out variables are affective characteristics - mathematics (ACM) and SES. This result is not unexpected because from the first stepwise regression of which the results are presented in Table 4.1, it is clear that the contributions of these variables are carried by KBOM, CEC, ASC and ACA. All these latter variables are entered in the present equation, the first three occupying the first three places.

The first two variables - KBOM and CEC - are the measures of cognitive preparedness of students. KBOM assesses the knowledge of basic operations in mathematics and CEC measures the level of prerequisites for learning tenth class mathematics. It is really unfortunate to find that knowledge of basic operations becomes the best predictor of tenth class mathematics achievement. In a real educational situation, all tenth class students are expected to know at least the basic concepts and operations in all subjects. Here, students differ considerably in the knowledge of basic operations. The fact of KBOM becomes the most important predictor of tenth class achievement reflects the quality of schooling.

KBOM is the first variable to enter the equation. This fact cannot be interpreted as the prediction capacity of KBOM is more than that of CEC because the correlations of KBOM and CEC with the dependent variable are almost the same. So in a different sample from the same population, there are good chances that the order may change - the correlation of CEC with the dependent variable may become slightly more than that of KBOM. What is important is not the relative importance of KBOM and CEC, but the combined effect. These two variables together can explain 65.54% of the variance in the dependent variable. This is the effect of cognitive entry measures on achievement. So the cognitive entry characteristics are the best predictors of learning intellectual skill and they can account for 65% of the variance.

When the variable academic self-concept (ASC) is also entered together with the cognitive variables, the multiple R^2 has increased to 0.69. The unique contribution made by ASC, which is assessed by the squared semi-partial correlation, is 3.5 per cent. These three variables -CEC, KBOM and ASC - are conceptually causal to achievement. They are theoretically causal; they are assessed prior to achievement; so, the results of the regression analysis can be meaningfully interpreted as causal influences.

It is the dichotomous variable tuition which has entered in the step four. In the regression 4.1.2., it was shown that tuition has unique contribution to make after partialling out the effects of affective variables. In the present regression, we notice that tuition has unique contribution to make after partialling out the effects of cognitive variables. These results imply that even after controlling the effects of cognitive and affective entry characteristics, tuition exerts a unique influence on mathematics achievement. What is being considered is not the quality of tuition, but whether a particular student goes for tuition or not. Results indicate that tuition has an independent positive effect on learning. This effect does not vanish even after the variables of cognitive and affective entry characteristics are properly controlled.

These four student variables-KBOM, CEC, ASC and TN - explains 70.6 per cent of the variance in mathematics achievement. An additional 1.4 per cent is explained by two school variables - PSE and KBOM-S. While the effect of PSE

(Psycho-social environment of the school) is direct, that of KBOM-S (School mean KBOM) seems to be indirect. The latter variable acts as a suppressor variable. In the regression 4.1.6, we have noted that school variables of PAS and PSE together can explain 19 per cent of the variance in the dependent variable. In that regression, past achievement of the school (PAS) was the first variable to enter. But when student variables have occupied prior positions in the present regression, PSE has much more unique contribution to make than PAS. It shows the uniqueness of the variable PSE. Even after the student variables of cognitive and affective entry characteristics and social class are controlled, the quality of the environment of the school does make some difference in achievement. This effect is present even after the major aspects of influence of 'Past achievement of the school' are also controlled (of course indirectly, which is clear from the above discussion). It has to be noted that, once four student variables are entered, the variable that got preference to enter at step five is PSE from a host of many possible variables - instructional variables, teacher variables, class variables and school variables. This implies that more explanatory power is carried by the variable PSE. So the psycho-social environment of the school is the most powerful school-related variable to explain the variations in achievement, once the effects of student variables are accounted for.

The variable KBOM-S is relevant only for the purpose of prediction. Because of its ability to control the unwanted variations of the other five previously entered independent variables, it could enhance the efficiency of prediction.

These six variables explain 72% of the variance in the dependent variable. Even if the analysis is stopped at this step, the efficiency of prediction will not be affected much. Other eight variables that are entered in the equation do add an R^2 of .012. Out of the eight variables, two are student variables - ACA and UTG; One is the instructional variable QII ; One is the teacher rating - the facilities for teaching mathematics (FTM); two are class variables - KBOM-C and ACA-C; and two are school variables - ACA-S and PAS.

Quality of instruction-interview has entered the equation as the fourteenth variable. The predictive capacity of this variable is very weak. Still, even after the entry of thirteen variables, QII has a significant unique contribution to make - whatever small the effect may be.

A note on class variables is needed. No class variable has entered the equation except the two suppressor variables. It may be considered to imply that class variables do not have anything to add once the student variables and school variables are entered. This may be due to the fact that in a given school, classes do not show much variation. The selection of students to a particular class is almost random in all the selected schools. Segregation is not practised in the schools of the sample. As such, class variables lose their relevance when the school variables are entered in the equation.

To summarise, student variables are the most important predictors of mathematics achievement. Measures of prerequisites are highly influential. KBOM, CEC, ASC and tuition are the variables in the category of student variables which carry unique contributions for the prediction and they represent the category in terms of predictive capacity. The effects of motivational variables are carried by KBOM, CEC and ASC. The effect of the instructional variable QII is marginal, though significant and unique from a statistical point of view. Psychosocial environment of the school is relevant even after the student variables are entered. School variables seem to be more inclusive and explanatory than class variables - the latter lose the relevance once the former are entered. Among the class variables, objective indices are more meaningful in prediction compared to teacher ratings. The suggested meaningful combination of six independent variables is able to explain 72 percent of the variance of the mathematics achievement.

In the next section, the analysis of the restricted sample is presented. The major focus of the analysis is on the role played by quality of instruction-observation and attention and participation of students. These variables can only be analysed in the restricted sample.

4.2 : Analysis of Mathematics Achievement - Student Level - Restricted Sample

The prime objective of conducting an analysis with the Restricted Sample is to understand the nature of influences of quality of instruction-observation (QIO) and attention and participation of students (APS) on mathematics achievement. From the correlational analysis, the obtained coefficients of correlation of QIO and APS with mathematics achievement (MA) are 0.0446 and 0.17 respectively. The former is significant at $p = .068$ and the latter is highly significant at $p = .000$. The sample size is 1116. Theoretically, QIO is highly related to MA, while statistically, the correlation came out to be quite less. Why? Several explanations can be given.

A coefficient of correlation is an index of the linear relationship between two variables. So, if the relationship is non-linear, the correlation cannot be trusted. Is the relationship between QIO and MA non-linear?

Let us consider some hypothetical examples to check the linearity between QIO and MA. Suppose there are three classes A, B and C where the quality of instruction is high, medium and low respectively. Suppose there are three students: 'x' who is a gifted student; 'y' who has a satisfactory level of prerequisites; and 'z' who does not have the essential prerequisites. Even if the quality of instruction is low, 'x' may be able to learn. So, quality of instruction, if it has to make some influence on 'x', has to be challenging enough and suited to the needs of gifted students. If this aspect is lacking, quality of instruction will not influence the achieveent of 'x'. The quality of instruction in class A may be comparatively high, but it may not be sufficiently high to challenge gifted students. In such a case, the achievement of gifted students in classes A, B and C will not differ.

It is for the student 'y',quality of instruction seems to be much influential. If 'y' is studying in class A, the learning and achievement will be much more than if he/she is attending class B. Similarly, the achievement of 'y' will be more in class B compared to that in the case of attending class C. So if all the students have the required prerequisites, instructional quality influences learning differentially and in such a case, the relationship will be linear. Unfortunately, it is not the case.

'z' is the representative of the majority of students in our sample. In a test of prerequisites, 'z' scores very low. 'z' will not be able to understand instruction even if he/she is studying in class A. For influencing achievement, instruction should have some special measures for teaching them prerequisites. Once 'z' learns the prerequisites, he/she becomes 'y' and then, instructional quality will be linearly related. So, if prerequisites are not taught, the achievements of 'z' will not differ much whether he/she is attending class A, class B or class C.

From the hypothetical discussion, it is evident that the quality of instruction will be linearly related to achievement only when all the students possess essential prerequisites and the high-quality instruction offers challenges for gifted students. This is not the present case. There is another possibility. The variable 'teaching of prerequisites' can be incorporated to the variable 'quality of instruction' i.e., high-quality instruction should teach prerequisites very well, medium-quality instruction should teach them less well and so on. In this condition also, the relationship of quality of instruction with achievement will become linear. But in the present sample, virtually no teacher, irrespective of his/her quality of instruction, teaches prerequisites. Whenever they give explanations for prerequsites, their intention is only in brushing up the memory of students who know the prerequisites. Even these explanations do not touch the students who are really poor with prerequisites. Lack of sufficient mastery of prerequisites is not an exception, but has become the characteristic of the majority of students.

Above discussion implies that the relationship of quality of instruction with achievement in the present sample does not seem to be linear. It seems to have some interactions with student entry characteristics. Even these interaction effects do not seem to be linear.

As the correlations of QIO and APS with achievement are relatively small, stepwise regression will not be helpful.

Further, multiple regression does not provide the effects of each value of the independent variables on the dependent variable. So it is decided to make use of multiple classification analysis (MCA). MCA is a modified form of multiple regression. If one looks more closely to MCA, it is an ordinary regression analysis using dummy variables (Sathe & Murthy, 1987). MCA techniques were originally developed by Yates in 1934 and elaborated by Anderson and Baneroft in 1952.

For the analysis of the restricted sample, MCA is preferred to ordinary regression for many reasons. In general, multiple regression technique requires that the variables involved in the model are on an interval or ratio scale and are measurable. Also, the relationships among the variables should be linear and additive. Nominal variables can be introduced to regression model as 'dummy' variables. But it is generally believed that the multiple regression analaysis with 'dummies' weakens the power of the contribution made by the predictor variables. In MCA, though the dependent variable should be on an interval scale (or at least dichotomous) predictor variables can be on interval, ordinal or nominal scales. Weak measurement of predictor variables is not the only advantage of MCA over the traditional regression approach, but problems like multicollinearity and non-linear relationships are adequately handled by MCA (Sathe & Murthy, 1987). Further, MCA can provide the effects of each category of the categorical independent variables on the dependent variable. So, the nature of the influence made by QIO and APS can be better understood with MCA.

MCA assumes additive effects of all the independent variables on the dependent variable. That is, the model assumes that there are no interactions among the independent variables. As such, MCA is a special case of analysis of variance with no interaction terms (Sathe & Murthy, 1987). So, to analyse interaction effects, if any, analysis of variance has been carried out. The details of the independent variables that are used in the analysis of variance and MCA are explained below.

Along with quality of instruction - observation (QIO), and attention and participation of students (APS), four other independent variables are included in the analysis - one measure of cognitive preparedness of students, one measure of affective readiness, tuition, and finally SES as a controlling variable. Instead of KBOM, CFC is selected as the measure of cognitive preparedness because of the reason that it is the strongest predictor of achievement in the restricted sample. ACM (affective characteristics - mathematics) is selected as the variable of affective readiness. So in the present analysis, there are five independent variables (CEC, ACM, TN, QIO and APS) and one co-variate (SES). The details of the catgorisation of these variables are presented in Table 4.9.

Table 4.9 : System of Categorisation of independent variables for ANOVA and MCA

Variable	Raw Scores	Categorised Scores
Cognitive Entry Characteristics (CEC)	12 or less 13 to 32 33 or more	1 2 3
Affective Characteristics - Mathematics (ACM)	5 or less 6 to 11 12 to 16 17 or more	1 2 3 4
Tuition (TN)	No tuition goes for tuition	1 2
Quality of Instruction-Observation (QIO)	3 or 4 5 or 6 7 or 8	1 2 3
Attention and participation of students (APS)	1 or 2 3 4 or 5	1 2 3

There is no change for the variable tuition. ACM is already categorised for prior analyses and the same scheme is followed. For QIO and APS, the ratings are further categorised so that three categories are formed for each variable. CEC is categorised according to the mean and standard deviation.

First of all, analysis of variance is carried out with these variables. There are five independent variables. There is one covariate - SES. The dependent variable is mathematics achievement - student level (MA). N is 1116. The results are presented in table 4.10. The analysis is done with the help of computer.

From the table, it is evident that all the main effects of the independent variables are highly significant. As the higher order interactions have been suppressed, we cannot derive anything about the possible interactions. R^2 is found to be .57 leaving the residual variance of .43. This result means that 43 per cent of the variance in the dependent variable is unexplained by these six independent variables (including the covariate SES). Both the variables QIO and APS are influential though they are the least influential among the six. The influence of APS is more than that of QIO.

Table 4.10 : ANOVA for MA in the Restricted Sample *

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
COVARIATE					
SES	65503.433	1	65503.433	314.1415	0.000
MAIN EFFECTS	240461.626	10	24046.163	115.3206	0.000
CEC	90919.184	2	45459.592	218.0152	0.000
TN	22745.386	1	22745.386	109.0824	0.000
ACM	15513.080	3	5171.027	24.79922	0.000
QIO	3043.138	2	1521.569	7.297143	0.001
APS	4211.670	2	2105.835	10.09917	0.000
Explained	305965.059	11	27815.005	133.3952	0.000
Residual	230201.356	1104	208.516		
Total	536166.416	1115	480.867		

* Due to Empty cells or a singular matrix, higher order interactions have been suppressed.

Multiple classification analysis (MCA) is carried out with the same five independent variables and the covariate SES to examine the contribution of each category of the predictor variables before and after adjustment of the controlling variables. MCA provides the grand mean of the dependent variable and a table of category means for each predictor variable expressed as deviations from the grand mean. Thus, they reflect the magnitude of the effect of each category of a predictor. These category effects are obtained in two different forms, viz. (i) unadjusted, and (ii) adjusted for variations in the predictors and for differences in the covariates as and when they are appropriate. Associated with the category effects of each predictor, a value of 'η' (eta) for unadjusted effects or 'β' (beta) for adjusted effects are also calculated. The square of eta and the square of beta respectively provide the variance explained by the particular variable before and after controlling the effects of other variables. Betas, adjusted for independent variables and covariates, can be seen as standardised partial regression coefficients.

The results of MCA are presented in Table 4.11.

Table 4.11 : Multiple Classification Analysis of MA by CEC, TN, ACM, QIO and APS with SES:

Restricted Sample

Grand Mean = 27.70

Variable	Category	N	unadjusted deviation	Eta	Adjusted deviation for independents & Covariates	Beta
	1	269	-16.97		-11.69	
	2	488	- 5.45		- 4.34	
CEC	3	359	20.13	0.67	14.66	0.48
	1	642	- 8.17		- 4.29	
TN	2	474	11.07	0.43	5.82	0.23
	1	207	-12.46		- 3.48	
	2	310	- 9.22		- 3.94	
	3	256	0.83		- 0.63	
ACM	4	343	15.24	0.51	6.13	0.19
	1	408	- 0.52		2.03	
	2	469	1.20		- 0.03	
QIO	3	239	- 1.47	0.05	- 3.40	0.09
	1	366	- 6.61		- 3.40	
	2	450	3.91		1.65	
APS	3	300	2.20	0.21	1.68	0.11
Multiple R						0.755
Multiple R^2						0.571

From the table, it is evident that all the variables accounted for 57 per cent of the overall variance in mathematics achievement. R^2 is almost the same as that obtained by ANOVA, the reason being that the interaction effects were not included by ANOVA. Usually, R^2 in ANOVA will be more than that of MCA - the major difference is that the former method includes interaction effects (not necessarily statistically significant), while the latter ignores them.

Among the five independent variables considered in the multiple classification analysis, the variable with the maximum explanatory power is CEC (cognitive entry characteritics), which explain nearly 45 per cent of the variability when the other factors and covariates are ignored (η^2) and about 23 per cent after the effects of other variables are controlled (β^2). The effect of this variable on mathematics achievement is systematic and positive : Unadjusted means (Grand mean + deviation) of the categories 1, 2 and 3 are 10.73, 22.25 and 47.83 respectively. Differences in the means are very sharp. After the adjustments, the effects are comparatively decreased.

The cases of TN and ACM are similar - the effects are systematic and sharp. While the unadjusted mean of tuition-going group is 38.77, it is only 19.53 for the 'no-tuition' group. Even after adjustments, a difference of ten points in the means is observable. In the case of ACM, the difference between means of highest and lowest categories is approximately 28 points before adjustment and is only 9.5 after adjustment. After adjustment, mean of the second category is lower than that of the first category. But they are almost equal. So if the first two categories are combined together, the influence of ACM will increase. The effect of ACM is sharply lowered after adjustent. This implies that much of the influence of ACM is shared by other variables. In stepwise regressions, we have seen that the explanatory power of ACM is fully carried by cognitive variables. Present result reinforces that finding.

While the variables CEC, TN and ACM show systematic results, quality of instruction (QIO) presents a different picture. Highest achievement is reported for category 2, followed by category 1 and the achievement of category 3 is the least. The differences are only marginal: a difference of 2.67 points between the highest (category 2) and the lowest (category 3). The implication is that the relationship between QIO and MA does not seem to be linear.

For categories 1 and 2, slight linear relationship is observed. But this relationship is almost nullified by category 3.After adjustments for other variables, the trend has become totally different. Subsequently, highest mean is obtained by category 1, followed by category 2 and then only category 3. Now, the relationship has become linear, but in opposite direction. It implies that the observed differences in the unadjusted condition are not stable. The categories seem to be almost arbitrary. The only conclusion that can be derived is that QIO per se is not related to achievement. If they have a relationship, it is not linear. Further, if there is a relationship, it is suppressed by other variables. A possible hypothesis that can be logically derived is that, quality of instruction interacts with entry characteristics to determine achievement, a situation which cannot be handled by MCA. As the multiple classification analysis is based on the additivity assumption, interaction effects, if they are present, weaken the efficiency of MCA.

Many other kinds of reasons can be offered for the observed low relationship between QIO and MA. The differences among teachers are quite less. Further, many students go for tuition and the quality of tuition may interfere with the influence of QIO.

Mathematics achievement is not assessed by a test developed by the investigator. This simple fact may affect the findings. Most of the questions in the public examination, test memory rather than learning. Instruction of low quality may be enough for answering such questions. Memorisation of a definition can be done without understanding the concept. If a test is used in which the questions are of understanding level, QIO may become more influential. As the assessment of instruction has been done with certain assumptions about the testing of concepts and rules, the quality of questions matters much.

With respect to attention and participation of students (APS), the second and third categories do not seem to differ much. But they differ from the first category. The influence has become more systematic after the adjustments for covariate and other independent variables are made. The relationship seems to be linear especially when the categories of 2 and 3 are clubbed together. Attention and participation of students is an important class variable - theoretically and statistically.

The important finding of the analysis is the low relationship between quality of instruction and achievement.

The relationship is neither direct, nor linear. Many kinds of reasons are offered for this trend. In the present study, entry characteristics of students seem to determine achievement. Instruction does not free achieveent from the impact of entry characteristics. The conclusion is that instruction is not differentially influencing achievement.

4.3 : Concluding Remarks

The following are the conclusions derived from the analysis:

1. Six important independent variables are identified, which together can account for 72 per cent of the variance in mathematics achievement. Four are student variables and the other two, school variables. The selected student variables are : KBOM, CEC, ASC and TN. The measures of cognitive preparedness of students - KBOM and CEC - are the most influential predictors. The effects of school-related affect and subject-related affect are carried by other variables. Academic self-concept (ASC) is an important dimension, both theoretically and statistically. The dichotomous variable 'tuition' has significant independent effect even after the effects of cognitive entry measures,

affective entry measures and socio-economic status are partialled out. Once the effects of student variables are accounted for, the variables of school become prominent. The psycho-social environment of the school -PSE - is the most powerful variable among school variables. This result reinforces the general finding by other researchers (for instance, Coleman et al, 1982) that the variables of the functioning of the school are more influential in determining student learning than the structural charcteristics of the school or the student body characteristics. The second selected variable from the category of school variables is the school mean KBOM (KBOM-S). It is meaningful only for the purpose of prediction. It acts as a suppressor variable.

2. The present study clearly demonstrates that student variables are the most important predictors of achievement. The second important set is school variables, followed by class variables. Teacher variables and instructional variables are least prominent as far as prediction is concerned.

3. Among student variables, measures of cognitive prerequisites have maximum effect on the dependent variable. These two measures - KBOM and CEC - together with the 'academic self-concept' and 'tuition'

represent the predictive influence of student variables.

4. Quality of instruction shows a doubtful picture. This variable does not seem to have an independent direct effect on achievement. Maybe the relationship is non-linear. Maybe this variable interacts with entry characteristics to determine the level of student learning. It can be concluded that instruction is not at all successful in freeing achievement from the impact of student entry characteristics.

In the present chapter, it is attempted to provide the results and interpretations of stepwise regressions, analysis of variance and multiple classification analysis. Findings of the present study are holistically presented in the next chapter to draw valid inferences and to discuss some important implications therefrom.