Chapter 5

ESTABLISHING CAUSAL RELATIONSHIPS

5.0.0 Introduction

The first research question that the present study attempted to answer was the following:

What are the specific causal relationships that exist between home environment, personality characteristics and school performance of urban, lower middle class Indian children of the first standard?

The design for building an empirical basis for answering this question was described in section 3.4.1 in chapter 3. A detailed model of causal relationships among selected vatriables was hypothesised on the basis of logical reasoning and previous research findings. The variables hypothesised to be involved in the web of causal relationships depicting the phenomenon under study were Mother's Education (ME), Gender of the child (G), Home Interaction Pattern (HIP), Educational Environment at Home (EEH), Social Competence (SC), Cognitive Development (CD), Adjustment to School (AS) and Academic Achievement (AA). The hypothesised model of causal relationships among these variables was presented in figure 3.7. It was further proposed that these causal relationships would be established by identifying or constructing suitable measuring instruments for each variable, administering these instruments on a sample of children from the identified population and thus deriving sample distributions of quantitative measures of the selected variables; and finally processing these distributions and testing the hypothesised causal model through Path Analysis.

In chapter 4, the procedural details pertaining to the identification or construction of measuring instruments for the selected variables were discussed. The details of data collection were also presented therein. The present chapter attempts to present the processing of the quantitative data collected and the verification of the hypothesised causal model through Path Analysis.

5.1.0 Processing of the data

As was discussed in chapter 4, the instruments identified or constructed for measuring the selected variables were such they that yielded various ranges of scores on various types of scales. A summary of the variables, instruments, scale and range is presented in table 5.1. As can be seen in the table, the quantitative measures were described on nominal, ordinal or interval scales for different variables. In each case, the categories were represented in numerical terms for the sake of processing the data quantitatively.[1] The quantitative data collected through procedures in chapter 4 were in terms of eight sample distributions. The sample distributions of raw scores for the eight variables are presented in appendix 8.

The raw data were processed in order to derive certain fundamental descriptions of the sample distributions. Descriptive statistics, viz., mean and standard deviation were computed for the sample distribution in raw scores of each variable. This information is presented in table 5.2.

---

1. A natural dichotomy like gender of the child ought not to have been treated like a continuum the way the other variables were treated. It was only for the sake of 'neatness' of the comprehensive analytic frame adopted here, that such an exercise was attempted. Literature on path analysis reports instances where similar attempts have been made (James, et al., 1982).

TABLE 5.1: Variables, instruments, scale and range

S.No.	Variable	Instrument	Scale	Range From	Range To
1.	Mother's Education	Years of formal Education	Ordinal	0	19
2.	Gender of the child		Nominal	1* Female	2* Male
3.	Home Interaction Pattern	Home Interaction Pattern Scale	Ordinal	-20	20
4.	Educational Environment Home	Educational Environment at Home Scale	Ordinal	25	75
5.	Social Competence	Social Competence Scale	Ordinal	32	96
6.	Cognitive Development	Battery of Piagetian Tasks	Ordinal	1	3
7.	Adjustment to School	Adjustment to School Inventory	Ordinal	0	20
8.	Academic Achievement	Teacher-made Achievement Tests	Interval	35	100

* This is not strictly a continuum or a range. Numerical equivalents have been fixed for the gender categories merely for the sake of treating the data quantitatively.

TABLE 5.2: Means and standard deviations of raw score distributions of variables

S.No.	Variable	Mean	SD
1.	Mother's Education	10.34	2.87
2.	Gender of the child	1.73*	0.45*
3.	Home Interaction Pattern	5.73	4.73
4.	Educational Environment at Home	49.32	12.20
5.	Social Competence	72.57	11.04
6.	Cognitive Development	1.98	0.65
7.	Adjustment to School	13.50	4.51
8.	Academic Achievement	61.06	12.99

* In the case of gender, the mean and SD do not mean anything since values are arbitrarily assigned and they are not on a continuum. This exercise is done however just in order to convert the raw scores to z scores and process the data further.

5.3.0 Formulation of the hypothesised path model

The task now is to verify the hypothesised model of causal relationships among the selected variables presented in figure 3.7 through Path Analysis. For this purpose, the hypothsised causal model will have to be converted into hypothsised path model, so as to distinguish the various paths of causal relationships that have to be put to test. Figure 5.1 presents the hypothesised path model derived from the causal model presented in figure 3.7. As can be seen in figure 5.1 nineteen paths of causal relationships (A to S) have been identified among the eight variables (1 to 8). The nineteen paths are clearly represented on a two-dimensional matrix of eight variables in table 5.3.

5.4.0 Path analysis

Having formulated the hypothetical path model and delienated the paths of causal relationships among variables to be empirically verified, the Path Analysis per se begins. The logic behind Path Analysis is to juxtapose a set of empirical information over an *apriori* model of causal relationships, so that those causal links in the model which do not have empirical substantiation get eliminated and those having empirical substantiation get established. As a consequence, the hypothetical path model gets trimmed while being verified empirically.

What kind of empirical information is required to confirm a path model? As James, Mulaik and Brett (1982) put it, "a structural model is confirmed if the predictions regarding correlations (variances / covariances) among ... variables are consistent with the observed (i.e. empirically derived) correlations (variances / covariances) among ... variables" (p. 59). They further clarify that "predictions regarding estimated structural parameters can be tested

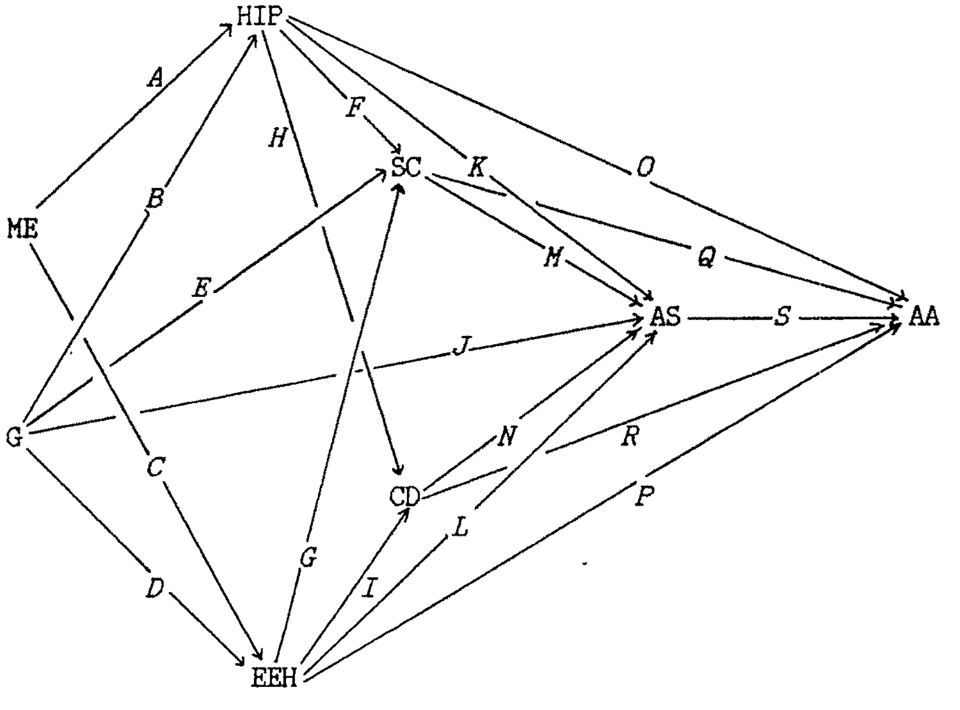

FIGURE 5.1: Hypothesised path model

TABLE 5.3: Paths in the hypothesised path model

			From							
	S.No.	Variable	1 ME	2 G	3 HIP	4 EEH	5 SC	6 CD	7 AS	8 AA
To	1	ME								
	2	G								
	3	HIP	A	B						
	4	EEH	C	D						
	5	SC		E	F	G				
	6	CD			H	I				
	7	AS		J	K	L	M	N		
	8	AA			O	P	Q	R	S	

in random samples from well defined populations by estimating statistically the value of structural parameters and conducting tests of significance on the estimates. If the estimate of a structural parameter is significantly different from zero, then the prediction is regarded as confirmed, which implies that the functional equation is consistent with the data" (pp.60-61). Stuessy (1988) attempts Path Analysis using standardised partial regression coefficients (beta coefficients or beta weights) as the structural parameter for confirmatory analysis of her path model. The idea was that path coefficients for the paths in a causal model can be estimated by regressing each endogenous variable in the model by each of the variables which were hypothesised to directly impinge on it everytime partialling out the influence of the other relevant variables. This whole exercise has to be done in terms of z scores. The path model then has to be trimmed by eliminating those paths where beta weights have level of significance greater than 0.05. This means that the standard error of each beta weight has to be computed and the 0.95 confidence interval has to be worked out to assess whether a certain beta weight is significantly different from zero. Those paths where path coefficients estimated in terms of beta weights are not significantly different from zero are then eliminated to trim and finalise the path model.

The raw scores of the selected variables presented in the appendix 3 were converted into z scores using the following formula (Garret, 1981, p. 313):

$$z = \frac{x - \bar{x}}{\sigma}$$

where z = standard score, x = raw score, $\bar{x}$ = mean of raw score distribution, and σ = SD of raw score distribution.

As a first step for calculating the beta weights as estimates of path coefficients, the inter-correlations of the variables were computed using the following formula (Garret, 1981, p. 142):

$$r = \frac{\Sigma XY - NM_xM_y}{\sqrt{[\Sigma X^2 - NM^2_x] [\Sigma Y^2 - NM^2_y]}}$$

where r is correlation coefficient, X and Y are obtained scores (here z scores), M_x and M_y are the means of the X and Y series respectively, N is the number of cases.

The matrix depicting the inter-correlations of the variables expressed in z scores is presented in table 5.4. Alongwith the coefficients of correlation, the probability is also presented. The coefficients of correlation found significant at 0.01 level and 0.05 level are being highlighted in table 5.5. The paths hypothesised are also juxtaposed against the correlations in the table 5.5. The picture that the table 5.5 presents is that out of the nineteen paths hypothesised, six paths do not have a corresponding significant correlation. The remaining thirteen paths have corresponding correlations significant at 0.01 level.

Further, path coefficients were estimated by calculating beta weights for each path, partialling out the influence of other convergent paths. For instance, the path coefficient for path E (p_E) was estmated by calculating beta weight $\beta_{52.34}$. This is the partial regression coefficient using z scores between G and SC, partialling out the influence of HIP and EEH.

TABLE 5.4: Correlation matrix of variables in z scores

S. No. Variable	1 ME	2 G	3 HIP	4 EEH	5 SC	6 CD	7 AS	8 AA
1 ME	1.000							
2 G	0.2390 P=0.038	1.000						
3 HIP	0.1579 P=0.123	-0.0608 P=0.328	1.000					
4 EEH	0.3824 P=0.002	0.1364 P=0.158	0.6667 P=0.000	1.000				
5 SC	0.3014 P=0.012	-0.1131 P=0.203	0.4723 P=0.000	0.6198 P=0.000	1.000			
6 CD	0.2591 P=0.027	0.2979 P=0.013	0.4138 P=0.001	0.4968 P=0.000	0.2807 P=0.018	1.000		
7 AS	0.4380 P=0.000	-0.0851 P=0.266	0.6482 P=0.000	0.7564 P=0.000	0.7163 P=0.000	0.3407 P=0.005	1.000	
8 AA	0.3165 P=0.009	0.1753 P=0.098	0.3597 P=0.003	0.5182 P=0.000	0.5807 P=0.000	0.2103 P=0.060	0.6988 P=0.000	1.000

TABLE 5.5: Hypothesised paths and significant correlations among variables

Variables	ME	G	HIP	EEH	SC	CD	AS	AA
ME								
G	*							
HIP	A	B						
EEH	**C	D	**					
SC	*	E	**F	**G				
CD	*	*	**H	**I	*			
AS	**	J	**K	**L	**M	**N		
AA	**		**O	**P	**Q	R	**S	

** = Significant at 0.01 level.
 * = Significant at 0.05 level.

The beta weights were calculated using the following formula (Garret, 1981, pp. 411-412):

$$\beta_{12.34\ldots n} = r_{12.34\ldots n} \frac{\sigma_{1.234\ldots n}}{\sigma_{2.134\ldots n}}$$

(using z scores)

where

1. $r_{12.34\ldots n}$ is partial correlation computed using the formula:

$$r_{12.34\ldots n} = \frac{r_{12.34\ldots(n-1)} - r_{1n.34\ldots(n-1)}\, r_{2n.34\ldots(n-1)}}{\sqrt{1 - r^2_{1n.34\ldots(n-1)}}\sqrt{1 - r^2_{2n.34\ldots(n-1)}}}$$

and 2. $\sigma_{1.234\ldots n}$ and $\sigma_{2.134\ldots n}$ are standard errors of estimate in terms of partial σs. They are computed using the following formula:

$$\sigma_{1.234\ldots n} = \sigma_1\sqrt{1 - r^2_{12}}\sqrt{1 - r^2_{13.2}}\sqrt{1 - r^2_{14.23}} \ldots \sqrt{1 - r^2_{1n.23\ldots(n-1)}}$$

The standard error of the beta weight was calculated using the following formula (Garret, 1981, p.415):

$$\sigma_{\beta_{12.34...m}} = \frac{\sigma_{1.234...m}}{\sigma_{2.34...m}\sqrt{N - m}}$$

in which m = number of variables correlated, and N = size of the sample, and (N - m) = degree of freedom.

The 0.95 confidence interval for $\beta_{12.34...n}$ is $\beta_{12.34...n}$ +/- 1.96 x $\sigma_{\beta\ 12.34...m}$ (Garret, 1981, p.415). If zero is within the 0.95 confidence interval, then the beta weight and therefore the path coefficient is not significant at 0.05 level.

The above mentioned exercise was done and the path coefficients were estimated and tested for significance for all the nineteen paths in the hypothesised path model. The summary of beta weights from which path coefficients were obtained for testing the hypothesised path model is presented in table 5.6. Out of the nineteen hypothesised paths only six were found significant at 0.05 level. Those paths found significant on empirical verification are highlighted in table 5.7.

One of the most significant evidences of the present analysis is that the Gender of the child (G) gets eliminated altogether from the path model. Similarly, Home Interaction Pattern (HIP) also gets eliminated. The paths found significant are highlighted in figure 5.2. It is further restructured and trimmed as represented in figure 5.3. What figure 5.3 presents is finally the

TABLE 5.6: Summary of beta weights from which path coefficients were obtained for testing the hypothesised path model

S.No.	Path	Variable From	Variable To	Beta Weight (Estimate of Path Coefficient)	Standard Error	0.95 Confidence Interval From	0.95 Confidence Interval To	Significance
1.	A	ME	HIP	0.1853	0.1408	-0.0907	0.4613	NS
2.	B	G	HIP	0.1040	0.1384	-0.1673	0.3753	NS
3.	C	ME	EEH	0.3754	0.1321	0.1165	0.6343	S*
4.	D	G	EEH	0.0475	0.1298	-0.2069	0.3019	NS
5.	E	G	SC	-0.2022	0.1135	-0.4247	0.0203	NS
6.	F	HIP	SC	0.0562	0.1515	-0.2407	0.3531	NS
7.	G	EEH	SC	0.6424	0.1528	0.3429	0.9419	S*
8.	H	HIP	CD	0.1485	0.1584	-0.1620	0.4590	NS
9.	I	EEH	CD	0.3975	0.1586	0.0866	0.7084	S*
10.	J	G	AS	-0.0824	0.0850	-0.2490	0.0842	NS
11.	K	HIP	AS	0.2036	0.1073	-0.0067	0.4139	NS
12.	L	EEH	AS	0.4274	0.1258	0.1808	0.6740	S*
13.	M	SC	AS	0.3375	0.0963	0.1488	0.5262	S*
14.	N	CD	AS	-0.0312	0.0937	-0.2149	0.1525	NS
15.	O	HIP	AA	-0.1661	0.1396	-0.4397	0.1075	NS
16.	P	EEH	AA	0.0182	0.1724	-0.3197	0.3561	NS
17.	Q	SC	AA	0.1567	0.1360	-0.1099	0.4233	NS
18.	R	CD	AA	0.0070	0.1157	-0.2196	0.2336	NS
19.	S	AS	AA	0.6760	0.1793	0.3246	1.0274	S*

S* = Significant at 0.05 level.
NS = Not significant.

TABLE 5.7: Paths found significant on testing the hypothesised path model

			From							
	S.No.	Variable	1 ME	2 G	3 HIP	4 EEH	5 SC	6 CD	7 AS	8 AA
To	1	ME								
	2	G								
	3	HIP								
	4	EEH	C							
	5	SC				G				
	6	CD				I				
	7	AS				L	M			
	8	AA							S	

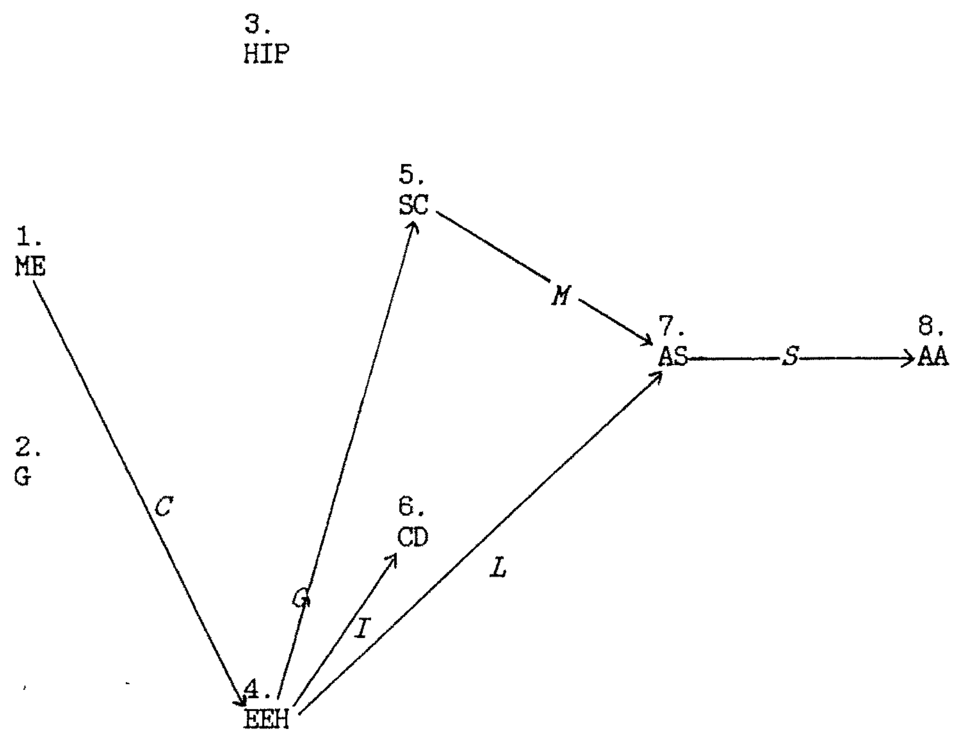

FIGURE 5.2: Paths found significant on testing the hypothesised path model

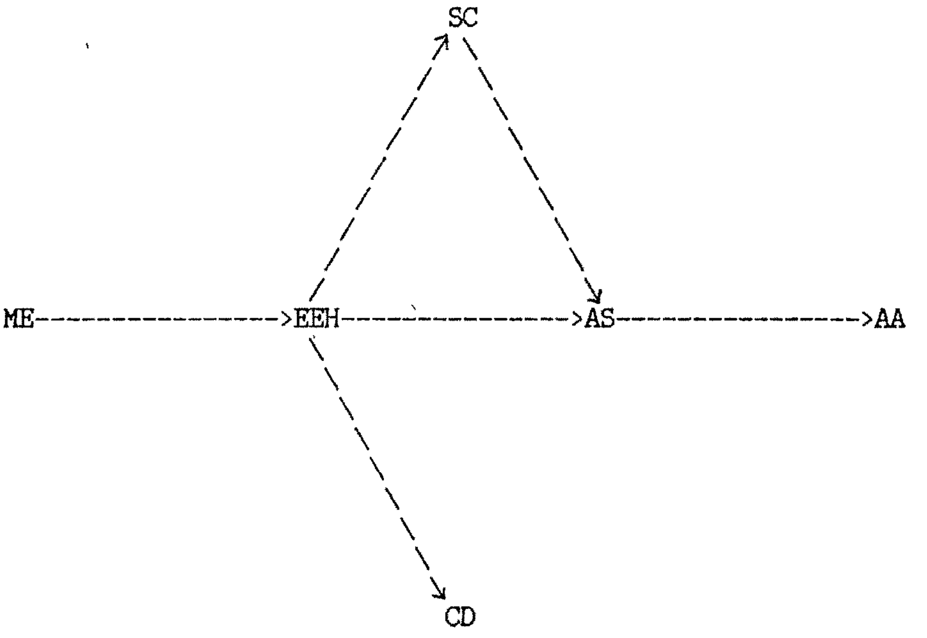

FIGURE 5.3: Empirically tested and trimmed path model

empirically validated and trimmed path model.

5.5.0 Interpretation and discussion

How does this exercise of path analysis aid to respond to the research question that was attempted to be answered? As a result of this exercise, it can now be stated with confidence that there is to be a clear causal link between:

a. Mother's Education and Educational Environment at Home;
b. Educational Environment at Home and Social Competence;
c. Educational Environment at Home and Adjustment to School;
d. Educational Environment at Home and Cognitive Development;
e. Social Competence and Adjustment to School; and
f. Adjustment to School and Academic Achievement.

There are three paths clearly emerging in the tested and trimmed path model. They are:

a. Mother's Education ---> Educational Environment at Home ---> Social Competence ---> Adjustment to School ---> Academic Achievement;

b. Mother's Education ---> Educational Environment at Home ---> Adjustment to School ---> Academic Achievement; and

c. Mother's Education ---> Educational Environment at Home ---> Cognitive Development.

Among the home variables, Educational Environment at Home emerges as a crucial 'anchor variable' in this network of causal relationships. Home Interaction Pattern on the other hand, gets eliminated in the final path model, although there was a very high correlation between Home Interaction Pattern and Educational Environment at Home (tables 5.5 and 5.6). The logical status of these two variables are very close to each other and it is quite possible that they may in actuality be two components of the same composite variable. This hypothesis can be verified only through factor analysis, which a future study may attempt.

Of the two background variables, Mother's Education emerges as an important exogenous variable which is the sole starting point of all causal paths in the trimmed path model. Implications of this are very significant. Emergence of Mother's Education as the only exogenous variable in the empirically verified and trimmed path model, in fact, gives an empirical basis to the popular saying: "Educate a woman; and thereby you educate a family". This evidence brings out rather clearly the nexus between female adult literacy and universalisation of primary education. Mother's role in the family is crucial. She occupies the 'star' position in the family sociogram. She spends most time with the children. Her vision of what her children should become, therefore, is most crucial in the child's development and school performance, and this vision is definitely influenced by her education. However the hypothetical causal model tested in this study did not include variables such as father's education, parental occupation, etc. Perhaps a more comprehensive path analytical study will reveal the relative importance of Mother's Education vis-a-vis other background variables.

The other exogenous variable in the hypothesised causal model, viz., Gender of the child, has been clearly eliminated from the final trimmed Path Model. This tendency was seen uneqivocally even in the inter-correlation matrix (tables 5.5 and 5.6). Though one expects in the Indian situation that there will be differential aspirations on the part of the family for female and male children and thus differential support that the family gives to their schooling, the present study does not indicate this. Perhaps, the children under study here are too young to be subjected to such gender based differentiation. Or, this must be due to the fact that urban families tend to attach more importance to female education than their rural counterparts do. A third and a rather far-fetched guess one may venture in explainig this, is the relatively liberal attitude towards women seen in urban Gujarati milieu. This final hypothesis, however, needs a thorough social-anthropological analysis before being given credance.

In the hypothetical causal model (figure 3.7) family variables were hypothesised to be causally linked with school performance variables, mostly through either of the two intermediary developmental variables, viz., Social Competence and Cognitive Development. However, in the tested and trimmed path model,in the paths linking family variables and school performance variables cognitve development does not emerge as an intermediary variable. Infact, the causal paths hypothesised between Cognitive Development and Adjustment to School, and Cognitive Development and Academic Achievement were eliminated in the tested and trimmed path model. The third path that gets delienated in the trimmed path model is Mother's Education -----> Educational Environment at Home -----> Cognitive Development. The path terminates in a blind end at Cognitive Development, and does not go further towards academic achievement. On the other hand, the significant path that emerges is Mother's Education ----> Educational Environment at Home ----> SocialCompetence ----> Adjustment to School ----> Academic Achievement. This shows very explicitly the predominance of social content in Academic Achievement as opposed to cognitive content. This is clearly in divergence with the popular psychometric view regarding predictors of Academic Achievemnt. This is also in contradiction with many previus research findings. However, this finding can be understood if one recalls that for the present study Academic Achievement is taken more as a measure of the child's conformity to the academic expectations of the school, and less as a measure of the child's scholastic achievement defined independent of the school milieu. One must recall here that in the present study it was deliberately decided that the instruments to measure Academic Achievement would be teacher-made tests. This deliberate choice was rooted on the preoccupation of the present study to understand the child in the school milieu 'as it is' and the extent to which the child responds to it; the idea was not to construct an abstract reality of school 'as it should be'.

What distinctly emerges out of the present study therefore, is that the demands the school makes on the child, at least in the first year of formal schooling, is more social in content and less cognitive. Hence, the child's response to it also emerges to be more social and less cognitive. This is precisely the reason why among developmental characteristics of the child, social competence proves to be a significant predictor of school performance, while cognitive development does not emerge in the same manner. This configuration is of course likely to change as the child advances in her schooling. In higher classes of primary school social adjustment to school will have been a battle already won and perhaps cognitive development may then emerge as a predictor of school performance. However, it calls for a longitudnal or a multi-cross sectional study of children in primary classes, for such a hypothesis to be verified.

One last point that needs mention here is the fact that the confidence range of the estimated path coefficients are rather wide indicating the lack of stability in these values (table 5.7). Even in cases where the path coefficients were significant, they were seen to be lacking is stability. This, in fact, is rooted in one limitation of the present study, and that is the limited sample size. At the same time it must be mentioned that path analysis was just one part of the present study and the evidences gathered through this are to be corroborated by qualitative information collected through idiographic approach. The empirically verified and trimmed path model, which is the outcome of the quantitative part of the study, should therefore be treated as a necessary first step in progressive spiral of scientific exploration of this phenomenon. The present trimmed path model may in future be subjected to further theoretical articulation and path analysis with a larger sample, in order to elicit a more stable picture of the phenomenon.

REFERENCES

Garret, H.E. (1981) *Statistics in psychology and education,* Bombay: Vakils, Feffer and Simons.

James, L.R., Mulaik, S.A. and Brett, J.M. (1982) *Causal analysis: Assumptions, models and data,* Beverly Hills: Sage, p. 59.

Stuessy, C.L. (1988) 'Path analysis: A model for the development of scientific reasoning abilities among adolescents'. *Journal of Research in Science Teaching,* 26(1), pp. 41-53.