Committee II
Theoretical Empiricism: A General
Rationale for Scientific Model-Building

Draft -- for
Conference Distribution only

ON PATH MODELING IN EDUCATIONAL RESEARCH

by

Norbert Sellin
University of Hamburg
Dept. of Education
West Germany

Discussion Paper

on

Richard Noonan's

EVALUATION OF SCHOOL SYSTEMS USING PARTIAL LEAST SQUARES (PLS):
AN APPLICATION IN THE ANALYSIS OF OPEN SYSTEMS

The Thirteenth International Conference on the Unity of the Sciences
Washington, D.C. September 2-5, 1984

Introduction

Richard Noonan's paper gives a comprehensive presentation which captures a general theoretical model for school system evaluation, a set of propositions about factors and processes influencing observed relations, and the empirical test of a complex model. The paper also demonstrates the use of hierarchically structured PLS models, which constitutes a newly developed methodology for the analysis of large and complex path models. It is an impressive piece of work. As it is impossible to cover all parts of Noonan's paper, my comments will concentrate on two aspects, namely (a) the set of general propositions about factors which have an impact on empirical path modeling results, and (b) the estimation of hierarchically structured PLS models. More specifically, my remarks will be clustered around the following topics :

(1) Multiplicity and Effectivity Principle. I shall argue that both propositions are not generally applicable because they refer to special data constellations. Section 1 presents a counter-example.

(2) Specification Error. An important aspect missing from Noonan's presentation is the notion of specification error. Section 2 discusses some implications of the specification error problem in respect to Noonan's PLS analysis.

(3) Estimation of Hierarchical Structures. The statistical implementation seems to be the most problematic aspect of the concept of hierarchical structures. Section 3 presents an empirical example and examines some peculiarities of Noonan's algorithm for estimating hierarchically structured latent variables. Section 4 briefly describes two alternative algorithms.

1. Multiplicity and Effectivity Principle

Let me start with some brief remarks on the multiplicity principle and the effectivity principle. The multiplicity principle states that the total effect of a given cause variable is a function of the number of paths through which the causal influence operates on a given criterion. Closely related to this statement is the effectivity principle which states that the total effect depends on the magnitude of the direct effects making up the causal chains through which the indirect effects operate. Both statements are closely related to basic principles of path modeling. First, the total effect of a given variable is defined as the direct effect (which may be zero) plus all indirect effects operating through other variables in the system. Second, indirect effects are defined as the product of the direct effects involved in the associated causal chains. Clearly, if the direct effect of a given variable is positive and if all indirect effects are also positive, the total effect wil increase as a function of the number and magnitude of indirect paths. However, this is only true if all direct and indirect effects can be assumed to operate in the same direction. So there are counter-examples to Noonan's principles. Consider the model depicted in Figure 1.

Figure 1

This example states that more drill will have a positive effect on academically engaged time but a negative effect on attitude towards subject, while both intervening variables will have a positive effect on cognitive achievement. Given that all path coefficients are, in absolute valuue , equal to, say, 0.5, the total effect of the variable DRILL is zero. It will be noted that there are, of course, many other data constellations implying a zero total effect. The key feature of the above example is the existence of contradicting effects; drill influences achievement positively via academically engaged time, but this positive effect is compensated by a negative indirect effect operating through attitude towards subject.

Neither the multiplicity principle nor the effectivity principle, as formulated by Noonan, take such contradicting influence structures into account. Both propositions are not generally applicable, since they refer to special data constellations, such as the effects of home characteristics, where direct and indirect effects can be assumed to operate in the same direction.

It may be noted that conflicting causal influences constitute an important characteristic of school systems (and open systems in general), a fact which has not been discussed in Noonan's paper. Noonan refers to conflict among goals attributable to the way in which available resources are allocated. It is argued here, however, that such conflict among goals may also be established by contradicting causal influences, such as unexpected or undesirable side effects. For example, there is not necessarily an inherent conflict between the level of academically engaged time, attitude towards subject, and the level of cognitive achievement. Rather, as illustrated in Figure 1, conflicts in this domain may be established by conflicting causal influences of specific characteristics and behaviors. The detection and analysis of conflicting causal influences and particularly the analysis of unexpected or undesirable side effects can be regarded as one of the most important aspects of school system evaluation.

3. Specification Error

The seven propositions formulated by Noonan are concerned with factors and processes influencing observed relations among student, teacher and school variables. More specifically, the propositions deal with factors that have an impact on empirical relations derived from path models. My major criticism of this part of Noonan's paper is that the set of propositions is incomplete. An important aspect missing from Noonan's presentation is the notion of specification error. Specification errors have profound effects on empirical path modeling results and may severely bias inferences about the existence or non-existence of causal effects. Moreover, specification errors may occasionally produce 'findings' which are completely wrong.

All this is well known, and it is also widely recognized that path models formulated and tested in educational research most likely involve misspecifications. In practice, however, the consequences of specification errors are largely ingnored. The following discussion deals with some implications of the specification error problem in respect to Noonan's PLS analysis and in terms of the application of path modeling techniques and the interpretation of empirical results. A useful typology of specification errors has been developed by Deegan (1976). This typology is reproduced in Table 1.

Table 1

The classification shown in Table 1 was developed for single regression equations. It is also applicable to more complex path models, however, if it is assumed that linear-recursive models are employed. Then, each model equation is equivalent to a simple regression equation and the appropriate estimation procedure is Ordinary Least Squares (OLS) regression applied to each model equation separately (see e.g. Land 1973). This is the dominant path analysis technique in educational research. It is also the technique employed in Noonan's PLS analysis for estimating relations among latent variables.

The typology in Table 1 refers to two sources of specification errors, namely (a) the inclusion of irrelevant variables, and (b) the omission of relevant explanatory variables from a given model equation. It is clear that this classification covers only a subset of possible misspecifications. It may, in fact, be argued that linear models are generally misspecified because rather imprecise theoretical expectations are translated into much more precise statistical models which cannot be expected to constitute a perfect representation of social processes. In many cases, however, there is good reason to believe that linear models render reasonably close approximations to reality, given that the tested hypotheses are not fundamentally wrong. The above classification refers to such fundamental errors. A thorough and statistically rigorous formal treatment of specification errors is to be found in the article by Deegan (1976) referred to above and in most textbooks dealing with regression and path analysis (see e.g. Duncan 1975). For the present purposes it should, therefore, suffice to examine just one highly simplified example. Let us assume that the basic model equation can be formulated as :

$$(1) \quad E(Y) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \quad ; \quad \beta_3 = 0$$

The assumption $\beta_3 = 0$ means that X_3 is a superfluous predictor variable. That is, the model equation is overspecified and involves, in Deegan's terminology, a Type A error. It can be readily shown, however, that OLS regression applied to (1) yields 'unbiased' coefficient estimates, given that the predictors are uncorrelated with the residual belonging to the dependent variable Y. The statistical treatment of specification errors is usually based on the distinction between 'true' or correctly specified models and <u>hypothesized</u> models which may or may not be misspecified. 'Irrelevant' predictors are simply defined as variables which have no effect on a given criterion, while 'relevant' predictors are defined as variables having non-zero effects. From a theoretical point of view, this definition is quite unsatisfactory, but this issue will not be discussed here. Suffice it to say that calculations on the basis of equation (1) would suggest that the corresponding 'true' model would include just two explanatory variables, namely X_1 and X_2. Now, a Type C error would occur if the hypothesized model were given by :

$$(2) \quad E(Y) = \beta_1^* X_1 + \beta_3^* X_3$$

To simplify the presentation, three additional assumptions will be made: (a) all variables are standardized to zero mean and unit variance, (b)

X_1 and X_3 are uncorrelated, and (c) OLS point estimates obtained from (1) are equal to the corresponding population parameters. The last assumption means that sampling considerations will be discarded. It is evident that inappropriate sampling can severely bias coefficient estimates, irrespective of whether the tested model is correctly specified or misspecified. On the basis of the above assumptions, the bias resulting from equation (2) can be evaluated from the expressions :

$$
(3) \qquad \begin{aligned} \hat{\beta}^*_1 &= \hat{\beta}_1 + r_{12}\hat{\beta}_2 \\ \hat{\beta}^*_3 &= 0 \;\; + r_{32}\hat{\beta}_2 \end{aligned}
$$

where the r's denote the correlations between the included predictors and and the omitted 'true' model predictor X_2. Equation (3) illustrates two things. First, non-zero effect estimates are generally obtained for irrelevant predictors (i.e. $\hat{\beta}^*_3 \neq 0$) and, second, the estimated effects of included 'true' model predictors are generally biased (i.e. $\hat{\beta}^*_1 \neq \hat{\beta}_1$). The magnitude and the direction of the bias depends on the values and the signs of the involved coefficients. $\hat{\beta}^*_1$, for instance, may be an overestimate or an underestimate of the 'true' model coefficient $\hat{\beta}_1$; note that this includes the possibility that $\hat{\beta}^*_1$ is zero, or that $\hat{\beta}^*_1$ is negative while the correct effect estimate is positive.

Turning to substantive considerations of specification errors, the first obvious question is which error type occurs most frequently in educational research. Unfortunately, the most frequent type of specification error is presumably Type C error. That is, empirical research in education probably deals most often with models which are simultaneously overspecified and underspecified. Much of the large-scale reserach in education, such as the IEA Six-Subject-Survey, can be characterized as huge fishing expeditions where data are collected on hundreds of variables covering a large number of dimensions in the hope of capturing something big and important. To be sure, schooling is a highly complex social phenomenon, and research on schooling therefore requires data collection on a fairly broad range of dimensions. However, educational research is, in practice, rarely guided by sufficiently clear theoretical expectations of relevant and irrelevant aspects of schooling, and is virtually never guided by explicit a priori conceptualizations of alternative models to be tested. As a consequence, the educational researcher is faced with huge data sets which probably contain a great deal of irrelevant information, but which nevertheless constitute a limited data base in that relevant information is often missing. In this situation, one basic task of the data analysis is to identify factors which have no impact on specified criteria. In terms of the application of path modeling techniques, a frequently used analysis strategy is to start with fully recursive models, that is, with models where all possible recursive relations are specified, and to apply standard significance tests in order to identify irrelevant explanatory variables. This approach can be criticized on several grounds. One aspect is that classical significance tests require several fairly strong assumptions on the distribution of residuals which are often highly unrealistic. But even if distributional assumptions seem to be justified, a general theoretical argument still applies, namely that this analysis stategy requires the assumption that only Type A error is present. That is, it must be assumed that each model equation involves all relevant explanatory variables. If this cannot be assumed or, alternatively, if it must be realistically assumed that Type C error is present, this approach is clearly nonsense, since statistically significant effects will most likely be obtained for irrelevant predictor variables. In terms of the above example, $\hat{\beta}_3^*$ will not only be non-zero, but may also turn out to be highly significant.

To be sure, the general critique formulated above does not fully apply to Noonan's paper. In fact, an important contribution of Noonan's presentation is its emphasis on theory and the argument that theory is required at all stages of research, from the data collection to the data analysis and the interpretation of results. The theoretical part of his paper states very clearly which factors are expected to have a direct influence on science achievement, for example, and which factors are expected to exhibit indirect effects only. It is to be noted, however, that the actual data analysis described in Noonan's paper deviates from the theoretical model presented earlier. For example, the macro-model equation for science achievement involves constructs which were not assumed to have a direct effect on science achievement (e.g. *SCHSTR, *TEASTR and *TEABEH). The corresponding equation includes, in fact, all causally prior variables involved in the macro-model, and Noonan's discussion of the corresponding results also refers to standard significance tests. In short, Noonan's data analysis has much in common with the general analysis strategy criticized above. This does not mean, however, that a strictly confirmatory approach would have been appropriate. In view of the theoretical knowledge available in the field of school system evaluation, a compromise between exploratory and confirmatory path modeling strategies is certainly required. The point I want to make here is that the data analysis should incorporate explicit considerations on possible misspecifications and their impact on given results. Of course, the 'true' model structure is not known, and formal derivations are, at first glance, not very helpful for assessing the influences of specification errors on specific results. Equation (3), for instance, indicates that virtually everything can happen if a given model involves Type C error; the estimated path coefficients may constitute underestimates or overestimates and may even change signs if the tested model is misspecified. It is argued here, however, that it is often possible to arrive at reasonable theoretical speculations about model misspecifications, such as speculations about omitted explanatory variables and their interrelations with included variables, and to incorporate such considerations into the interpretation of results. As a matter of fact, Noonan's interpretation of the effect of *TEABEH on science achievement constitutes an example of such theoretical speculations. He argues that the direct effect of *TEABEH reflects an indirect effect operating via student learning behaviors omitted from the tested model. In other words, Noonan refers to Type C error involved in the model and speculates about relations between the construct *TEABEH and omitted variables reflecting student behaviors. Whereas it may well be plausible that the effect of *TEABEH on student achievement reflects an important aspect of educational processes, it is imperative to note that theoretically relevant student variables are omitted from the model, and that this omission has several fundamental implications in terms of the application and interpretation of path analysis techniques. Among other things, this omission prohibits the use of classical significance tests, unless it can reasonably be assumed that the existing Type C error does not affect other parameter estimates.

This assumption, however, appears to be fairly strong and therefore should not be taken for granted; quite to the contrary, it would have seemed appropriate to extend considerations of possible model misspecifications to other model relations (e.g., the negative effect of curriculum on science achievement).

All this is not to say that Noonan was unaware of the problems outlined above. In fact, he implicitly refers to the problems of misspecifications in several parts of his presentation (e.g., in his discussion of the 'Rules Principle'). The point I want to make is that these difficulties should have been made explicit, all the more so, since little is known about the effects of specification errors in PLS modeling. As the PLS procedure heavily relies on regression-based methods, it can be expected that the well-known results derived for conventional path analysis techniques carry, at least in part, over to PLS models, but in view of the importance of specification error problems, further research in this field is clearly needed.

4. Estimation of Hierarchical Structures

This section examines some conceptual and statistical aspects of Noonan's algorithm for estimating hierarchically structured PLS models. As will be shown below, Noonan's algorithm departs from key principles of the basic PLS design. This has consequences in terms of the correspondence between so called macro-models, which include hierarchically structured latent variables (lv's), and micro-models, which aim at an investigation of particular model parts. The introduction of hierarchically structured lv's requires some modifications of the basic PLS procedure, and it is therefore not to be expected that macro-models and corresponding micro-models yield numerically equivalent results. It is desirable, however, that macro-models and corresponding micro-models generally imply similar conclusions in terms of the relative importance of lower level lv's and in terms of the relative importance of manifest variables making up specific constructs. The following discussion deals with peculiarities of Noonan's procedure which may lead to a substantive lack of correspondence between macro and micro modeling results.

To illustrate possible differences between macro-models and corresponding micro-models, the PLS model presented in Figure 2 will be used as an empirical example. The model is based on data from the Classroom Environment Study conducted by the International Association for the Evaluation of Educational Achievement (IEA). A major aim of this study is to examine relations among observed teacher behaviors, student behaviors and student achievement (cf. Ryan 1981). The model shown in Figure 2 is based on data from 65 fifth-grade mathematics classes observed in an Asian country. A short description of the included constructs and the associated manifest variables is given in Table 2. The model relates an observer rating of the percentage of academically engaged students in class (ENGPCT) to four intervening constructs comprising indicators of different types of teacher-to-student interactions. These intervening constructs are related to two exogenous blocks labelled as TPERC and ACTIVITY. TPERC involves two questionaire items reflecting the teacher's perception of the average class abiltity at the beginning of the observed learning unit. Note that the scaling of the manifest variables implies a high value of TPERC if the teacher perceived the class as of comparatively low ability. The block ACTIVITY involves two observational indicators reflecting general activities (seatwork and lecture) in which the classes were involved. ACTIVITY is scaled in such a way that LECTURE is negatively weighted while SEATWRK is positively weighted. A high value indicates, thus, that a comparatively large amount of seatwork and a relatively small amount of lecturing activities was observed. It may be noted that the two activity categories were found to account, on the average, for 80 to 90 percent of all observed class activities. It is not possible and for the present purposes not necessary to describe the data collection and the data preparation procedures in detail. It should suffice to note that each class was observed at six lessons, and that the observations yielded about 5,000 data points per class. For this reason, some data aggregation was necessary. In the present case, the recorded teacher-to-student interactions were aggregated to the teacher or class level by transforming the total frequencies with which each interaction was observed to percentages of all coded interactions. Hence, the manifest variable LE, for example, indicates the percentage of all interactions classified as 'verbal lecture'. A similar procedure was used to aggregate the variable ENGPCT and the activity categories included in the block ACTIVITY. The model shown in Figure 2 is, thus, based on highly aggregated observational data reflecting purely quantitative aspects of classroom instruction. It is to be emphasised that the presented model constitutes a simplified submodel taken from more extensive PLS analyses. One simplifaction was to specify all constructs as 'outward' blocks. The model is primarily intended to serve as a numerical example and, therefore, no detailed interpretation will be given.

Figure 3 displays a hierarchical model that corresponds to the PLS model shown in Figure 2. The intervening blocks TEACH to MANAGE are summarized by a higher order construct labelled as ACTIVE. This higher order construct can be interpreted as reflecting the extent to which the teacher was actively involved in different types of classroom instruction. This interpretation is based on the fact that the included interaction categories accounted, on the average, for about 60 percent of all coded interactions while the remaining 40 percent were usually due to student responses and, to a large portion, classified as 'silence/absence of interactions'. The included interaction categories covered, therefore, nearly 100 percent of the observed teacher initiated interactions. The hierarchical model depicted in Figure 3 has been specified in accordance with two basic principles noted by Noonan and Wold (1983: 284): (1) all influences on a given hierarchy operate via lower level constructs, and (2) all influences of lower level lv's are summarized by the top level construct; i.e. it is assumed that no influences emanate from lower level lv's directly to variables outside the hierarchy. For comparative purposes,

Figure 2 and Figure 3 display the estimated path coefficients, the R^2-values, and the total effects on ENGPCT. The coefficients displayed in Figure 2 were estimated by the basic PLS procedure, and the coefficients shown in Figure 3 were estimated by Noonan's algorithm, with ACTIVE specified as an 'outward' block. It can be seen that the model results differ in several respects. The most striking difference concerns the R^2-value of ENGPCT; the R^2-value determined from the hierarchical model is larger than the R^2-value obtained from the micro-model. Noonan states that R^2-values of hierarchical PLS models are generally smaller than R^2-values determined from corresponding micro-models. The above example demonstrates that this assertion is not generally true. A loss of predictive power occurs with regard to the <u>overall</u> model results, however, since the mean R^2-value of the hierarchical model is equal to 0.142 while the mean R^2-value of the micro-model is equal to 0.173. This is because the hierarchical model yields smaller R^2-values for the intervening constructs TEACH, QUEST and QUINT. In terms of the estimated effect coefficients, a comparatively large difference is observed for the total effect of TEACH on ENGPCT for example. The micro-model effect is close to zero (-.026) while the macro-model yields a relatively strong negative effect (-.110).

The differences alluded to above occur because Noonan's algorithm departs from key features of the basic PLS procedure used to estimate the micro-model. To show this, it is necessary to present some aspects of both estimation procedures in greater detail. A thorough exposition of the basic PLS procedure and Noonan's procedure for estimating hierarchical structures is to be found in Noonan and Wold (1983).

The basic PLS procedure involves two steps. The first step is the iterative estimation of the weights defining the latent variable estimates, and the second step is the non-iterative estimation of path coefficients and loadings by standard procedures. The weights are estimated using so called adjacent constructs. Adjacent constructs are generally defined as linear composites of all lv's with which a given lv is directly connected, irrespective of whether these adjacent lv's are regressors or regressands. For example, in Figure 2 the adjacent constructs of the intervening lv's TEACH to MANAGE are defined as linear composites of the blocks TPERC, ACTIVITY and ENGPCT. The weights are determined in accordance with two estimation modes, called 'outward' and 'inward' mode. The weights of 'outward' blocks are computed by simple regressions of each manifest variable on the corresponding adjacent construct; the weights of 'inward' blocks are determined as multiple regression coefficients obtained from a multiple regression of the adjacent construct on the set of manifest variables. Using different estimation modes or modifying the adjacent constructs generally results in different weight estimates and, therefore, generally results in numerically different path coefficients and loadings. As noted by Noonan and Wold (1983: 222; see also Wold 1982), the definition and use of adjacent constructs is the key feature that determines the 'holistic' nature of the basic PLS design. This is because the adjacent constructs directly or indirectly transfer information coming from all blocks involved in a PLS model. For example, in Figure 2 the estimation of the block TPERC utilizes not only information coming from the intervening lv's TEACH to MANAGE but also information coming from the blocks ACTIVITY and ENGPCT, as mediated by the adjacent constructs belonging to the intervening lv's.

Noonan's algorithm for estimating hierarchically structured models departs from the basic PLS design in respect to the determination of adjacent constructs. As indicated in Figure 3, the manifest variables associated with the blocks TEACH to MANAGE are summarized in the higher order construct ACTIVE. Noonan's procedure works, intuitively speaking, 'downward' from the top level of a given hierarchy to the lowest level. The PLS iteration is performed at the top level using <u>subsequent</u> lv's only for determining the corresponding adjacent construct (compare Noonan and Wold 1983: 284). That is, in Figure 3 the weights of the manifest variables belonging to the blocks TEACH to MANAGE are determined with regard to ENGPCT only. These weights are then inserted into the lower level blocks and are transformed so as to give the lower level lv's unit variance. The difference between adjacent constructs used to estimate the intervening blocks is the source of the deviations between the micro-model and the macro-model results described before. Three additional aspects of Noonan's procedure should be noted. (1) Causal chains consisting of hierarchically structured lv's are estimated by moving from the last endogenous part of the model, over intervening parts, to the exogenous part. For example, the model shown in Figure 3 could be estimated by completing, first, an iteration sequence using the blocks ACTIVE and ENGPCT only and by continuing with a second iteration sequence for estimating the blocks TPERC and ACTIVITY. Above all, this implies that information coming from exogenous blocks is not used for estimating intervening and endogenous constructs. (2) Noonan's procedure allows to specify causal relations among lower level lv's. These relations are not incorporated into the iteration process. This feature may constitute an additional source of deviations between micro and macro modeling results. (3) The same estimation mode is applied to all lv's in a given hierarchy. That is, the estimation mode used to determine the top level construct is indirectly also applied to all lower level lv's.

The critical point of Noonan's algorithm for estimating hierarchically structured PLS models is not the mere occurrence of numerical differences between macro-models and corresponding micro-models. Rather, it is the departure from fundamental principles of the basic PLS design which ought to be questioned. While it is true that the introduction of hierarchically structured lv's helps to test large and complex path models, the option chosen by Noonan to modify the original PLS estimation procedure seems to be somewhat less than optimal. My main objection against Noonan's algorithm is that it introduces a reductionist component into the estimation of hierarchical PLS models. As indicated above, causal chains are during iteration divided into separate parts which are consecutively estimated by moving from the the last endogenous segment to the exogenous segment of a given model. This is because the estimation of hierarchical blocks is exclusively directed towards subsequent lv's while information coming from exogenous blocks is discarded. That is, Noonan's procedure treats each hierarchy as if it were an exogenous lv, irrespective of whether a given hierarchy is specified as an exogenous or as an internening block. It is primarily this feature which seems to be questionable. As illustrated by the above example, Noonan's algorithm may occasionally increase the predictive power of intervening hierarchies (because hierarchies are approximated as predictors only), but may simultaneously result in a less optimal prediction of mediating lower level lv's. The basic PLS design, on the other hand, can be expected to approximate intervening blocks as 'best mediating factors', since these blocks are explicitly estimated as predictors <u>and</u> predictands. This property appears to be jeopardized in Noonan's approach. Another critical point of Noonan's procedure concerns primarily practical aspects. As noted by Noonan, hierarchical models are primarily intended to aid in identifying micro-models for more intensive investigations. Due to the differences between estimation procedures, however, macro and micro modeling results are not necessarily consistent in the sense that macro-models and corresponding micro-models generally imply similar conclusions in terms of the relevance of lower level lv's and manifest variables. Hence, PLS analyses on the basis of hierarchical models may occasionally suggest model modifications or interpretations of results which differ substantively from conclusions based on the analysis of corresponding micro-models.

4. Alternative Algorithms

This section briefly describes two algorithms which constitute possible alternatives to Noonan's estimation procedure. For convenience, Noonan's algorithm will be labelled as Algorithm I, and the two alternative algorithms will be labelled as Algorithm II and Algorithm III, respectively. An obvious alternative to Algorithm I is to estimate the lowest level lv's directly, and to construct hierarchies by moving 'upwards', from lower level lv's to higher order constructs. This is the basic idea of Algorithm II and Algorithm III. Algorithm II assumes the same basic model structure as illustrated by Figure 3. That is, influences on the hierarchy are assumed to operate via the lowest level constructs, and influences on variables outside the hierarchy are assumed to operate through the top level construct. For estimating the lowest level lv's, it is necessary to construct appropriate adjacent composites. This is done by considering the predictors of a given lower level lv and the predictands of the top level lv as adjacent lv's. For the model in Figure 3, this results in exactly the same adjacent constructs of the blocks TEACH to MANAGE as used in the corresponding micro-model. Note that the above rule allows to use different estimation modes for lower level lv's and that causal relations among lower level lv's would be incorporated into the iteration process. In a second step, the lower level lv's need to be combined into higher order constructs. This is accomplished by treating lower level lv's in exactly the same way as manifest variables and by applying the normal PLS estimation procedure. That is, higher order constructs may be specified as 'outward' or 'inward' blocks, and the associated lower level lv's are simply treated as indicator variables. The predictands of the top level lv are considered as adjacent lv's; i.e. the estimation of higher order constructs is directed towards subsequent blocks only. Note that the estimation mode applied to higher order constructs may be different from the estimation mode applied to lower level lv's. In short, Algorithm II simply applies the basic PLS procedure to all blocks involved in a given hierarchy and determines adjacent composites in accordance with the assumed flow of information.

Figure 4 illustrates the hierarchical model that corresponds to Algorithm III. This algorithm works in much the same way as Algorithm II, but allows to specify direct effects on the top level construct. Accordingly, predictors and predictands of the top level lv's are defined as adjacent to the lower level lv's as well as adjacent to the top level lv.

For comparative purposes, Table 3 and Table 4 present selected results (total effects on ENGPCT and R^2-values) obtained on the basis of the example used before. It is not possible to discuss the estimation procedures and the respective results in detail. Therefore, the following brief comments should suffice. (1) The largest R^2-value of ENGPCT is obtained by Noonan's algorithm, with ACTIVE specified as an 'inward' block. Since Noonan's algorithm involves a multiple regression of ENGPCT on all manifest variables belonging to the intervening lv's, this is, in fact, the maximum R^2 that can be achieved on the basis of the blocks TEACH to MANAGE. It will be noted, however, that a loss of predictive power occurs, again, with regard to the overall model results. (2) Due to the use and definition of adjacent constructs, Algorithm II results in a very close correspondence between macro-model and micro-model coefficients. If ACTIVE is specified as an 'inward' block, the macro-model is, in fact, equivalent to the underlying micro-model. This is a direct consequence of the way in which Algorithm II has been designed. It must be noted, however, that the numerical equivalence is due to the fact that ENGPCT involves just one manifest variable. Differences will normally occur if a given endogenous construct involves multiple indicators. In general, however, Algorithm II can be expected to produce macro-model results which are fairly close to the results obtained from corresponding micro-models. (3) The outcomes of Algorithm III appear to be rather unsatisfactory. The procedure results in the largest amount of loss of information and, therefore, some further mosifications seem to be necessary. The examples and the results presented here are, of course, by no means sufficient to come to definitive conclusions, and further research on the estimation of hierarchical structures will therefore be necessary.

References

Deegan, J. : The Consequences of Model Misspecifications in Regression Analysis. Multivariate Behavioral Research, 1976: 237-248

Duncan, O.D. : Introduction to Structural Equation Models. New York 1975

Land, K.C. : Identification, Parameter Estimation and Hypothesis Testing in Recursive Sociological Models. In: Goldberger, A.S. and Duncan, O.D. (Eds.): Structural Equation Models in the Social Sciences. New York 1973

Ryan, D.W. : An Organizing Framework for the IEA Classroom Environment Study: Teaching for Learning. Internal Report. The Ontario Institute for Studies in Education. Toronto 1981

Noonan, R. and Wold, H.: Evaluating School Systems using Partial Least Squares. Evaluation in Education: An International Review Series, Vol. 7, Pergamon Press 1983

Wold, H. : Soft Modeling: The Basic Design and Some Extensions. In: K.G. Jöreskog and H. Wold (Eds.): Systems under Indirect Observation Part II, Amsterdam 1982, 1-54

Table 1 Typology of Specification Errors*

Hypothesized model ...

		Incorrectly includes variables: No	Incorrectly includes variables: Yes
Incorrectly omits variables	No	Correct model results	Type A error
	Yes	Type B error	Type C error

* Adopted from Deegan (1976)

Table 2 Basic Block Structure; Figure 2

Block/Lv	Lv-Label	Mv/Mv-Label
TPERC	Teacher perception of class ability	RNEED: Majority of class needs remedial work (1=yes; 0=no)
		ABIL: Low average class ability (1=yes; 0=no)
ACTIVITY	Major class activity	LECTURE: Lecture/explain/demonstration
		SEATWRK: Seatwork on written assignments
TEACH	Teaching students	LE: Verbal lecture
		LEM: Lecture with materials
		EX: Use of examples
QUINT	Questioning/Evaluation/Feedback	PB: Probes
		RD: Redirect questions
		EF: Effectiveness questions
		AC: Acknowledge correct answer
QUEST	Questions	QTGP: Questions - teacher to group
		QTST: Questions - teacher to student
MANAGE	Classroom management/discipline	DI: Discipline
		PR: Procedural interactions
		DR: Directives
ENGPCT	Pct academically engaged students	ENGP: Observer rating - percentage of academically engaged students in class

Table 3 Total effects on ENGPCT: Disaggregated model and hierarchical models

	Disaggr. model	I		II		III	
		A	B	A	B	A	B
TPERC	104	100	117	097	104	070	075
ACTIVITY	-175	-197	-155	-214	-175	-227	-238
TEACH	-026	-110	-145	021	-026	084	099
QUINT	213	257	225	238	213	195	167
QUEST	212	165	167	225	212	224	258
MANAGE	-299	-160	-330	-198	-299	-074	-144
ACTIVE	-	445	490	423	437	371	394

Note: A: Hierarchical block ACTIVE outward
B: " " " inward

Coefficient multiplied by 1000

Table 4 R^2-values: Disaggregated model and hierarchical models

	Disaggr. Model	I		II		III *	
		A	B	A	B	A	B
TEACH	182	009	095	182	182	060	074
QINT	162	144	116	162	162	197	170
QUEST	317	302	284	317	317	294	272
MANAGE	012	058	037	012	012	001	000
ENGPCT	191	198	240	178	191	138	155
Mean	173	142	154	170	173	138	134

* R^2-values for blocks TEACH to MANAGE determined by substitutive prediction

Note: A: Hierarchical block ACTIVE outward
B: " " " inward

Coefficients multiplied by 1000

Figure 1

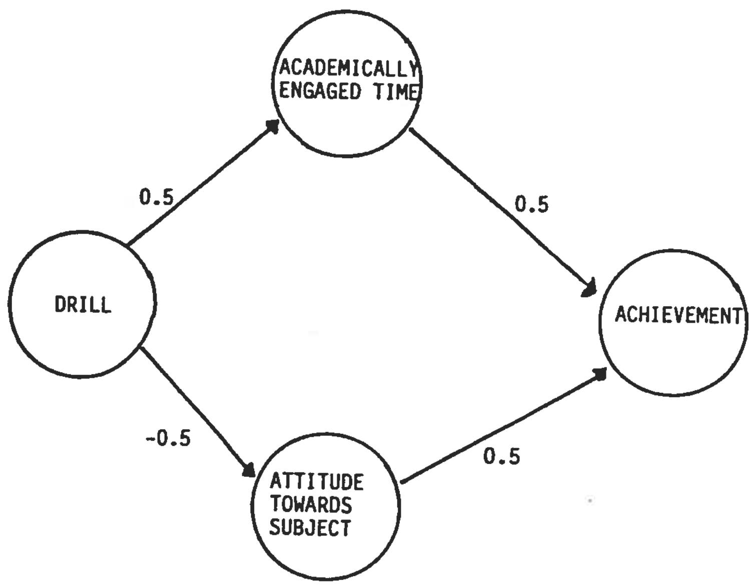

Figure 2 Micro-model; total effects on ENGPCT in parentheses; all coefficients multiplied by 1000

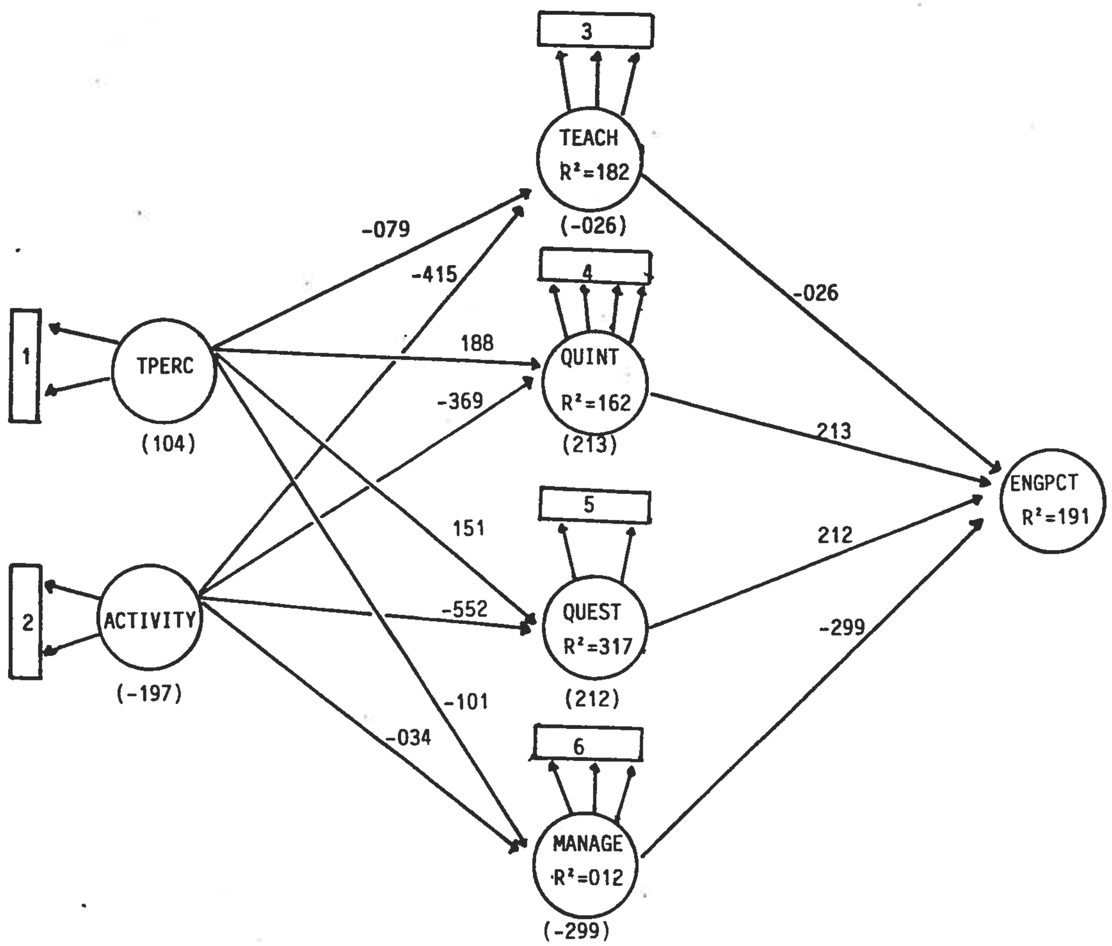

Figure 3   Macro-model; total effects on ENGPCT in parentheses; all coefficients multiplied by 1000

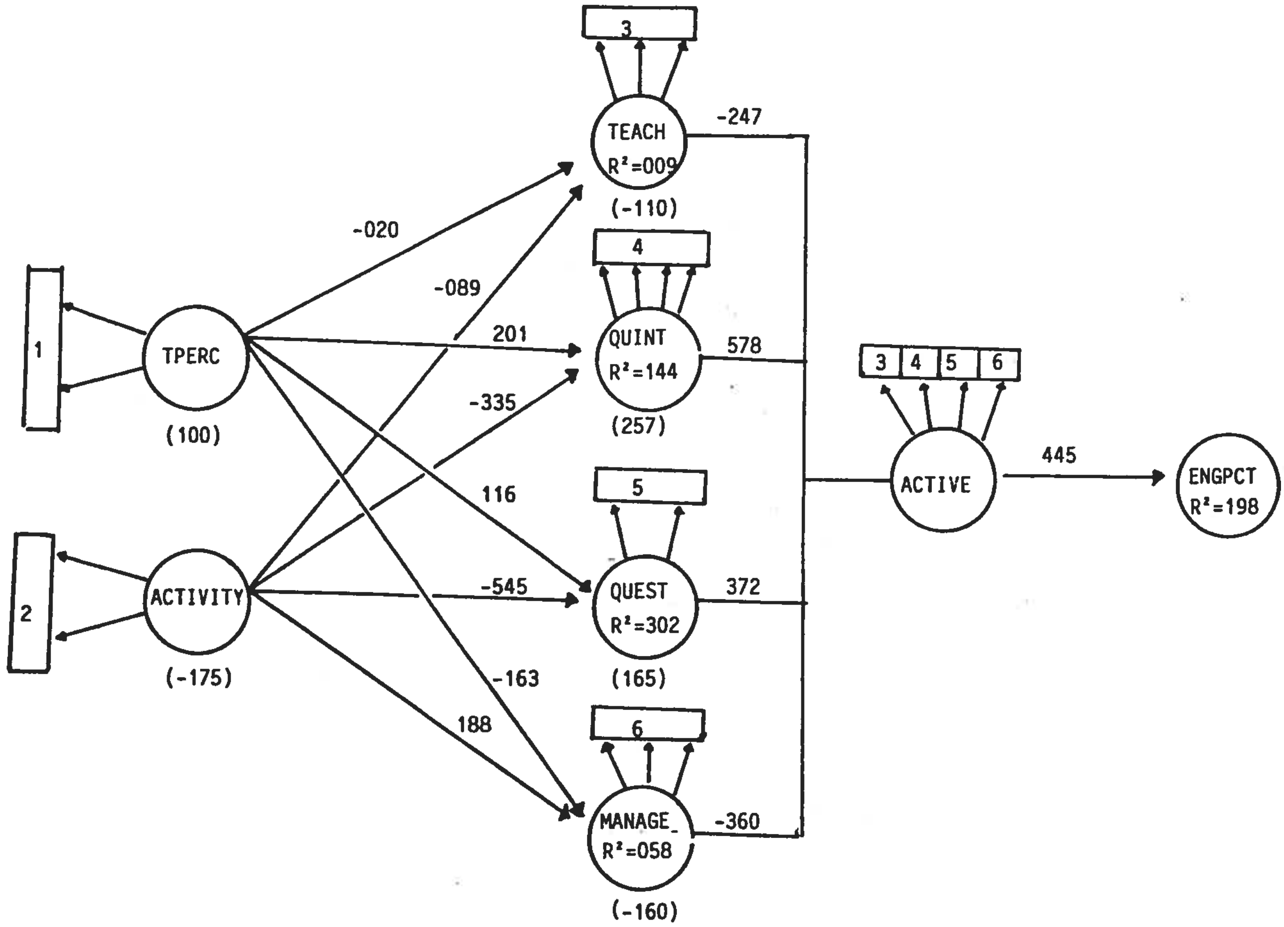

Figure 4