Chapter 10

Patterns of School Performance and Cognitive Development in Early Adolescents - An Application of Prediction Analysis

Alexander von Eye, Michael J. Rovine and Christiane Spiel

Michigan State University,
The Pennsylvania State University, and University of Graz

This chapter is concerned with the prediction of performance in German and Mathematics in early adolescents. Measures of Fluid and Crystallized Intelligence are used as predictors. The question is asked whether Fluid and Crystallized Intelligence measures allow one to make the same prediction of performance in German and Mathematics. This question is approached using both correlational methods and Prediction Analysis (PA) which allows one to statistically test user-specified prediction hypotheses that concern only subtables of cross-classifications. This chapter is also concerned with positioning PA within the frame of statistical methods that use *latent* variable concepts, in particular, Latent Class Analysis (LCA). Specifically, reference is made to definitions of the term *latent*.

1. Predicting Performance in School

Performance in school is of great importance for individuals' personal, academic, and professional development. Reasons for this importance include, at the personal level, the impact that performance in school has on students' self worth. At the academic level performance indicators are used for selection to tracking systems in European high schools. Professional prospects largely depend on performance in school. Predicting performance in school is therefore a widely discussed issue. Early prediction allows parents and teachers to intervene before students are labeled as poor performers (Kühn, 1984; Tramontana, Hooper, & Selzer, 1988).

Theories of cognitive development in early childhood, Kindergarten age, and school age allow one to describe growth and development in domains pertinent to formal school training. Examples of such theories include Piaget's theory of cognitive development (1936; zur Oeveste, 1987) and psychometric approaches to intelligence development (Horn, 1988). Sirsch and Spiel (1994) showed that both approaches allow one to predict performance in school. The psychometric approach seemed to perform slightly better when predicting marks in school.

In this chapter, we predict performance in school using Cattell's (1971; cf. Horn, 1988) distinction between Fluid and Crystallized Intelligence. *Crystallized Intelligence* can be defined as knowledge-based cognitive performance. As such it is largely dependent upon learning. Tests of Crystallized Intelligence typically assess verbal knowledge, following instructions, general information available for science, business, and culture, passage comprehension, and problem definition. *Fluid Intelligence* can be defined as broad reasoning. It is to a lesser extent dependent upon learning. Tests of Fluid Intelligence typically assess inductive reasoning, concept formation, visual auditory learning (code comprehension and translation of language into symbol codes), visual conceptualization, effectiveness in using problem-solving strategies, and numbers reversed memory. This chapter asks, whether the psychometric measures of Fluid and Crystallized Intelligence allow one to differentially predict performance in Mathematics and German (detailed hypotheses are specified in section 2.1).

2. The Study

The data used for the following analyses were collected in the Longitudinal Vienna Developmental Study („Wiener Entwicklungsstudie“) on the cumulative effects of risk factors on performance in school and cognitive development (Spiel, 1995, in press a).

Subject sample. The sample involved N = 93 elementary school students (45 boys and 48 girls). Average student age in fourth grade was 11 years, with range 10.22 - 12.06 years.

Variables and Procedures. For the following analyses we use the variables Performance in German (G) and in Mathematics (M), both at grade four, Fluid Intelligence (F), and Crystallized Intelligence (C).

The intelligence measures were composed using subscale values of the intelligence test AID for children [Adaptives Intelligenz Diagnostikum (Adaptive Intelligence Battery), a revision of the Wechsler Intelligence Scale for Children (WISC-R), Wechsler, 1974; Kubinger & Wurst, 1988]. This test can be used for children between 6 and 16 years of age. Assessment is in individual sessions. From theory and confirmatory factor analysis two subscales were formed that can be labeled as factors of Fluid and Crystallized Intelligence (for details see Spiel, 1995, in press b). The instrument's subscales conform to the Rasch model. The AID is adaptive in nature, that is, the selection of items and, thus, the difficulty of the resulting test instrument depends on a child's performance.

Indicators of performance in Mathematics and German were taken from school records at the end of fourth grade. We did not use grades, because some of the children attended special education classes for the less gifted. Marks are not comparable across tracks. Performance was weighted using complexity of curriculum as weight (see Spiel, 1995).

	Variables			
	Crystallized Intelligence (C)	*Fluid Intelligence (F)*	*German (G)*	*Mathematics (M)*
Minimum[a]	21.5	33.25	3	3
Maximum[a]	62.67	67.25	10	10
Mean	48.15	50.27	8.17	8.34
Sd	8.32	7.92	1.54	1.43
Median	50.33	49.00	8.00	9.00

[a] Values given for performance in school are empirical extrema. Theoretical minima for both performance in German and Mathematics are 0, theoretical maxima are 10.

Table 1: Descriptive values of variables in study

Missing Data Imputation. There were missing data on all variables. We estimated and imputed missing data using the multiple regression method provided by the EQS structural equation modeling program, Version 4 (Bentler, 1992). Descriptive information about the variables used in the following analyses appears in Table 1.

2.1 Hypotheses

We analyze the four variables, C, F, G, and M from two perspectives. The first involves standard correlation techniques. We ask how strong the correlations among the four variables are. The second perspective, complementing the first, involves using PA for differential prediction of performance in school from measures of Fluid and Crystallized Intelligence. We posit the following hypotheses:

1. Above average Crystallized Intelligence allows one to predict above average performance in both Mathematics and German;
2. below average Crystallized Intelligence allows one to predict below average performance in both Mathematics and German;
3. above average Fluid Intelligence allows one to predict above average performance in both Mathematics and German; one might argue that this is not justifiable because performance in school is largely training dependent and thus, above average fluid intelligence alone - which is supposed to be less training dependent - may not be enough to obtain above average results; we suspect, however, that children in this bracket tend to perform above average also, because intelligence factors tend to correlate and the skills tested are trained in school also;
4. below average Fluid Intelligence allows one to predict below average performance in both Mathematics and German; the reason for this prediction is that the training provided by schools may be less efficient for children with below average Fluid Intelligence.

It may be worth noting that these hypotheses do not distinguish between girls and boys and between performance in Mathematics and performance in German. While interesting, hypotheses concerning these distinctions cannot be pursued in addition to the above four hypotheses using the present sample size. Cross-classifying three dichotomous variables creates a table with eight cells. This is not too many for a sample size of N = 93. Cross-classifying five variable creates a table with 32 cells. This creates a possibly sparse table.

Parameter estimates from sparse table can rarely be trusted. Therefore, we confine ourselves to the smaller tables.

2.2 Methods for Data Analysis

This section gives a brief description of the statistical methods used for analysis of the present data. First, we describe the correlational methods, then, prediction analysis. For the following analyses, variables were dichotomized at their medians with 1 = below median and 2 = above median.

2.2.1 Correlations Between Binary Variables

To estimate the correlations between the dichotomous variables we selected similarity coefficient S4 from the SYSTAT CORRelation module (1992). This coefficient is defined as follows:

$$S4 = \frac{a + d}{a + b + c + d}\ , \qquad (1)$$

where a, b, c, and d denote cell frequencies of the 2 x 2 cross-tabulation of two variables, row-wise, starting with the upper row. This coefficient estimates the proportion of pairs where the values of both values agree. It is a metric coefficient that creates symmetric, Gramian, that is, positive semi-definite matrices as long as one has complete data. It yields estimates numerically very close to the Pearson correlation coefficient (see Rovine & von Eye 1990, 1996).

2.2.2 Prediction Analysis of Cross-Classifications

To test the hypotheses put forth in section 2.1 we used prediction analysis (PA; cf., von Eye, Brandtstädter, & Rovine, 1993, in press). PA is a method for analysis of point predictions in cross-classifications of predictors and criteria. Point predictions link selected predictor states to selected criterion states. For example, consider the predictor, A with categories a1, a2, and a3, and the criterion, B, with categories b1, b2, and b3. It should be noticed that both categories a_i and b_j with i, j = {1, 2, 3}, can be categories of *composite variables* (see von Eye & Brandtstädter, 1988), that is, can result from combining categories from two or more predictors, or two or more criteria. For the variables, A and B, a sample set of predictions is as follows:

$$a1 \rightarrow b1,$$

$$a2 \rightarrow b2 \vee b3,$$

$$a3 \rightarrow b1 \vee b3.$$

In words, this set of partial predictions predicts the occurrence of b1 from a1, the occurrence of b2 OR b3 from a2, and the occurrence of b1 OR b3 from a3, where the capitalized words indicate logical operators. Sets of prediction hypotheses of this type create three types of cells in the cross-classification of predictors and criteria:

1. *Hit Cells:* these cells contain events (cases, subjects, respondents etc.) that meet a prediction; hit cells for the above example are a1b1, a2b2, a2b3, a3b1, and a3b3;
2. *Error Cells:* these cells contain cases that do not meet any of the predictions; error cells for the above example are a1b2, a1b3, a2b1, and a3b2;

3. *Irrelevant Cells* (von Eye & Brandtstädter, 1988): these cells contain cases not involved in predictions; reasons for the existence of such cases include that particular predictor levels are part of no prediction; the above example contains no irrelevant cells.

von Eye et al. (1993) proposed testing sets of prediction hypotheses using the following non-standard log-frequency model (Rindskopf, 1990):

$$\log m_{ij} = \lambda + \sum_i \lambda_i^A + \sum_j \lambda_j^B + \sum_k \lambda_k^X \ . \qquad (2)$$

The λ parameters for this model are calculated for three groups of vectors (see Tables 3 and 4). The first group, indexed with i, contains vectors for all possible main effects and interactions among predictors. The second, indexed with j, contains all possible main effects and interactions among criterion variables. Thus, this model is *saturated in both predictors and criteria*. However, thus far, it assumes independence between predictors and criteria.

The third group of vectors, indexed with k, contains the *unmeasured variables* that describe the links between predictors and criteria. These variables reflect the prediction hypotheses. If predictors and criteria are independent of each other, the third group of vectors is not necessary for a satisfactory rendering of the data.

Contrasts for the vectors indexed with i and j can be specified according to the usual specifications for effect coding. The contrast vectors indexed with k are formulated according to the following specifications:

$$\lambda_k^X = \frac{\gamma_i}{h_i} \ , \qquad (3)$$

if cell ij is one of the h hit cells, and

$$\lambda_k^X = -\frac{\gamma_i}{e_i} \ , \qquad (4)$$

if cell ij is one of the e error cells. γ_i is defined as in effect coding. Using these specifications we obtain, instead of (2), the following model:

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B + \gamma_i \left(\frac{\delta_{ij}}{h_i} - \frac{1-\delta_{ij}}{e_i} \right) \qquad (5)$$

where $\delta_{ij} = 1$ if cell ij is a hit cell and $\delta_{ij} = 0$ else.

This model implies that in each row, the deviations from independence are the same for all hit cells; in each row, the deviations from independence are the same for all error cells; these deviations can be different for hit and error cells. If the model fits, all statistically important deviations from independence are captured.

Comparing PA and LCA. To compare LCA and PA we first express the latent class model (for two manifest variables, A and B, and one latent variable, X) in log-linear model terms:

$$\log m_{ijk} = \lambda + \sum_i \lambda_i^A + \sum_j \lambda_j^B + \sum_k \lambda_k^X + \sum_{ik} \lambda_{ik}^{AX} + \sum_{jk} \lambda_{jk}^{BX} \ . \qquad (6)$$

Obviously, the PA model presented in (2) and the LCA model presented in (6) overlap in the first four terms on the right hand side of the equation. The terms that capture the interaction between the latent and the manifest variables are not explicitly part of the PA model (2). However, there is a basic difference between the two models [that also prevents the interaction terms in (6) from being subsumable in the fourth term in (2)]. This difference is that the vectors in the fourth term in (2) restrict the interaction terms λ_{ij} for the interaction between A and B to a smaller set of parameters. Thus, they parameterize the interaction between the manifest variables. In contrast, Variable X in (6) can be thought of as an additional, *latent* variable or dimension of the cross-classification under study. This variable is used to explain the interaction between the manifest variables (for details see Rost & Langeheine, chapter 1). Thus, terms that represent this interaction do not need to be included in the LCA model. (The relation of this basic difference between the PA and the LCA models will be taken up again further in the discussion.)

3. Results

Table 2 displays the correlations among the four variables under study in the data example.

Results suggest that the correlations among the dichotomous variables are high throughout. Differences between correlations are not statistically significant. Therefore, one cannot justify the statement that the correlations between Crystallized Intelligence and school performance indicators are higher than the correlations between Fluid Intelligence and school performance indicators (see discussion of Hypothesis 3, above). Particularly high is the correlation between the two subjects, German and Mathematics.

Variables	*Variables*		
	C	F	G
F	0.699		
G	0.731	0.667	
M	0.774	0.688	0.806

Table 2: Intercorrelations among the variables Fluid Intelligence, F, Crystallized Intelligence, C, school performance in German, G, and in Mathematics, M

The following *results from PA* complement the correlational results in that they provide finer grade specifications of the links between intelligence measures and the school performance indicators. Table 3 displays PA results for the prediction from Crystallized Intelligence (C) to school performance in German (G) and Mathematics (M).

Cell Index	*Frequencies*		*Design Matrix*					
CGM			*Main Effects*			*Interaction*	*PA Hypotheses*	
	Obs.	Exp.	C	G	M	G x M	I	II
111	21	21.00	1	1	1	1	0	1
112	2	1.44	1	1	-1	-1	0	-1/3
121	11	11.56	1	-1	1	-1	0	-1/3
122	13	13.00	1	-1	-1	1	0	-1/3
211	1	1.00	-1	1	1	1	-1/3	0
212	0	0.56	-1	1	-1	-1	-1/3	0
221	5	4.44	-1	-1	1	-1	-1/3	0
222	40	40.00	-1	-1	-1	1	1	0

Table 3: Observed and expected cell frequencies and design matrix for PA of the cross-classification of C, G, and M

Overall fit for the prediction model given in the design matrix in Table 3 is excellent (log-likelihood $G^2 = 1.396$; df = 1; $p = 0.2374$ (all p-values given here and in the following paragraphs refer to one-sided tests)). Both parameter estimates for the prediction hypotheses are statistically significant. The test statistic for the parameter for the first of the above hypotheses is $z = 1.815$ ($p = 0.0348$), the test statistic for the second parameter is $z = 3.378$ ($p = 0.004$). The model of independence between C on the one hand side, and M and G on the other, that is, the model without the two PA hypothesis parameters, is not tenable ($G^2 = 41.852$; df = 3; $p < 0.01$).

Table 4 displays PA results for the prediction from Fluid Intelligence to performance in German and Mathematics. Overall fit for the prediction model given in the design matrix in Table 4 is excellent ($G^2 = 0.028$; df = 1; $p = 0.8671$). However, only one of the parameters for the Prediction Hypotheses 3 and 4 is statistically significant. Specifically, the test statistics are $z = 2.076$ ($p = 0.0189$) for the third hypothesis, and $z = 1.597$ ($p = 0.0551$) for the fourth hypothesis. Thus, these results suggest that only the vector for Hypothesis 3 in the design matrix allows one to capture statistically significant variation that, otherwise, would lead to rejection of the model of independence among the predictor, F, and the criteria, G and M ($G^2 = 18.763$; df = 3; $p < 0.01$).

Cell Index	*Frequencies*		*Design Matrix*					
FGM			*Main Effects*			*Interaction*	*PA Hypotheses*	
	Obs.	Exp.	F	G	M	G x M	III	IV
111	19	19.00	1	1	1	1	0	1
112	1	1.11	1	1	-1	-1	0	-1/3
121	9	8.89	1	-1	1	-1	0	-1/3
122	18	18.00	1	-1	-1	1	0	-1/3
211	3	3.00	-1	1	1	1	-1/3	0
212	1	0.89	-1	1	-1	-1	-1/3	0
221	7	7.11	-1	-1	1	-1	-1/3	0
222	35	35.00	-1	-1	-1	1	1	0

Table 4: Observed and estimated expected cell frequencies and design matrix for PA of the cross-classification of F, G, and M

4. Discussion

The following discussion focuses on two issues. The first concerns the substantive results presented in this chapter. The second issue concerns the role played by and the characteristics of Prediction Analysis in comparison to standard latent class models.

4.1 Predicting Performance in School

The present results suggest a very strong correlation between the predictors, Fluid and Crystallized Intelligence, and the two criteria, school performance in German and Mathematics. In addition, the predictor intercorrelations and the criteria intercorrelations are high (see Table 2). Prediction Analysis, however, provides a more differentiated picture. The support for the theory-derived hypotheses suggests that Fluid and Crystallized Intelligence differ in their predictive characteristics. Specifically, above average Crystallized Intelligence allows one to predict above average performance in both German and Mathematics. Below average Crystallized Intelligence allows one to predict below average performance in both subjects. For Fluid Intelligence, only the second of these predictions is supported. Thus, we conclude that, as hypothesized, Fluid Intelligence relates to behaviors not typically shaped in school.

4.2 The Role and Characteristics of PA

PA is a method for custom-tailored testing of hypotheses that concern the relationship between predictors and criteria. The variant of PA proposed by von Eye et al. (1993) involves specifying non-standard log-linear models. These models include vectors for unmeasured covariates that express the prediction hypotheses. There are no vectors for the interaction between unmeasured, latent variables and the manifest variables.

The LCA and the PA models attempt at parameterizing deviations from some model of independence. The PA model is saturated in predictors and criteria. In other words, this model can be contradicted only if there are predictor-criteria relationships. Using PA, researchers model these relationships.

In contrast to what one might assume, there is no generally agreed-upon definition of the term *latent* in latent class models. This applies to the nature of latent variables as categorical versus continuous (Molenaar & von Eye, 1994) but also to the characteristic of variables as latent, that is, unmeasured. Whereas in most works (e.g., Clogg & Shockey, 1988) the definition of *latent* implies that variables are neither measured nor measurable, there are discussions of latent variable concepts for measurable variables (Sobel, 1994). PA typically operates with measured variables. However, the concepts carried by partial predictions (see section 2.2.2 of this chapter) often are *operationalizations of statements derived from latent concepts that cannot be measured*. Examples of such latent concepts include *equi-finality* and *equi-causality* as investigated in developmental research (von Eye et al., in press). Thus, PA can be viewed as a method that allows one to analyze at the manifest variable level predictions derived from latent variables that cannot be measured.

Acknowledgments

Christiane Spiel's work on this study was supported in part by the Austrian „Fonds zur Förderung der Wissenschaftlichen Forschung“ (No: P7630 - S0Z) and by the Johann Jacob Foundation. Alexander von Eye's work on this study was supported in part by a grant for the SEEK Project from the Michigan State University and a grant from the Mary Louis foundation. This support is gratefully acknowledged. The authors are indebted to Ferdinand Keller and Thorsten Meiser for constructive comments on an earlier version of this chapter.

References

Bentler, P.M. (1992). *EQS. Structural Equations Program Manual*. Los Angeles, CA: BMDP.

Cattell, R.B. (1971). *Abilities: Their structure, growth and action*. New York: World Book.

Clogg, C.C., & Shockey, J.W. (1988). Multivariate analysis of discrete data. In J.R. Nesselroade & R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology*, 2nd ed. (pp. 337 - 365). New York: Plenum.

Horn, J. (1988). Thinking about human abilities. In J.R. Nesselroade & R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed.)(pp. 645 - 685). New York: Plenum.

Kubinger, K.D., & Wurst, E. (1988). *Adaptives Intelligenz Diagnostikum, Manual* [Adaptive Intelligence Battery]. Weinheim: Beltz.

Kühn, R. (1984). Vorhersagbarkeit von Schulnoten mit Hilfe zweier Intelligenztests. [Can one predict grades in school from intelligence test measures*?] Zeitschrift für erziehungs- und sozialwissenschaftliche Forschung, 1*, 169 - 180.

Molenaar, P.C.M., & von Eye, A. (1994). On the arbitrary nature of latent variables. In A. von Eye, & C.C. Clogg (Eds.), *Latent variables analysis. Applications for developmental research* (pp. 226 - 242). Newbury Park: Sage.

Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant* [The evolving intelligence in children]. Neuchâtel: Delachaux & Niestlé.

Rindskopf, D. (1990). Nonstandard log-linear models. *Psychological Bulletin, 108*, 150-162.

Rovine, M. J., & von Eye, A. (1990). A reinterpretation of the correlation coefficient: Correlation as a count of matches. In K. Berk & L. Malone (Eds.*), Computer science and statistics: Proceedings of the 21st Symposium on the Interface* (pp. 287-291). Alexandria, VA: The American Statistical Association.

Rovine, M.J., & von Eye, A. (1996) Correlation and categorization under a matching hypothesis. In A. von Eye & C.C. Clogg (Eds.), *Analysis of categorical variables in developmental research*. Boston: Academic Press.

Sirsch, U., & Spiel, C. (1994). Die Bedeutung formal-operatorischen Denkens für die Schulleistung [The significance of formal-operatoric thinking for performance in school]. In R. Olechowski & B. Rollett (Eds.), *Theorie und Praxis. Aspekte empirisch-pädagogischer Forschung - quantitative und qualitative Methoden* (pp. 111 - 117). Frankfurt/M.: Peter Lang.

Sobel, M.E. (1994). Causal inference in latent variable models. In A. von Eye, & C.C. Clogg (Eds.), *Latent variables analysis. Applications for developmental research* (pp. 3 - 35). Newbury Park: Sage.

Spiel, C. (1995). *Erfolg in der Schule ein dynamisches Modell der Effekte biologischer und sozialer Risiken auf Schulleistungen* [Academic success - a dynamic model of effects that biological and social risks have on performance in school]. University of Vienna: unpublished 'Habilitationsschrift.'

Spiel, C. (in press, a). Long-term effects of minor biological and psychosocial risks on cognitive competence, school achievement, and personality traits. In S. Harel & J.P. Shonkoff (Eds.*), Early childhood intervention and family support programs: Accomplishments and challenges*. Brookes.

Spiel, C. (in press, b). Methods for re-analysis of data from longitudinal studies. In R.K. Silbereisen and A. von Eye (Eds.), *Growing up in times of social change*. New York: Springer.

SYSTAT *for Windows: Statistics, Version 5 Edition* (1992). Evanston, IL: SYSTAT Inc.

Tramontana, M.G., Hooper, S.R., & Selzer, S.C. (1988). Research on the preschool prediction of later academic achievement: A review. *Developmental Review, 8,* 89 - 146.

von Eye, A., & Brandtstädter, J. (1988). Formulating and testing developmental hypotheses using statement calculus and non-parametric statistics. In P. B. Baltes, D. Featherman, & R. M. Lerner (Eds.), *Life-span development and behavior* (Vol. 8, pp. 61-97). Hillsdale, NJ: Erlbaum.

von Eye, A., Brandtstädter, J., & Rovine, M. J. (1993). Models for prediction analysis. *Journal of Mathematical Sociology, 18*, 65-80.

von Eye, A., Brandtstädter, J., & Rovine, M. J. (1996). Models for prediction analysis in developmental research. *Journal of Mathematical Sociology* (in press).

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children - Revised*. New York: The Psychology Corporation.

zur Oeveste, H. (1987). *Kognitive Entwicklung im Vor- und Grundschulalter* [cognitive development in preschoolers and grade schoolers]. Göttingen: Hogrefe.