



Universiteit
Leiden
The Netherlands

Crossvalidation of the WISC-R factorial structure using three-mode principal components analysis and perfect congruence analysis

Kroonenberg, P.M.; Berge, Jos M.F. ten

Citation

Kroonenberg, P. M., & Berge, J. M. F. ten. (1987). Crossvalidation of the WISC-R factorial structure using three-mode principal components analysis and perfect congruence analysis. *Applied Psychological Measurement*, *11*, 195-210. Retrieved from <https://hdl.handle.net/1887/11643>

Version: Not Applicable (or Unknown)
License: [Leiden University Non-exclusive license](#)
Downloaded from: <https://hdl.handle.net/1887/11643>

Note: To cite this publication please use the final published version (if applicable).

Cross-Validation of the WISC-R Factorial Structure Using Three-mode Principal Components Analysis and Perfect Congruence Analysis

Pieter M. Kroonenberg
University of Leiden

Jos M. F. ten Berge
University of Groningen

By using three-mode principal components analysis and perfect congruence analysis in conjunction, the factorial structure of the 11 correlation matrices of the Wechsler Intelligence Scale for Children-Revised was analyzed within a single framework. This allows a unified description showing both the strong similarities between the standardization samples and some small differences related to age. Furthermore, claims about comparability between the WISC-R factorial structure, the structures of other independently conducted studies, and those of several translations of the WISC-R were evaluated. Again the overall similarity was striking, albeit most studies showed lower explained variances. Some age effects seemed to be present here as well. The contribution of three-mode principal components analysis was found to lie primarily in the simultaneous analysis of the standardization samples, while perfect congruence analysis allowed the evaluation of the strengths and the correlations of the common WISC-R components in all studies without encountering rotation problems.

Since its publication the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) has been widely applied in practice, and has been an object of extensive research (for reviews see Kaufman, 1979, 1981; Quattricchi & Sherrets, 1980). One topic in these studies has been the factorial structure of the test. Kaufman (1975) investigated the 11 correlation matrices from the stan-

dardization samples (age groups from 6½ to 16½) and derived a median factor-loading matrix for these samples. In a considerable number of studies, the WISC-R factorial structure was investigated for groups of children who deviate in various ways from the standardization samples, and their solutions were compared mostly with Kaufman's median factor loadings (Kaufman, 1975, Table 4). Establishment of a common factor space, for both the Wechsler samples and for the similarity of the factor structure of other studies with a standard solution, is important in light of claims that the WISC-R can be used for other different groups, and that translations of the WISC-R are essentially identical to the original test after translation.

One aim of the present paper was to investigate the extent to which the standardization samples can be considered to have one single underlying factor space by (1) performing an integrated analysis of all samples at the same time, and (2) comparing them separately to the components of a weighted average correlation matrix. A second aim was to investigate the extent to which studies conducted in the field with the WISC-R resemble the standardization data with respect to the factorial structure of the subtests. A third aim was to show how perfect congruence analysis for weights (PCW; ten Berge, 1986b) and a special type of three-mode principal components analysis (Kroonenberg, 1983a; Tucker, 1972) can be used, both in conjunction and separately, to study the problems exhibited by such

comparisons. Previous studies generally used visual inspection or congruence coefficients to study similarities between structures, but no single integrated analysis of the Wechsler data exists, nor have attempts been made to evaluate the results of other studies, and those of translations of the WISC-R within an integrated context.

Several interesting relations exist between the major techniques used here, and a first investigation into them was presented by Kroonenberg and ten Berge (1985), but these will not be discussed here. Other techniques serving similar purposes might be considered, such as factor matching techniques (for an overview, see ten Berge, 1977) or simultaneous factor analysis (Jöreskog, 1971). These techniques will not be used here, but will be briefly compared with the approaches presented below.

The techniques applied and outlined here have applicability far beyond the present data. In general, when components from correlation (or covariance) matrices are to be compared, both perfect congruence analysis and three-mode principal components analysis can perform useful functions. Other possible applications are developmental studies and cross-sectional studies of the same variables over time. Three-mode principal components analysis (or factor analysis) has also been shown to be useful with raw data (Kroonenberg, 1983b).

Method

Data

The primary dataset (referred to as the *Wechsler set*) consisted of the 11 correlation matrices contained in Wechsler (1974, pp. 36-46), one for each of the standardization subsamples of the WISC-R. To expand the database, letters were sent to authors of papers which presented factor analyses of the WISC-R but did not include correlation matrices when published. In particular, volumes of such journals as *Journal of Consulting and Clinical Psychology*, *Psychology in the Schools*, and *Psychological Abstracts* were searched for papers on the factorial structure of the WISC-R; 57 papers were found. Authors were approached with a request to make available the original matrices upon which their

analyses were based. Some authors supplied the desired information; others only sent reprints of their papers or indicated that the correlation matrices were no longer available, while most did not respond at all.

In all, 32 correlation matrices were available for reanalysis, several of which could not be used because only 10 of the 12 subtests were administered. The remaining correlation matrices can be divided into two groups, those with and those without *Mazes*. Here only 11 subtests were analyzed (excluding *Mazes*) to keep as many studies as possible in the investigation (Wechsler indicated that *Mazes* is an optional subtest). Table 1 gives the summary information and references for the matrices used in this study. The 12 matrices from the nine American studies will be referred to as *Other* correlation matrices from *Other* studies to distinguish them from the Wechsler set, and the 15 matrices from the three translated tests as the *Cross-cultural* matrices from *Cross-cultural* studies.

Nature of the Data: Age-Scaled Scores

Generally, sets of covariance matrices rather than sets of correlation matrices are analyzed. The major rationale put forward for analyzing covariance matrices (e.g., Harshman & Lundy, 1984, p. 141) is that if more than one subpopulation is analyzed using a single set of variables, the unit of measurement for each variable should be the same in each subpopulation; that is, standardization should be performed across populations (see also Meredith, 1964). Correlation matrices from several populations, however, are based on standardizations within subpopulations.

The present paper originated in current usage of the WISC-R, and thus displays some of the defects of this usage. In particular, it is common practice in most field studies to have a mix of children of varying ages, as can be seen from column 6 in Table 1. Because of the age mix, the use of age-scaled scores is unavoidable, in contrast to age-homogeneous groups where raw scores are available. As standard deviations from age-scaled scores

Table 1
 Description of Correlation Matrices

Authors	Description	p	m	n	Age	IQ ^a	pub ^c	Deleted Subtest ^b
Wechsler (1974)	standardization samples	11	12	200	6½-16½	100	+	-
Chan (1984)	Hong Kong standardization samples	11	12	100	5-15	100	+	-
Cummins & Das (1980)	educable mentally retarded	1	11	95	13-15	73	-	M
Groff & Hubble (1982a)	'average ability' sample	1	12	72	8-10	(100)	-	-
Groff & Hubble (1982b)	younger retarded	1	11	63	9-11	70	-	M
Groff & Hubble (1982b)	older retarded	1	11	78	14-16	67	-	M
Hofman & Pijl (1983)	standardization-like samples	1	12	231	6-7	103	+	-
Petersen & Hart (1979)	'no significant problems'	1	12	238	8-9			
Pijl (1982)	Entire standardization sample Dutch WISC-R	1	11	248	7-12	-	-	M
Reschly (1978)	Anglo-Americans	1	12	2013	6½-7½	100	+	-
"	Blacks	1	12	252	6-16	-	-	-
"	Native American Papago	1	12	235	6-16	-	-	-
"	Mexican-American	1	12	240	6-16	-	-	-
Sandoval (1982)	Mexican-American	1	12	223	6-16	-	-	-
Stedman et al. (1978)	'school-related problems'; 90% Spanish surname	1	12	307	5-11	-	-	-
Titze & Tewes (1984)	Average from standardization samples German WISC-R (HAWIK)	1	11	106	6-13	50	+	-
Van Hagen & Kaufman (1975)	retarded	1	11	1898	6½-16½	100	+	M

Notes: p=number of correlation matrices; m=number of subtests; n=number of children for each correlation matrix;
^a average IQ; - not available;
^b M Mazes;
^c published: + yes, - no.

have no conceptual interpretation, an analysis of correlations of subtests, rather than covariances, seems to be appropriate for the Other studies and the Cross-cultural studies.

In principle, the WISC-R covariance matrices could be analyzed, but because the Wechsler set serves below as a basis for comparison for the correlation

matrices from mixed-age groups, the Wechsler correlation matrices were used. Note, however, that for the Wechsler set, analyzing covariance or correlation matrices does not lead to essentially different solutions as long as the published matrices are used, because for each subgroup the raw scores of all subtests have been rescaled to a standard

deviation of 3 and a mean of 10. As the age-scaled scores are nonlinear (monotone) transformations of the raw scores, it is conceivable that the results for the raw scores would not be the same as those for the age-scaled scores. However, the relationship between analyses from age-scaled and raw scores is a different problem from the one addressed here, and its study would require the original data of the 2,200 "standardization children."

Another question not addressed in the present paper is the appropriateness of the conceptual models for the present data implied by the three-mode model, and the perfect congruence approach. In particular, it is not unlikely that age-scaling, even though appropriate for comparing IQ, might lead to distortions in the structure underlying the subtests both within and across ages, and that combining age-scaled scores of different age groups might add further distortions. It is clear that a thorough treatment of the problem of age-scaled scores requires both a full-fledged theoretical analysis and an empirical analysis of a large database, preferably the standardization set itself. The present paper has a narrower focus, and takes the limitations of current practice as its starting point.

Three-mode Principal Components Analysis

Three-mode principal components analysis is a generalization of regular principal components analysis to a situation in which measurements have been collected under several conditions, for more than one point in time, or for several groups. Here, a short description is given of the analysis of several correlation matrices based on scores of the same m variables of different groups or of the same group under several conditions. Thus there is a set of n correlation matrices \mathbf{R}_k ($k = 1, \dots, n$), and it is desired to find one single set of component loadings for all groups simultaneously. At the same time, it is desired to assess the strength (or explained sum of squares, variance) of the components and their correlations in each of the various groups.

The correlation matrices are modeled with the so-called TUCKER2 model (Kroonenberg & de

Leeuw, 1980; Tucker, 1972) as

$$\mathbf{R}_k = \hat{\mathbf{R}}_k + \mathbf{E}_k = \mathbf{G}\mathbf{C}_k\mathbf{G}' + \mathbf{E}_k, \quad (1)$$

where $\hat{\mathbf{R}}_k = \mathbf{G}\mathbf{C}_k\mathbf{G}'$ is the estimated correlation matrix based on the model with, say, s components, $s \leq m$,

\mathbf{E}_k is the matrix of residual correlations, and

\mathbf{C}_k is a symmetric matrix, which may be called the "group characteristic matrix" (Tucker, 1972, p. 6).

\mathbf{G} is the component space common to all subgroups and can be seen as a compromise solution for the individual component spaces of the separate samples. As Kroonenberg and de Leeuw (1980) showed, after \mathbf{G} has been determined by an alternating least squares algorithm, \mathbf{C}_k may be computed as

$$\mathbf{C}_k = \mathbf{G}'\mathbf{R}_k\mathbf{G}, \quad (2)$$

which shows that \mathbf{C}_k is a Gramian matrix. Under the assumption of positiveness of the diagonal elements of \mathbf{C}_k , \mathbf{C}_k may be decomposed as

$$\mathbf{C}_k = (\mathbf{D}_k)^{1/2} \Phi_k (\mathbf{D}_k)^{1/2}, \quad (3)$$

where $\mathbf{D}_k = \text{Diag}(\mathbf{C}_k)$, and Φ_k has elements ϕ_{pqk} such that $\phi_{pqk} = c_{pqk}/(c_{ppk})^{1/2}(c_{qqk})^{1/2}$, which implies that $\phi_{ppk} = 1$ if $p = q$ ($p, q = 1, \dots, s$). Substituting Equation 3 into Equation 1 gives

$$\mathbf{R}_k = (\mathbf{G}\mathbf{D}_k^{1/2})\Phi_k(\mathbf{G}\mathbf{D}_k^{1/2})' + \mathbf{E}_k, \quad (4)$$

so that the diagonal elements of \mathbf{D}_k represent the explained variance or fit of \mathbf{G} in the k th group, if \mathbf{G} is taken to be orthonormal, and the ϕ_{pqk} ($p \neq q$) represent the correlation of the components (scores) on the p th and q th component in group k . The diagonal elements $\hat{\mathbf{R}}_k$ are the estimated communalities on the basis of the model, and the off-diagonal elements of the symmetric matrix \mathbf{E}_k indicate the "residual correlations," that is, they indicate how well the original correlations have been reproduced by the model.

The Wechsler set was subjected to an analysis using Equation 1 with $s = 3$ and $s = 4$ components for the subtests, respectively. Previous examples of such analyses can be found in Kroonenberg (1983a, chap. 11; 1983c). The three-mode analyses were performed using the TUCKALS2 program¹

¹The TUCKALS2 program may be obtained from the first author.

(Kroonenberg & Brouwer, 1985). A theoretical discussion and comparison of the methods used in this paper to analyze correlation matrices simultaneously was presented by Kroonenberg and ten Berge (1985).

Perfect Congruence Analysis for Component Weights

Perfect congruence analysis for weights (PCW; ten Berge, 1986a, 1986b) is essentially a method for cross-validating component weights. The method can also be interpreted as a generalization of the Multiple Group Method of factor analysis (see, e.g., Gorsuch, 1983, pp. 81-89; Nunnally, 1978, pp. 398-400). A parallel procedure may be derived for component loadings; however, ten Berge (1986b) has shown PCW to be superior.

The procedure is based on the fact that components are linear combinations of variables, and are defined by the weights in those linear combinations. The weights are derived from an initial study, and these weights are then used in a second study to determine the values of the persons on the components in the second study, parallel to the cross-validation procedure in ordinary regression. Because (except for scaling constants) the same weights are used, the weights in the two studies as measured by Burt's (1948) and Tucker's (1951) congruence coefficient are perfectly congruent.

As every component is uniquely defined by its component weights, the weights can be taken to define the interpretation of the component, as has been argued by Harris (1985, pp. 317-320). It follows that any component from a first (previous) study can be recovered in a second (new) study where the same variables have been used. Component weights from the first study can simply be applied to the variables in the second one to define new components with the same interpretation. Components with the same interpretation may behave differently across studies (populations) in several respects:

1. The components may differ in terms of the amount of variance they explain. That is, components may have different sums of squared

correlations with the variables across populations. Such differences are well-known in practice and can be reported as interesting in their own right.

2. The components may correlate differently with other components in the same study. This is also familiar from applied studies. For instance, in selected populations of gifted students, verbal ability and numerical ability tend to correlate relatively low. Also, components constrained to be orthogonal in a first study typically do correlate in other studies, if the orthogonality constraint is dropped.
3. The reliability of a component, and its validity with respect to external criteria, may differ across studies.

More formally, PCW can be explained as follows.² Let \mathbf{R}_2 denote a $m \times m$ correlation matrix of k variables in a second study, and let \mathbf{B}_1 be a $m \times s$ matrix of component weights or component-score coefficients, defining s components of the same variables in the first study. The $m \times s$ matrix \mathbf{B}_2 of component weights, defining the same components in the second study, can be obtained as

$$\mathbf{B}_2 = \mathbf{B}_1 [\text{Diag}(\mathbf{B}_1' \mathbf{R}_2 \mathbf{B}_1)]^{-1/2} \quad (5)$$

This normalization of the columns of \mathbf{B}_1 guarantees that \mathbf{B}_2 defines standardized components in the second study. The normalization does not affect the "behavior" of these components, but merely serves to simplify the presentation. It should be noted that the weight matrices \mathbf{B}_1 and \mathbf{B}_2 are perfectly congruent (proportional) columnwise, hence the name of this analysis procedure.

From \mathbf{R}_2 and \mathbf{B}_2 it is easy to compute the $m \times s$ structure matrix \mathbf{S}_2 , with correlations between the variables and the components in the second study, as $\mathbf{S}_2 = \mathbf{R}_2 \mathbf{B}_2$, and the $s \times s$ component correlation matrix $\Phi_2 = \mathbf{S}_2' \mathbf{B}_2 = \mathbf{B}_2' \mathbf{R}_2 \mathbf{B}_2$.

The j th column sum of squares of \mathbf{S}_2 conveys how much variance is explained in the second study by a single component which has the same interpretation as component j from the first study. Thus

²A PASCAL program for PCW may be obtained from the second author.

the variances follow from $\text{Diag}(\mathbf{S}'_2\mathbf{S}_2) = \text{Diag}(\mathbf{B}'_2\mathbf{R}'_2\mathbf{B}_2)$. The sum of the explained variances, $\text{tr}(\mathbf{S}'_2\mathbf{S}_2)$, is a meaningless quantity unless the components are orthogonal in the second study (i.e., unless Φ_2 is \mathbf{I}). Instead, the amount of variance explained by the s components jointly must be computed as $\text{tr}(\mathbf{S}'_2\mathbf{S}_2\Phi_2^{-1})$. In addition, the off-diagonal elements of Φ_2 can be inspected to assess the correlations between the components in the second study.

A considerable advantage of PCW over other transformation (rotation) techniques is that arbitrary decisions no longer need be made about the number of components to be retained before transformation, about the "appropriate" transformation technique, about the amount of congruence that is necessary for components to be called the same, etc. (For further discussion of these points see ten Berge, 1986b.)

In many applied studies \mathbf{B}_1 is not known or has not been reported. Typically, only component loadings are reported. However, this need not be a problem. If the components from a first study are truncated and/or rotated principal components with $k \times r$ pattern matrix \mathbf{P}_1 , then \mathbf{B}_1 can be explicitly obtained as

$$\mathbf{B}_1 = \mathbf{P}_1(\mathbf{P}'_1\mathbf{P}_1)^{-1} \quad (6)$$

(see ten Berge, 1986a).

The starting point for all PCW analyses in this paper was the weight matrix $\hat{\mathbf{B}}$ derived from the varimax-transformed loadings $\hat{\mathbf{G}}$ of the three-mode principal component solution using Equation 6 with

$$\hat{\mathbf{G}} = \mathbf{G} \left[\text{Diag} \left(\sum_{k=1}^n \mathbf{C}_k/n \right) \right]^{1/2} \mathbf{T} \quad (7)$$

where \mathbf{T} is the varimax transformation matrix. In the present study these loadings were nearly identical to the loadings obtained from the average correlation matrix (Wechsler, 1974, p. 47), but in general this need not be the case. When correlation matrices are less homogeneous than in the Wechsler set, considerable differences may occur. In case of nonhomogeneity, a three-mode analysis seems preferable because averaging may lead to cancellation by opposite signs or to a general leveling off of the correlations.

Three-mode Principal Components Analysis Versus PCW

In this paper three-mode principal components analysis and perfect congruence analysis are used as complementary techniques. The former is used to derive a common space for the variables of the Wechsler set, estimated communalities, and residual correlations. On the other hand, PCW is used to derive the explained variance of the common components and their correlations in the standardization samples of the Wechsler set, the Other studies, and the Cross-cultural studies. The techniques can be said to address slightly different questions. Three-mode analysis, as used here, aims to provide a comprehensive description of a set of correlation matrices, and perfect congruence analysis aims to provide information about how samples conform to known components. This explains why no three-mode analysis was performed over all available data. The principal aim of this paper was to evaluate claims of researchers that their solutions resemble those from the Wechsler set. In an overall analysis the characteristics of the other studies would also contribute toward the solution, causing it to diverge from that of the Wechsler set.

Comparison with LISREL

Several other techniques may be used to analyze sets of covariance (or correlation) matrices. Most prominent among these is simultaneous factor analysis (Jöreskog, 1971), which may be carried out with programs such as LISREL (Jöreskog & Sörbom, 1981), EQS (Bentler, 1985), or COSAN (Fraser, 1980). There are several differences between PCW and simultaneous factor analysis (ten Berge, 1986b). First, the goodness-of-fit test for simultaneous factor analysis provides a basis for the conclusion that the null hypothesis is *not* true, while it is generally desired to prove that a certain model fits adequately. Thus the latter method attempts to show that the model cannot be rejected, which should not be too difficult with small samples. Second, simultaneous factor analysis is not really geared toward investigation of factors which seem to have disappeared, while PCW does provide this infor-

mation, that is, for every factor of the original sample the explained variance can be assessed in the new sample, irrespective of its size in this new sample. Furthermore, if the exact factors of a first study do not fit a second one, with LISREL-type procedures a complicated search is necessary to find out to what extent and in what respect two studies are similar and different.

Another difference exists between three-mode principal components analysis and simultaneous factor analysis. The former technique is primarily an exploratory method used here to derive a common component space for all correlation matrices simultaneously, while the latter is a confirmatory technique requiring a hypothesized component space from the start. It is possible to use programs like LISREL in an exploratory fashion (see, e.g., Kroonenberg & Lewis, 1982), and LISREL is often used in this way, but such usage is full of methodological and technical pitfalls, such as possible nonconvergence and convergence to imaginary solutions with negative variance estimates, especially in small samples (see Boomsma, 1983).

Results

In this section, first the Wechsler set is investigated, primarily (but not exclusively) using three-mode principal components analysis for the overall analysis, and perfect congruence analysis for a comparative study of the subsamples (Study I). After having established the common component space and the variability within the Wechsler set, these results are used to evaluate the similarities and differences of other American studies employing the Wechsler scale (Study II). Next, a similar analysis follows for three translations of the WISC-R (Study III). In the latter two studies, perfect congruence analysis is the main tool for investigation.

Study I: The WISC-R Component Space

As most of the Other studies used only 11 of the 12 subtests (see Table 1) by excluding Mazes, in this paper results will be reported only for 11 sub-

tests. Table 2 shows the varimax rotated loadings and component weights derived by three-mode principal components analysis from the 11 correlation matrices in the Wechsler set.

Not unexpectedly, the structure of the components agrees with previous analyses of these data; that is, there exists a loading pattern, which can be seen to refer to the constructs *Verbal Comprehension* (VC), *Perceptual Organization* (PO), and *Freedom from Distractability* (FD) for the three components respectively. Note that "factorially pure" components cannot be found if the components are restricted to remain orthogonal. (As most previous studies have not seriously investigated oblique factor patterns, neither does the present study, nor does it consider the real meaning of the third factor; see Kaufman, 1981, for references to this debate.) Previous authors have sometimes used Kaufman's (1979, Table 4) median loadings for comparison. The component spaces presented here, however, can better serve this purpose, as they are based on a direct analysis of all correlation matrices simultaneously rather than on medians from separate component (factor) analyses. Some of Kaufman's components were even derived from four-dimensional solutions (see below), and they were therefore not necessarily oriented in the optimal manner for combining them into one solution.

Deriving a common solution is, of course, only fruitful and valid if the common component space is shared by all subsamples to a considerable degree. In order to verify this, it is necessary to investigate how well the common solution fits the individual subsamples. Table 3 gives the explained variance of each component for each of the subsamples, as well as the correlations between the varimax components based on a perfect congruence analysis. It is clear from these tables that the common components fit all subsamples more or less equally well (range of overall variance explained approximately 62% to 68%; mean 65%), and that even the components themselves are equally present in all subsamples. This result is especially satisfactory, as Kaufman (1979) needed four-dimensional solutions for the 6½- and 16½-year-olds to find the common FD factor. It is interesting to note

Table 2
WISC-R Varimax-Rotated Component Loadings and Weights
(11 subtests)

Subtest	Loadings ^b			Weights		
	VC	PO	FD	1	2	3
Information	.751	.272	.236	.315	-.096	-.036
Similarities	.745	.336	.153	.313	-.042	-.116
Arithmetic	.559	.150	.512	.156	-.150	.283
Vocabulary	.811	.264	.217	.361	-.121	-.068
Comprehension	.758	.288	.076	.359	-.069	-.183
Digit Span	.294	.023	.743	-.034	-.175	.581
Picture Completion	.341	.677	.018	-.020	.345	-.164
Picture Arrangement	.297	.619	.128	-.059	.307	-.045
Block Design	.284	.730	.297	-.153	.367	.097
Object Assembly	.149	.824	.124	-.221	.498	-.032
Coding	.040	.254	.741	-.272	.069	.623
Variance Explained ^a	3.060	2.480	1.616			
Percentage of Total	27.8	22.5	14.7			
Sum		65.0				

^a Derived from the diagonal of the average 'group characteristic matrix' (see Method section).

^b VC= Verbal Comprehension; PO= Perceptual Organization; FD= Freedom from Distractability

that Kaiser's criterion of only accepting eigenvalues greater than 1 would be ill-advised for the present data, as the FD component would only have been included in four subsamples, although it is present in all of them (details not shown).

In addition to the overall similarity of the subsamples, some relationships with age may be distinguished. In particular, age has substantial correlations with "Explained variance of VC" (.54), "Total variance explained" (.51), and "Correlation between PO and FD" (-.69); see Table 3. One conclusion from these correlations and the actual values in Table 3 seems to be that with increasing age, VC gains in importance, and the PO,FD correlation decreases with age. Compared to the overall similarities, however, the age effect is a rather small one.

Analyzing the residual correlations (i.e., observed correlation minus fitted correlation derived

from the three-mode principal components analysis) shows that the mean absolute residual per subsample is approximately .08 (not taking the diagonals into account). The diagonals of the fitted correlation matrix can be taken as estimates of the communalities of the three components for each of the subsamples using the common components (see Table 4). Inspection of the individual residual correlations shows that there is a systematic overestimation for all subsamples of the correlation between Coding and Digit Span by the three component solutions. A stable fourth principal component for all subsamples can be found, which accounts for another 6.8% variance on the average in all subgroups (range 6.0%-7.5%). This component has substantial loadings only for Coding and Digit Span (and a somewhat smaller one for Arithmetic), correcting the overestimation of their correlation in the three-component solution.

Table 3
 Explained Variances and Correlations of the Wechsler Varimax Components,
 Verbal Comprehension (VC), Perceptual Organization (PO), and Freedom
 from Distractability (FD) in Standardization Samples

Age groups	Explained Variance				% ^a	Correlations		
	VC	PO	FD	Total		(VC,PO)	(VC,FD)	(PO,FD)
6½	2.5	2.6	2.0	6.9	63	-.06	-.02	.21
7½	2.9	2.5	1.6	6.9	63	-.01	.01	.05
8½	2.5	2.6	1.6	6.9	63	-.04	-.15	.08
9½	3.5	2.8	1.7	7.5	68	.10	.07	-.05
10½	2.9	2.4	1.5	7.0	64	.04	-.10	-.05
11½	3.4	2.6	1.6	7.4	67	.07	.00	-.08
12½	3.4	2.9	1.7	7.5	68	.12	-.01	.03
13½	3.3	2.3	1.8	7.5	68	-.07	.06	-.05
14½	3.5	2.5	1.5	7.4	67	.03	.04	-.02
15½	2.9	2.2	1.6	7.1	65	-.12	-.02	-.03
16½	3.2	2.3	1.8	7.2	65	-.04	.13	-.08
Average ^b	3.1	2.5	1.6	7.2	65	.00	.06	-.00
Pooled ^c	3.1	2.5	1.6	7.2	65	.00	.00	.00
Correlation with age	.54	-.46	-.21	.51		-.15	.49	-.69

Note: % = Percentage explained variance.

^b average = explained variance averaged over subsamples.

^c pooled = explained variance derived from average correlation matrix (Wechsler, 1974, p.47).

Study II: Importance of Wechsler Components in Other Studies

As discussed above, each component of the Wechsler set returns in each of the Other studies, and the crucial question is how strong the components are in the new study, and what the correlations are between the components in the Other studies. This information can most easily be obtained from perfect congruence analyses on each of the Other studies.

This information is given in Table 5, as well as some comparable information from the Wechsler set. From this table it may be deduced that the amount of explained variation, both overall and for the separate components, generally falls short of the Wechsler ranges. The Stedman, Lawlis, Cort-

ner, and Achterberg (1978) and Van Hagen and Kaufman (1975) studies are the most notable exceptions. It is difficult to say for certain why this is so, but it is striking that all Other studies use mixed-age groups, whereas the Wechsler standardization samples are all homogeneous with respect to age.

It is tempting to attribute the lower explained variances to shrinkage as in cross-validation using, for instance, regression techniques. Ten Berge (1986b) investigated the possible role of shrinkage for PCW, but found in simulation studies that shrinkage does not occur. In fact, the opposite effect (although it was almost negligible) was systematically observed. Also noteworthy are the results of Cross-cultural studies, which show hardly any shrinkage. These speculations could be eval-

Table 4
Communalities of WISC-R Subtests for Standardization Samples;
IN = Information; SI = Similarities; AR = Arithmetic;
VO = Vocabulary; CM = Comprehension; DS = Digit Span;
PC = Picture Completion; PA = Picture Arrangement;
BD = Block Design; OA = Object Analysis; CD = Coding
(Three-mode PCA; 11 Subtests)

Age Group	Subtests										
	Verbal						Performance				
	IN	SI	AR	VO	CM	DS	PC	PA	BD	OA	CD
6½	.62	.62	.56	.69	.58	.60	.53	.48	.74	.72	.66
7½	.65	.65	.56	.73	.62	.59	.55	.48	.69	.70	.58
8½	.61	.62	.52	.68	.60	.61	.56	.48	.71	.73	.68
9½	.73	.73	.63	.81	.69	.65	.62	.52	.74	.74	.59
10½	.66	.67	.56	.74	.65	.66	.57	.47	.66	.68	.65
11½	.72	.73	.61	.81	.70	.66	.60	.50	.70	.71	.61
12½	.72	.73	.61	.80	.70	.65	.61	.52	.74	.72	.64
13½	.75	.73	.66	.84	.70	.70	.57	.49	.70	.73	.62
14½	.76	.76	.62	.85	.73	.60	.59	.50	.70	.71	.54
15½	.69	.68	.59	.78	.66	.64	.55	.47	.67	.72	.61
16½	.70	.68	.64	.78	.65	.68	.57	.48	.69	.72	.58
Low	.62	.62	.52	.68	.58	.59	.53	.47	.67	.68	.54
Median	.70	.68	.61	.78	.66	.65	.57	.48	.70	.72	.61
High	.76	.76	.66	.85	.73	.70	.62	.52	.74	.74	.68
Range	.15	.14	.14	.17	.15	.11	.09	.05	.07	.06	.14

uated by reanalyzing the original standardization data and creating mixed-age correlation matrices to compare their results with those of the Other studies. As mentioned above, such issues as the effect of age-scaled scores could be addressed at the same time.

Visual inspection shows no obvious trends between either explained variance or component correlations and sample characteristics on which the Other studies could be compared, such as ethnicity, retardedness, or number of children in the study (see also Table 1). There is, however, some indication that VC might not be at its full strength for younger children. This effect is present both in the Wechsler set (Table 3) and in the Other studies using younger children (Groff & Hubble, 1982a; Petersen & Hart, 1979), indicating that, relatively speaking, VC is not as clearly defined for younger children as for older ones.

Another point to notice is that even though the order of the components in the Wechsler set, both overall and individually, is very clearly VC, PO, and FD, in several of the Other studies (Cummins & Das, 1980; Groff & Hubble, 1982b—younger retarded; Van Hagen & Kaufman, 1975) there is no marked difference between VC and PO, while in Reschly (1978—Blacks), PO and FD were of more or less equal importance. Furthermore, in nearly all Other studies there is a (higher) negative correlation between VC and PO than in the Wechsler set, while all PO,FD and all but four VC,PO correlations fall within the ranges of the Wechsler set. In this respect, all three samples of Groff and Hubble (1982a, 1982b) stand out by their high component correlations: They belong both to the four samples with the highest VC,PO correlations and to the four with the highest VC,FD correlations.

Negative correlations between intelligence com-

Table 5
 Explained Variance and Correlations of the Wechsler Components,
 Verbal Comprehension (VC), Perceptual Organization (PO), and
 Freedom from Distractability (FD) in Other Studies

Other Study	Explained Variance					Correlations		
	VC	PO	FD	Total	% ^a	(VC,PO)	(VC,FD)	(PO,FD)
Cummins & Das	2.1	2.2	1.3	6.2	56	-.44	-.11	-.15
Groff & Hubble								
- Average Ability	2.2	1.7	1.3	5.8	52	-.46	-.23	-.07
- Younger Retarded	1.6	1.7	1.4	5.3	48	-.36	-.23	-.09
- Older Retarded	2.6	1.9	1.4	6.2	56	-.46	-.43	.02
Petersen & Hart	2.2	1.8	1.3	5.9	54	-.23	-.36	.02
Reschly								
- Anglo-American	2.6	1.8	1.4	6.5	59	-.23	-.02	-.07
- Black Americans	3.0	1.8	2.0	6.7	61	-.21	.10	.16
- Native Papago	2.4	1.8	1.4	6.2	56	-.28	-.12	.08
- Mexican Americans	2.5	1.8	1.4	6.3	57	-.21	-.07	.00
Sandoval	3.0	1.9	1.8	6.7	61	-.12	.06	.09
Stedman et al.	3.1	2.0	1.5	7.0	64	-.11	-.08	.06
Van Hagen & Kaufman	2.4	2.5	2.0	7.0	64	-.22	.11	.05
Wechsler								
Low	2.5	2.2	1.5	6.9	63	-.12	-.15	-.08
Average	3.1	2.5	1.6	7.2	65	.00	.01	-.00
High	3.5	2.9	1.9	7.5	68	.12	.13	.21

ponents may seem awkward at first sight. However, they are not surprising in the present context, where the components have been constrained to be orthogonal in the Wechsler set. Orthogonality can only be achieved by assigning small negative weights to those clusters of tests that belong to "other" components. These negative weights provide the correction for overlap needed to obtain orthogonal components from correlated clusters of tests. In those samples where the overlap is smaller than in the Wechsler set, overcorrection, and hence negative correlations between components, may be expected. This is precisely what seems to have happened. Tables 5 and 6 reveal that correlations between components vary around 0 whenever the explained variance (pointing to overlap) is of the same magnitude as in the Wechsler set (see, e.g., Hong Kong samples 11 and 12; Psychological Corporation, 1981). Conversely, the negative correlations tend to prevail in samples where the overlap is small (see, e.g., Groff & Hubble, 1982a, 1982b). It seems, therefore, that the negative correlations

are an artifact of using orthogonal components in the Wechsler set.

In conclusion, the factorial structures in all Other studies clearly resemble that of the WISC-R, but there are nevertheless marked differences between studies and with the Wechsler set, particularly lower explained variances and varying component correlations. Of course, it should be realized that the Other studies form too small and too heterogeneous a sample to permit an unequivocal statement about sample characteristics and factorial structure for special subgroups. Without further information from other studies, it is difficult to make general statements about the differences.

Study III: Importance of Wechsler Components in Cross-Cultural Studies

There exist several translations/adaptations of the WISC-R for which the correlations have been published (e.g., Chan, 1984—Hong Kong-WISC-R,

Table 6
Explained Variances and Correlations of the Wechsler Varimax Components
(Verbal Comprehension (VC), Perceptual Organization (PO), and
Freedom from Distractability (FD) in Cross-Cultural Studies

Study	Age	Explained Variance					Correlations			
		VC	PO	FD	Total	%	(VC,PO)	(VC,FD)	(PO,FD)	
Hong Kong	5	1.9	1.8	1.2	5.5	50	-.24	-.26	.06	
	6	2.4	1.8	1.5	6.0	55	-.08	-.12	-.02	
	7	2.3	2.1	1.7	6.5	59	-.16	.00	-.05	
	8	2.5	2.3	1.5	6.6	60	-.08	-.24	.17	
	9	2.8	2.5	1.7	7.1	65	.02	-.04	-.03	
	10	2.8	2.1	2.0	7.0	64	-.18	.03	.16	
	11	3.4	2.4	2.2	7.6	69	-.06	.01	.23	
	12	3.2	2.1	2.0	7.5	68	-.13	-.07	.18	
	13	3.3	2.0	2.3	7.3	66	-.16	.21	.02	
	14	2.6	2.2	1.8	7.1	65	-.16	-.02	.03	
	15	2.6	1.7	1.9	6.4	58	-.14	-.05	.11	
	Average		2.7	2.1	1.8	6.8	62	-.12	-.05	.08
	Pooled		2.7	2.0	1.7	6.7	61	-.13	-.05	.08
	HAWIK-R		3.1	2.6	1.9	7.1	65	.01	.01	.20
	Dutch WISC-R		2.8	2.1	1.5	6.8	62	-.12	-.06	.03
Hofman & Pijl	6&7	2.4	1.8	1.3	6.3	57	-.27	-.20	.02	
Hofman & Pijl	8&9	2.3	1.9	1.4	6.2	56	-.23	-.10	.03	
Wechsler	Low	2.5	2.2	1.5	6.9	63	-.12	-.15	-.08	
	Average	3.1	2.5	1.6	7.2	65	.00	.00	-.00	
	High	3.5	2.9	2.0	7.5	68	.12	.13	.21	

Cantonese; Pijl, 1982—WISC-R, Dutch; Titze & Tewes, 1984—HAWIK-R, German). In transferring the WISC-R to other countries, adaptations have had to be made to eliminate specifically American elements from the test. For instance, in their introduction Titze and Tewes (1984) stated that "with respect to the items the HAWIK-R was largely developed anew. The fundamental concepts of the subtests, however, had to be maintained for copyright reasons [translation PMK]" (p. 5). The Hong Kong Cantonese and Dutch versions had to undergo similar adaptations, even though they seem to be closer to the original. Given the variations, it is interesting to see to what extent the Wechsler components can be cross-validated in the Cross-cultural studies. The results are presented in Table 6, together with the results of a Dutch study using sam-

ples closely resembling the Dutch standardization samples.

The most detailed information is available from the Hong Kong Cantonese WISC-R (Chan, 1984; Psychological Corporation, 1981); notwithstanding several exceptions, especially for the younger age groups, the translation seems to allow good recovery of the Wechsler components. Overall the explained variance is not much lower than in the Wechsler set, but the correlations are somewhat higher, and the variations between the groups are larger. Still, the Cantonese version resembles the American WISC-R more closely than do the Other studies. Note that in this case there is no mixing of age groups. The most deviant group are the five-year-olds, who are in fact younger than the youngest Wechsler group. The low explained variance

for VC with the younger children, and the rather higher correlations of the other components with VC, conform to the abovementioned pattern for younger children in the Other studies and the Wechsler set.

The German results from the correlations averaged over the subgroups (Titze & Tewes, 1984) agree quite nicely with the Wechsler results of the average correlation matrix (except for the higher correlation between the PO and FD factors), despite the considerable adaptation that took place. Apparently, the fundamental concepts were captured adequately with the new items.

The Dutch WISC-R (Pijl, 1982) does not fare quite as well, but it should be remembered that the correlation matrix available is based on the scores of all age groups together; as suggested above, this might lead to lower agreement with the Wechsler data. Note, furthermore, that the application of the Dutch WISC-R in Hofman and Pijl (1983) stands in a similar position with respect to the Dutch standardization sample, as do the Other studies with respect to the Wechsler set: less explained variance and higher correlations between components. Finally, in Hofman and Pijl's youngest age group, the above mentioned pattern typical for younger children emerges yet again.

Discussion

Using all the data from the standardization samples for the Wechsler Intelligence Scale for Children-Revised, a single component space has been derived which was shown to be equally representative for all age groups. As mentioned before, other methods than the ones used here may be employed for the same purpose. Using the common component space from the Wechsler set as a standard, it was possible to assess the relevance of these components for a number of independently conducted studies and various translations of the WISC-R. Previous comparisons between studies depended primarily upon visually assessing the adequacy of a solution in terms of the results of similar studies. Analysis of a series of studies within the same framework permits a more rigorous comparison.

An attractive aspect of the perfect congruence approach is that the relevance of the Wechsler components can unequivocally be assessed in other studies using the amount of explained variance, according to the fundamental proposition that perfect congruence is always possible. Although the technique is applied to component weights rather than loadings, Harris (1985, pp. 317-320) makes a forceful case for weights as the basis for interpretation (see also ten Berge, 1986a, 1986b). As demonstrated above, the perfect congruence approach is easy to implement. In future research with the WISC-R, each researcher can directly compare the factorial structure found with the structures given in Table 2. Some claims about the factorial structure of the WISC-R which have been made by authors of some of the Other studies are discussed below.

The main difference some authors claim to have found is that no more than two of the three factors (i.e., VC and PO) found by Kaufman (1975) in the Wechsler set were present in their data. For instance, Reschly (1978, p. 419) found a third factor for his Anglo-American group but not for Blacks, Chicanos, or Native American Papagos. Sandoval (1982) indicated that the FD factor emerged with Anglos, but not with Blacks and Mexican-Americans, while Petersen and Hart (1979) found no clear FD factor in their three groups: "emotionally handicapped", "learning disabled and slow learners", and "no significant problems". Of the latter two studies, only the last mentioned correlation matrix was still available for reanalysis.

Comparing the above findings with the results in Table 5 to the extent that data are available, there seems to be no basis for Reschly's claim, as in all groups the FD factors have explained variances for this factor comparable to those of the Wechsler set; in fact, the Black Americans have one of the highest explained variances. Similarly, Sandoval's Mexican-Americans have a FD with an explained variance well within the Wechsler range. Also, in Petersen and Hart's "no significant problem" group, the FD factor can clearly be found (albeit with one of the smaller explained variances), but from their information on the third fac-

tors, their other groups may have less important third factors. Note that in contrast to Kaufman (1975), four-factor solutions for the 6½- and 16½-year-olds were not required in the present study to establish the three-factor structures for the Wechsler set.

Blaha and Vance (1979) and Vance, Wallbrown, and Fremont (1978) claimed that the ability structure for retarded persons may well be more complex than the structure for normals. Within the framework used thus far, it is not possible to assess their suggestion. To see if other factors of learning-disabled children are present in the Wechsler set, the presence of these factors, either in the Wechsler set or in their average correlation matrix, could be assessed by using the perfect congruence approach in the other direction. From the data available (Table 5), there is little indication that retarded persons have a different structure. The only exception may be that in three of the four samples of retarded children (Cummins & Das, 1980; Groff & Hubble, 1982b; Van Hagen & Kaufman, 1975), the VC and PO factors have nearly equal explained variances, in contrast to the unequal ones in the Wechsler set.

Kaufman (1981), in his review of the state of the art of the WISC-R and learning disabilities, stated:

It is time to call a halt to virtually all factor-analytic investigations of the WISC-R. Enough! We understand the factor structure of the instrument. We do not need to know more about slight differences in the two or three factors for various ethnic or exceptional groups . . . small differences in factorial composition from sample to sample cannot be attributed to ethnic membership or type of exceptionality; they are just as likely to be due to an irrelevant, uncontrolled variable or, most likely of all, to the chance fluctuations that are known to characterize correlation matrices. Future research in this area should focus on what the factors mean in either a theoretical or clinical sense. (p. 571)

Kaufman's plea is understandable, and undoubtedly correct from a clinical point of view. At the same time it should, however, be realized that his claim with respect to chance fluctuations and other

causes for differences between studies is conjecture and not firmly rooted in evidence.

As can be seen from the above discussion, the controversy is real, and authors have claimed to find important differences for special groups. The present results seem to provide the empirical evidence in favor of Kaufman's point of view, although a larger selection of correlation matrices would certainly make the point more strongly. Furthermore, the consistently lower explained variance for VC of younger children should be investigated, possibly by both test constructors and users. A reasonable conjecture for the age effect is that it is related to factor differentiation (see, e.g., Atkin, Bray, Davison, Herzberger, Humphreys, & Selzer, 1977). However, to show this with any degree of certainty, it is likely that even larger sample sizes are needed, as was noted by Atkin et al. (p. 75). Another aspect which should be investigated is the effect of age-scaled scores and the effect of mixed-age groups on the factorial structure of the WISC-R.

References

- Atkin, R., Bray, R., Davison, M., Herzberger, S., Humphreys, L., & Selzer, U. (1977). Ability factor differentiation, grades 5 through 11. *Applied Psychological Measurement, 1*, 65-76.
- Bentler, P. M. (1985). *Theory and implementation of EQS, a structural equations program*. Los Angeles: BMDP Statistical Software.
- Blaha, J., & Vance, H. B. (1979). The hierarchical structure of the WISC-R for learning disabled children. *Learning Disabilities Quarterly, 2*, 71-75.
- Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality*. Unpublished doctoral thesis, University of Groningen, The Netherlands.
- Burt, C. (1948). The factorial study of temperament traits. *British Journal of Psychology, Statistical Section, 1*, 178-203.
- Chan, D. W. (1984). Factor analysis of the HK-WISC at 11 age levels between 5 and 15 years. *Journal of Consulting and Clinical Psychology, 52*, 482-483.
- Cummins, J. P., & Das, J. P. (1980). Cognitive processing, academic achievement, and WISC-R performance in EMR children. *Journal of Consulting and Clinical Psychology, 48*, 777-779.

- Fraser, C. (1980). *COSAN user's guide*. Toronto: Ontario Institute for Studies in Education.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale NJ: Erlbaum.
- Groff, M. G., & Hubble, L. M. (1982a). *Factor analysis of WISC-R scores of an average ability sample*. Unpublished manuscript, Douglas County Family Service Clinic, Roseburg OR.
- Groff, M. G., & Hubble, L. M. (1982b). WISC-R factor structures of younger and older youth with low IQs. *Journal of Consulting and Clinical Psychology, 50*, 148-149.
- Harris, R. J. (1985). *A primer of multivariate statistics* (2nd ed.). New York: Academic Press.
- Harshman, R. A., & Lundy, M. E. (1984). The PARAFAC model for three-way factor analysis and multi-dimensional scaling. In H. G. Law, C. W. Snyder Jr., J. A. Hattie, & R. P. McDonald (Eds.), *Research methods for multimode data analysis* (pp. 122-215). New York: Praeger.
- Hofman, R. H., & Pijl, Y. J. (1983). *Vergelijkbaarheid van WISC-R en R-AKIT* [Comparability of WISC-R and R-AKIT]. Haren, The Netherlands: R.I.O.N.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409-426.
- Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods*. Chicago: National Educational Resources.
- Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years. *Journal of Consulting and Clinical Psychology, 43*, 135-147.
- Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.
- Kaufman, A. S. (1981). The WISC-R and learning disabilities assessment: State of the art. *Journal of Learning Disabilities, 14*, 520-526.
- Kroonenberg, P. M. (1983a). *Three-mode principal component analysis: Theory and applications*. Leiden, The Netherlands: DSWO Press.
- Kroonenberg, P. M. (1983b). Annotated bibliography of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology, 36*, 81-113.
- Kroonenberg, P. M. (1983c). Correlational structure of the subtests of the Snijders-Oomen Non-verbal Intelligence Scale. *Kwantitatieve Methoden, 4* (11), 40-51.
- Kroonenberg, P. M., & Brouwer, P. (1985). *User's guide to TUCKALS2, version 4.0* (WEP-reeks, WR 85-12-RP). Leiden, The Netherlands: University of Leiden, Vakgroep W.E.P.
- Kroonenberg, P. M., & de Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika, 45*, 69-97.
- Kroonenberg, P. M., & Lewis, C. (1982). Methodological issues in the search for a factor model: Exploration through confirmation. *Journal of Educational Statistics, 7*, 69-90.
- Kroonenberg, P. M., & ten Berge, J. M. F. (1985, July). *Cross-validation of components from several correlation matrices with perfect congruence analysis and three-mode principal component analysis*. Paper presented at the 4th European Meeting of the Psychometric Society, Cambridge, UK.
- Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29*, 177-185.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Petersen, C. R., & Hart, D. H. (1979). Factor structure of the WISC-R for a clinic-referred population and specific subgroups. *Journal of Consulting and Clinical Psychology, 47*, 643-645.
- Pijl, Y. J. (1982). *Revisie en normering van de Nederlandse WISC-R* [Revision and normalization of the Dutch WISC-R]. Haren, The Netherlands: R.I.O.N. Psychological Corporation. (1981). *Hong Kong Wechsler Intelligence Scale for Children Manual*. New York: Author.
- Quattrocchi, M., & Sherrets, S. (1980). WISC-R: The first five years. *Psychology in the Schools, 17*, 297-312.
- Reschly, D. J. (1978). WISC-R factor structures among Anglos, Blacks, Chicanos, and Native-American Papagos. *Journal of Consulting and Clinical Psychology, 46*, 417-422.
- Sandoval, J. (1982). The WISC-R factorial validity of minority groups and Spearman's hypothesis. *Journal of School Psychology, 20*, 198-204.
- Stedman, J. M., Lawlis, G. F., Cortner, R. H., & Achterberg, G. (1978). Relationships between WISC-R factors, Wide-Range Achievement Test scores and Visual-Motor Maturation in children referred for psychological evaluation. *Journal of Consulting and Clinical Psychology, 46*, 869-872.
- ten Berge, J. M. F. (1977). *Optimizing factorial invariance*. Unpublished doctoral dissertation, University of Groningen, The Netherlands.
- ten Berge, J. M. F. (1986a). Some relationships between descriptive comparisons of components from different studies. *Multivariate Behavioral Research, 21*, 29-40.
- ten Berge, J. M. F. (1986b). Rotation to perfect congruence and the cross-validation of component weights across populations. *Multivariate Behavioral Research, 21*, 41-64, 262-266.
- Titze, I., & Tewes, U. (1984). *Messung der Intelligenz bei Kindern mit dem HAWIK-R* [Measuring intelligence with the German version of the WISC-R]. Bern: Huber.
- Tucker, L. R. (1951). *A method for synthesis of factor*

- analysis studies* (Personnel Research Section Report No. 984). Washington DC: Department of the Army.
- Tucker, L. R. (1972). Relations between multidimensional scaling and three-mode factor analysis. *Psychometrika*, 37, 3-27.
- Van Hagen, J., & Kaufman, A. S. (1975). Factor analysis of the WISC-R for a group of mentally retarded children and adolescents. *Journal of Consulting and Clinical Psychology*, 43, 661-667.
- Vance, H. B., Wallbrown, F. H., & Fremont, T. S. (1978). The abilities of retarded students: Further evidence concerning the stimulus trace factor. *Journal of Psychology*, 100, 77-82.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Corporation.

Acknowledgments

Thanks are due to D. W. Chan, J. P. Das, M. G. Groff, C. R. Petersen, D. J. Reschly, and J. Sandoval for supplying some of the analyzed data, to Jan de Leeuw for suggesting the project, and to two reviewers for their penetrating comments.

Author's Address

All correlation matrices analyzed in this study are available on request. Send requests for reprints or further information to Pieter M. Kroonenberg, Department of Education, University of Leiden, P.O. Box 9507, 2300 RA Leiden, The Netherlands.