MEASURING CIVIC ASPECTS ON CLASS LEVEL -FROM THE INDIVIDUAL STUDENT LEVEL TO CLASS LEVEL USING SIX NEW SCALES

Peter Allerup
The Danish University of Education, Denmark

Abstract

This paper presents analyses based on class level and defines six new scales based in part on selected items from the international pool of items and, in part on items from a Danish national option consisting of supplementary items of particular interest. Items from the six scales were subject to validity tests similar to those for international CIVIC scales. It is shown how these six scales effectively contribute new information in relation to the international CIVIC scales. The aim is to construct new scales with outcomes closely related to the class level of analysis, and it is demonstrated how six-dimensional profiles erable classification of the classes into distinct categories. In fact, the use of a simple mathematical distance measure in the six-dimensional space immediately isolates two opposite and extreme profiles, which can both be straightforwardly interpreted as reasoned pictures of well known class characteristics. Other class profiles lie between these extremities as other combinations across the six scales. The profiles may be analyzed further by relating either of the two extremities to other characteristics of the class, like ratio of boys to girls and type of school.

INTRODUCTION

The Nordic countries Sweden, Norway and Denmark participated in the IEA CIVIC 18 year-old study using the three international booklets and, as a common national option, a fourth booklet. This fourth booklet was constructed as a questionnaire along the same lines as the international booklets, and it was included in the ordinary CIVIC rotation of booklets among the students. Thus, the fourth booklet's mixture of open and multiple choice-questions collected representative student information within the areas designed for the booklet. These were (Bruun and Johansen, 2004, Bruun et.al., 2003) issues of special relevance for students in the Nordic countries. At the same time it included specific questions, which could provide information about what happens

at the class level, i.e., 'between' students in a class rather than 'within' the individual students.

The aims of the present analyses are to show first, how information based on individual questions or items can be analyzed as scale scores from six new CIVIC scales: PROPAT, INDKOL, DEPOAK, ETNRET, ANTRET and GLOBAL. The scales are constructed partly from selected items from the international pool of questionnaire items, and partly from items in the fourth booklet. Emphasis will be given to efforts (see also Johansen and Bruun, 2004) to interpret individual student measurements taken by these six scales in relation to the specific class from which the students come. Second, the correlation between the six new scales and the international CIVIC scales will be studied in order to see if independent information is collected by the new scales. Finally, it will be demonstrated how a multidimensional statistical technique enables simple comparisons to take place among classes in the space spanned by all six scales, i.e., considering the scale scores from all six scales simultaneously.

THE DEFINITION AND PROPERTIES OF SIX SCALES

The international CIVIC scales were constructed using a two-step procedure (Amadeo et.al., 2002), where the first step comprised standard factor analyses of individual student responses and the second step involved statistical analysis *of homogeneity* of responses fixed by factor dimensions. While the results of first type of analysis rely solely on *individual correlations* between items in the total set of items investigated, the second type tests something very different, viz., the sufficiency of student scores. In fact, it is here tested, by means of Rasch Model analysis, if the scale scores exhaust *all individual* related information kept in the items, across which the individual score is calculated.

After careful content analysis, selected items from the standard CIVIC set of items and items taken from the fourth booklet were chosen for comparisons of the approx. 150 classes in the Danish data set. Originally, items were selected according to seven different research criteria, or seven different ways of describing student attitudes in a classroom. During the analyses of Rasch Model homogeneity it became evident, that the psychometric properties of items of one of the scales were so poor that the scale was abandoned. The six remaining scales were all accepted by analyses of homogeneity after a few adjustments, and labels to the six scales could be defined:

Scale INDKOL:	Individual values versus collective values
Scale ETNRET:	Inclusion of civil rights of ethnic minorities and immigrants
Scale ANTRET:	Inclusion of civil rights of anti-democratic groups.
Scale DEPOAK:	Political engagement in a wide rage of political activities.
Scale GLOBAL:	Global orientation
Scale PROPAT:	National protectionism (political, economical and cultural protectionism)

The statistical analyses used for validating the item homogeneity were initially the same (Andrich, 1978; Lehmann, 2004) as the analyses used for calibration of the

international CIVIC scales for both populations. It means looking for test of fit of a special version of the general Rasch Model (Rasch,1960) for four response categories:

$$(1) \qquad p(X) = \frac{e^{\theta + \Sigma t + \sigma}}{\Sigma e^{\theta + \Sigma t + \sigma}}$$

.

where the underlying response scale in the six scales consists of the thresholds "strongly agree", "agreee", "disagree" and "strongly disagree", as elsewhere in the international CIVIC scales. This model allows for interpretation of fit in terms of the thresholds $\Sigma\tau$, individual (attitude) level σ and item prevalence θ. Furthermore, a powerful and detailed test of fit (Allerup,1994,1997) is available, if student responses are dichotomized to "agree"/ "disagree" by collapsing the four response categories into two. In this case test of fit means evaluation of the so-called two parameter logistic Rasch model

$$(2) \qquad p(X) = \frac{e^{\theta + \delta\sigma}}{\Sigma e^{\theta + \delta\sigma}}$$

The model is here equipped with item discriminations, δ together with the usual item- and individual (attitude) parameters θ and σ. The later version allows for conditional inference and exact testing of fit.

Table 1: Correlations (Spearman) Between New Scales And The International CIVIC Scales

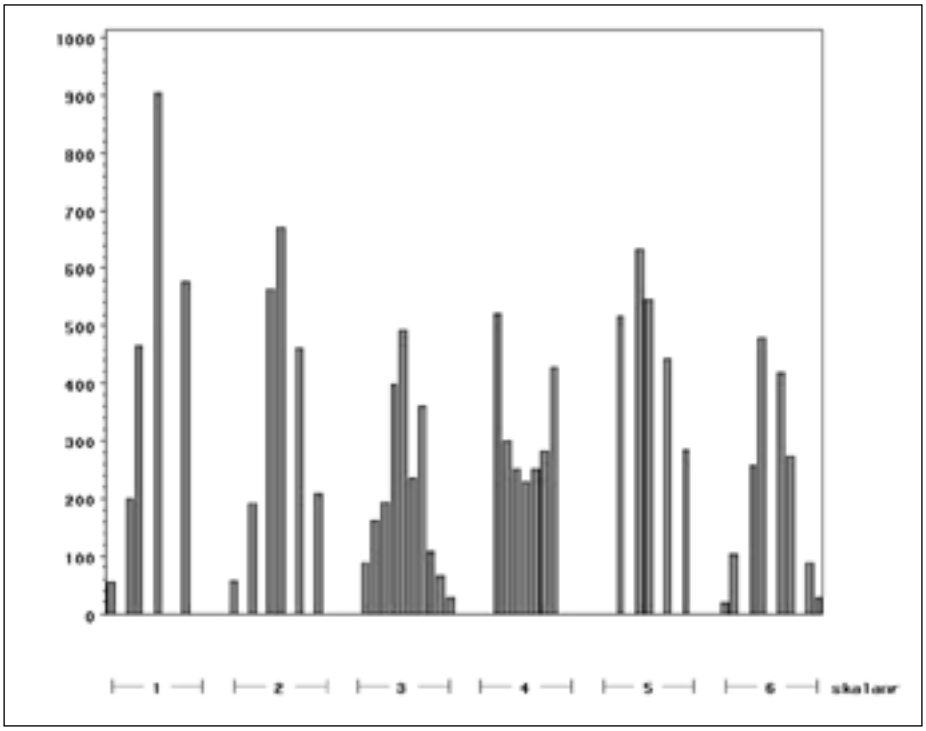
OldCIVIC/New	*PROPAT*	*INDKOL*	*DEPOAK*	*ETNRET*	*ANTRET*	*GLOBAL*
TOTCG	-0.15	0.09	0.21	0.22	-0.30*	0.07
KNOCG	-0.10	0.05	0.19	0.20	-0.26	0.08
SKICG	-0.11	0.06	0.17	0.20	-0.24	0.07
ECOCG	-0.12	0.11	0.17	0.15	-0.25	0.03
CTCON	-0.02	-0.04	0.30*	0.15	-0.37*	0.16
CTSOC	0.03	-0.21	0.27	0.24	-0.15	0.36*
GOVEC	0.08	-0.30*	0.11	0.20	-0.04	0.22
GOVSO	0.03	0.00	0.14	0.24	-0.16	0.19
TRUST	-0.02	0.00	0.05	0.04	-0.08	0.10
PATRI	0.30*	0.03	-0.07	-0.14	0.04	-0.05
WOMRT	-0.06	-0.14	0.12	0.43*	-0.10	0.23
IMMIG	-0.32*	-0.21	0.28	0.85*	-0.22	0.36*
CONFS	0.02	-0.12	0.10	0.25	-0.06	0.20
POLAT	-0.11	-0.03	0.82*	0.16	-0.80*	0.16
CCLIM	-0.07	0.01	0.13	0.18	-0.17	0.10

Both versions of the theoretical model were tested successfully against data, and fit of items underlying each of the six scales resulted therefore in the estimation of six latent scores (σ_1, ..., σ_6) for each student. For simplicity, average values are set to (10,...,10) as for the international scales.

The distributional characteristics of latent σ-scores from the six scales are shown in figure 1. Only the scale ETNRET: "Inclusion of civil rights of ethnic minorities and immigrants" seems to exhibit a true non-normal behavior. In fact, one gets an impression of polarization of attitudes from the bimodal distribution.

The generally weak correlations found in table 1, and the unstructured appearance of significant co-variations (*flagged correlations are significant on p = 0.05) between the new scales and the international CIVIC scales indicate that independent information is collected by the new scales.

Figure 1. Distributions Of Latent σ-Scores For New Scales Nos. 1...6.



In table 2 the interclass correlations (Rao, 1965) are displayed for the international CIVIC scales and for the new scales. It can be read from the table that the average level of interclass correlation is *higher* for the new scales compared to the international scales. This could be taken as an indication that the new scales are dealing with attitudes, which are created in relation to *other* students in a classroom, i.e., the attitudes of the individual students are dependent on attitudes of the other students in the same class.

It may be summarized from the definition of the six new scales that the scales measures something different compared to the international CIVIC scales, and that student-by-student measurements tend to correlate to a higher degree compared to parallel measurements taken by the international CIVIC scales.

Table 2: Interclass Coefficients, IC For International CIVIC Scales And For New Scales.

Scale	*IC-coefficient*
CTCONMLE	0.04
CTSOCMLE	0.04
GOVECMLE	0.05
GOVSOMLE	0.01
TRUSTMLE	0.02
PATRIMLE	0.03
WOMRTMLE	0.05
IMMIGMLE	0.12
CONFSMLE	0.06
POLATMLE	0.03
CCLIMMLE	0.09
PROPATMLE	0.06
INDKOLMLE	0.09
DEPOAKMLE	0.10
ETNRETMLE	0.12
ANTRETMLE	0.07
GLOBALMLE	0.09

DEFINING RESPONSE PROFILES

One aim of these analyses was to be able to define student *response profiles* based on measurements taken by the new scales. On the practical level this means that instead of studying *marginal* comparisons for each of the new scale σ' s at a time, comparisons should be based on *simultaneous* measurements on all six σ's: $(\sigma_1,...,\sigma_6)$.

The definition of a profile can, theoretically, be initiated from (nearly) any definition of a distance measure on the six-dimensional $(\sigma_1,...,\sigma_6)$'s. Whenever a group or a cluster of students can be defined as a result of measuring distances between the $(\sigma_1,...,\sigma_6)$'s, a profile is defined. The technique is well known from various cluster analysis and discriminant analysis procedures. Few of these procedures, however include distributional characteristics and intercorrelations between the variables.

A distance measure, which explicitly takes into account the correlations among the new scales, i.e., among the σ's in $(\sigma_1,...,\sigma_6)$, could be the Mahalanobis Distance (Rao, 1965). This is a distance measure in the space spanned by the six scales, which make use of the 'average scale level' and 'spread of responses on the scale' (standard deviation). Furthermore, it takes into account the correlation structure of the scales by attributing more 'length' to the distance between two fixed points placed in high correlating scales (co-ordinate axes) compared to independent scales. The mathematical expression for the distance between the individual multivariate student response $\sigma_v = (\sigma_{1v},...,\sigma_{6v})$, for student No. v, and the total average $\bar{\sigma} = (\bar{\sigma}_1, ..., \bar{\sigma}_6)$ is

$$(3) \qquad M_v(\sigma_1,..., \sigma_6) = (n - 1)\ (\sigma_v - \bar{\sigma})'\ \ (X'X)^{-1}\ (\sigma_v - \bar{\sigma})$$

As stated, the measure is suitable for measuring distances between multivariate student responses σ_v in relation to their common multivariate mean value $\bar{\sigma}$, but it can also be used for the calculation of distances between class means $\bar{\sigma}_1$, ..., $\bar{\sigma}_k$. In this case $(\sigma_v - \bar{\sigma})$ is changed to $(\bar{\sigma}_q - \sigma_k)$ for classes q and k, i.e., the distance between mean values for these two classes. Thus, by means of (3) it is possible to calculate individual student distances in relation to the class mean (*within* class distances) as it is possible to calculate individual student distances in relation to the grand mean for all students, $\bar{\sigma} = (\bar{\sigma}_1, ..., \bar{\sigma}_6)$, which was fixed to (10,...,10). Moreover, these analyses allow for the possibility of calculating distances between class mean values (*between* class distances).

The Mahalanobis Distance measure proved useful earlier (Allerup, 2003) when analysing the response behaviour across *all* eleven international CIVIC scales, using the international data set. In fact, figure 2 displays a distinct pattern of regularity found earlier for the international CIVIC data set: On average girls 'distribute' their responses on the eleven scales in a more 'narrow' way around the neutral point between "agree" and "disagree" compared to the boys, i.e., they tend not to be using the *extremes "highly agree" /"highly disagree"* on the underlying response scale as much as the boys.

Application of this technique conducted on class means, $\bar{\sigma}_1$, ..., $\bar{\sigma}_k$, for the six new scales produces three examples shown in table 3 of pair of classes, which have extreme Mahalanobis Distances. The pairs are denoted "1" and "2" and it is seen, how the class means differ between the two classes forming a couple of extreme 'pairs'. Examples from other 'corners' of the multidimensional ellipsoid spanned by the new scales (a rugby football) could of course be given. A total of 142 classes is available in the Danish Civic data set.

*Figure 2: Within Country Mahalanobis Distances, All students in the international data set, grouped by gender: Girls (1 ⇔ *), Boys (2 ⇔ •)*

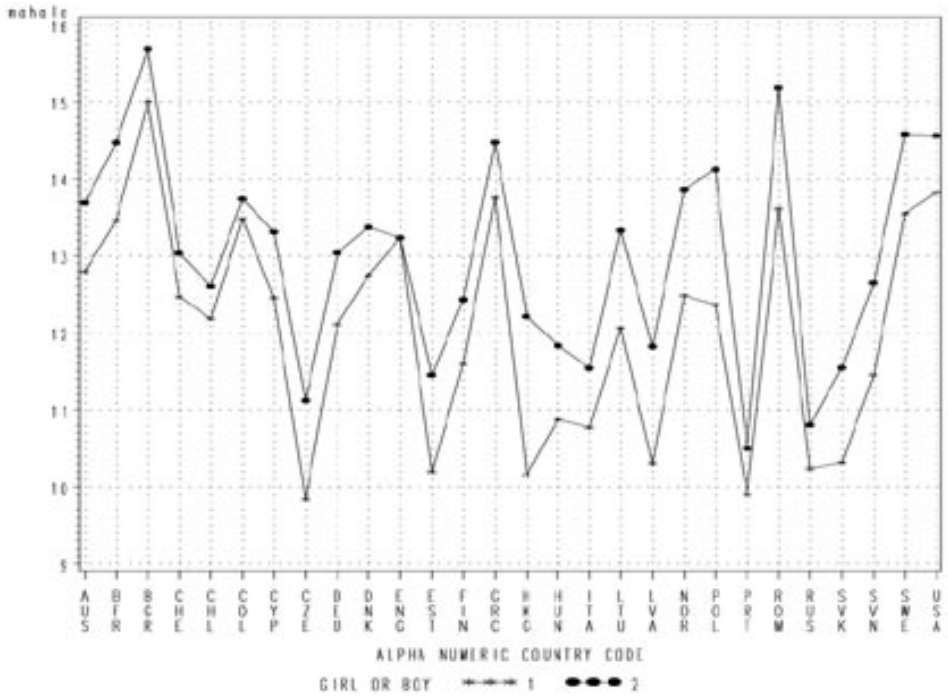
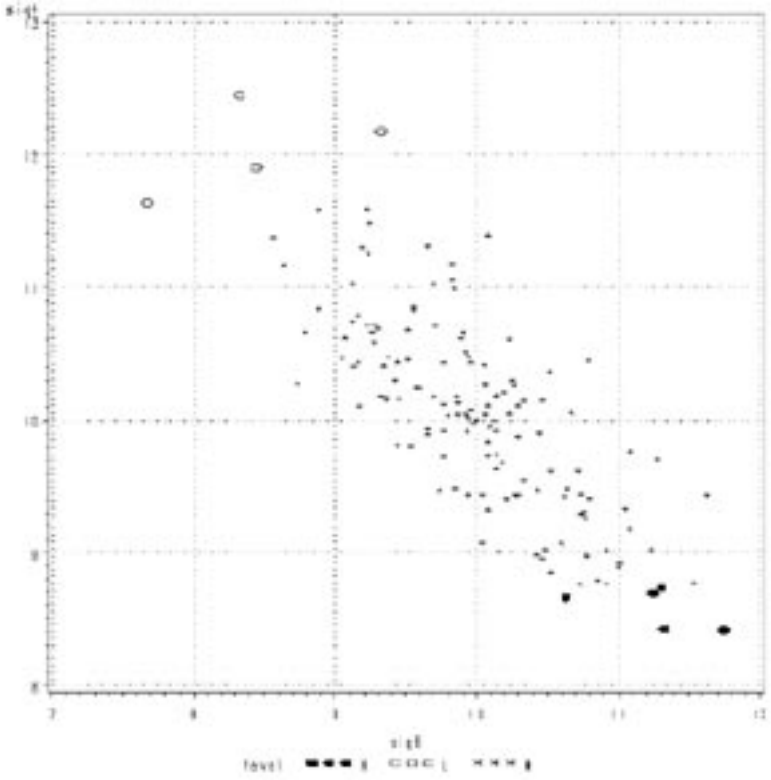


Table 3: Three Examples On Pairs Of Classes With Extreme Mahalanobis Distances

New scales	*Three pairs of classes (1,2)*					
	A1	*A2*	*B1*	*B2*	*C1*	*C2*
PROPAT	8.47	9.33	10.97	11.27	8.42	9.52
INDKOL	9.54	9.91	6.51	9.92	10.47	8.05
DEPOAK	11.63	9.02	11.48	8.69	10.47	12.17
ETNRET	11.81	9.57	12.03	9.39	10.60	11.64
ANTRET	7.67	11.23	9.25	11.24	9.95	9.32
GLOBAL	12.01	7.69	10.61	8.51	10.24	9.08

A visualization of how extreme class means are placed in the six-dimensional ellipsoid can be seen from a two-dimensional "cut" along the DEPOAK - and ANTRET axes. Notice that the usual variance reduction on the mean, $\hat{\sigma}^2/n$ is now increased to $\hat{\sigma}^2/n + (n-1)/n\sigma^2\rho$, with ρ as the intra-class correlation. The mean is therefore calculated with less statistical accuracy. In figure 3 each point represents a class mean and the three pairs A,B and C can be found between the combinations of "o" and "•" in the graph.

Figure 3: Latent σ-scores for DEPOAK (Y-Axis) and ANTRET (X-axis). Each point represents a class mean. Extreme classes are marked by "o" and "●".



It can be summarized from these profile analyses that it is possible to 'map' the 142 (six-dimensional) class means $\bar{\sigma}_1$, …, $\bar{\sigma}_k$ as points in the ellipsoid spanned simultaneously by the six latent σ's from the new scales. Is has been *possible* to establish a simple distance measure between points in this rugby-football sphere that takes into account correlations between the six scales. Thus, whenever a pair of extreme points can be detected, a *possible* profile is at hand - the members of a pair providing opposite "ends". However, this crude mathematical definition of a profile must be accompanied by some sort of CIVIC-related interpretation, before the "profile" can be used operationally as creator of new information about the CIVIC classes. The suggested profiles need labels.

CREATING LABELS ON THE PROFILES

The procedure leading to table 3 is easy to repeat on all 142 classes, with the result, that more than three pairs show up as points in the extreme "ends". It is, of course, then due to subjective judgment when one wants to classify the points on one end as "extreme 1" (denoted category "I" in the sequel) and their extreme partners on the other end as "extreme 2" (or category "K" in the sequel). In-between these two

extremes exists a variety of $\bar{o} = (\bar{o}_1, \ldots, \bar{o}_6)$ combinations. For simplicity they are at present put into one category: "B".

A sensible grouping of the distance measures (mainly respecting statistical needs for a minimum number of observations in each category) leads to a grouping of all 142 classes into the following groups: Category "I", 29 classes (20,42 %), category "B" in the middle, 83 classes (59,15 %) and category "K"in the opposite end containing 29 classes (20, 42 %).

It is remarkable that the application of this crude clustering of classes immediately points to two easy-interpretable profiles, "I" and "K", listed below (the list should throughout have a prefix like "Relatively ..." to stress that the categories are not created from absolute definitions).

Profile "I"	*Profile "K"*	*Scale*
high degree of Individualism	high degree of collectivistism	INDKOL
high degree of National protectionism	low degree of National protectionism	PROPAT
low degree of political participation	high degree of political participation	DEPOAK
low degree of inclusion of rights of ethnic minorities	high degree of inclusion of rights of ethnic minorities	ETNRET
low inclusion of rights of anti-democratic groups	high inclusion of rights of antidemocratic groups	ANTRET
low global orientation	high global orientation	GLOBAL

A very brief and oversimplified interpretation is then at hand by simply taking profile "I" classes to be collections of right-winged students, who are individualistic ally minded, against the inclusion of rights of ethnic- and antidemocratic groups, do not participate in politics, are highly protective nationally and have little feelings for global issues. This is a clear profile, but before any conclusions are drawn as to the frequency of classes in this category (20.42%), one should remember the arbitrariness by which the borders between the three categories were fixed.

A first labeling of the profiles could, consequently, be undertaken by *solely* looking at the captions of the six scales.

Usually a way of expanding the understanding of the categories I, B and K is undertaken by analyzing the categories in relation to various background variables for the classes. The following analyses take "frequency of boys in the class", the "school type" and, as a historic variable, which proved sensitive in many IEA studies, the "number of books" in the home.

The relation with "frequency of boys" in the class can be studied in table 4.

Table 4: Distribution Of 142 Classes (Three Missing) Over Categories I, B and K and Percentage Boys In The Class.

Percentage boys	*I profile*		*B profile*		*K profile*		*Total N*
	N	*%*	*N*	*%*	*N*	*%*	
Less than or equal 30%	5	17.2%	20	24.4%	16	57.1%	41
More than 30%, but less than 48%	10	34.5%	31	37.8%	10	35.7%	51
More than 48%	14	48.3%	31	37.8%	2	7.1%	47
Total	29	100%	82	100%	28	100%	139

It is clear from table 4, that a strong correlation exists between type of category and the number of boys in the class. Out of 28 K profile classes 57.1% have less than 30% boys in the class. Only in two cases (7.1%) a K profile class hold more than 48% boys. In summary, classes holding many girls tend to be K profiles.

Students from the CIVIC-old study come from four different types of school: Alment gymnasium as the general public high school, Hf being a type of school, designed for students older than alment gymnasium and usually not full-time students, and Hhx, Htx types of high schools which emphasize business - and technical (science and mathematics) issues, respectively.

The relation between school type and the I, B, K - categorization is displayed in table 6.

Table 5: Distribution Of 142 Classes (One Missing) Over Categories I, B And K And school Type.

School type	*I profile*		*B profile*		*K profile*		*Total*
	N	*%*	*N*	*%*	*N*	*%*	*N/ %*
Alment gymnasium	0	0.0%	48	66.6%	24	33.3%	72/100%
Hf	7	28.0%	14	56.0%	4	16.0%	25/100%
Hhx	17	53.1%	14	43.8%	1	3.1%	32/100%
Htx	5	41.7%	7	58.3%	0	0.0%	12/100%
Total	29	20.6%	83	58.9%	29	20.6%	141/100%

A remarkable correlation between school type and the three categories can be seen in table 5. It can be seen that the alment gymnasium schools are dominated by K profile classes while the school types Htx and Hhx are dominated by I profile classes. This is in accordance with the fact that Htx is a very subject-oriented school type with emphasis on hard core subjects like science and mathematics, which mainly attracts boys, while the alment gymnasium with its emphasis on other soft subjects as well, mainly attracts girls.

The last illustration comes from combining the variable "number of books at home" with the three categories (see in table 6).

Table 6: Distribution Of 142 Classes (One Missing) Over Categories I, B and K And The Number of Books At Home.

Number of books at home	*I profile*	*B profile*	*K profile*	*Total N*
approx 50-100	17.2%	9.6%	0%	13
approx 101-200	75.9%	57.8%	48.3%	84
More than 200	6.9%	32.5%	51.7%	44
Total	29	83	29	141
	100%	100%	100%	

The 'classic' IEA background variable "Number of books at home" is effective also in this case: Significantly more K profile classes contain students with many books at home (complete absence of observations in the lowest book-category 50-101), while classes dominated by I profiles have students with generally fewer books at home.

CONCLUSIONS

Based on content analysis of items from the international CIVIC set and items from the Danish national option of items, six groups of items were isolated. The content analysis aimed at identifying aspects of student attitudes and student behavior that could be related to interactions between students in a fixed class.

The six groups of items were subject to extensive scale analysis, part of which were the same scale analyses undertaken prior to the construction of the international CIVIC scales. The resulting scales collect independent information compared to the information collected by the international CIVIC scales; higher interclass correlations, measured by the six scales, supported the impression that responses to the six scales reflected student interactions on the *class level.*

It was part of the intention, that subsequent analyses of the six scales should be undertaken considering *simultaneous* measurements from the six scales instead of looking at *marginal* analysis results, emerging from analysis of the six scales, one at a time. Simultaneous measurements or *profile* analyses as they are denoted. Only by such analyses could the correlations among the six scales in a profile be properly handled.

A well-known statistical distance measure, the Mahalanobis Distance, matched scale requirements for the six scales, including correlations among scales. This distance measure was applied to 142 class means in the Danish data set. An empirical finding was that two extreme class profiles emerged from the analyses and could be identified by 'name', using the labels from the six scales. Further analysis studying the extreme groups in the environment of background variables expanded the understanding of the nature of the profiles.

It seems, therefore, that intensive work with item content analysis, prior to the scale analysis, pays off in the sense that empirical findings found later by such scales provide coherent, immediately interpretable information for the research. Another conclusion which can be drawn from this study is that the potential of *simultaneous profile analyses* is greater than that of *marginal analyses.*

When new studies involving student attitudes are planned, the more direct inclusion of student interaction should be considered. This can be achieved by forming matrices using *all students* in a class as entries for questions dealing with "interaction" issues between students. Only from such detailed pictures, where every student in a class "rates" other students in the same class, can one truly evaluate and interpret measures of interclass correlation.

References

Amadeo, J.-A., Torney-Purta, J., Lehmann, R., Husfeldt, V. & Nikolova, R. (2002). *Civic Knowledge and Engagement. An IEA Study of Upper Secondary Students in Sixteen Countries*. Amsterdam: IEA.

Allerup, P. (1997). *Statistical Analysis of Data from the IEA Reading Literacy Study; in: Applications of Latent trait and latent Class models in the Social Sciences.* Waxmann.

Allerup, P. (1994) *Rasch Measurement, theory of*; The International Encyclopedia of Education, vol8, Pergamon.

Allerup, P. (2003). *Democratic Values and Civic Knowledge - Analysis of Student Attitudes* (paper in preparation).

Andrich, D. (1978). *A rating formulation for ordered response categories.* Psykometrika 43, 561-573.

Bruun, J. Johansen J. (2004): *Interpreting civic aspects at school class level in Denmark - an educational anthropological perspective*. paper at IRC-2004, Cyprus.

Bruun, J., Johansen, J. og Allerup, P. (2002). *Rum for politisk dannelse*. København: Danmarks Pædagogiske Universitets Forlag (in Danish).

Lehmann, R. et al. *Technical Report: IEA Civic Education Study*. Amsterdam: IEA (in preparation).

Rao, C. R. (1965). *Linear Statistical Inference and its Applications*. New York: Wiley.

Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests.* K_benhavn: Danmarks Pædagogiske Institut.