University of Tartu
Department of Psychology

Aire Raidvee

DOUBLE SKEW-DUAL SCALING (DSDS) AND RASCH ANALYSIS OF RAVEN MATRICES' (RSPM) DATA

Master's thesis

Supervisors: Jüri Allik, *PhD*
Aavo Luuk, *candidate (psychology)*

Running head: Application of DSDS on RSPM data

Tartu 2005

CONTENTS

Abstract

The aim of the current work is to shed light onto an original method called *double skew-dual scaling* (DSDS) (Dzhafarov, 1999), with emphasis on its psychometric properties. An overview of DSDS is given and the method is applied to random as well as real data. The latter consists of Estonian schoolchildren's results on Raven's matrices. For reference, the same data are Rasch-analyzed. The results of DSDS are compared to these of traditional methods. Relationships of dominance and subdominance values across grades and sexes are examined. Possible applications of DSDS and interpretations of dominance and subdominance indices are discussed. Source codes for implementing DSDS in Statistica and R are included.

Keywords: Methodology of psychometrics; Double skew-dual scaling; Rasch measurement; Raven's progressive matrices

Summary in Estonian

Kokkuvõte

Kaksik-kald-duaalse skaleerimise rakendamine Eesti koolilaste Raven'i maatriksite andmetele võrdluses Rasch'i analüüsiga

Käesoleva töö eesmärk on tutvustada algupärast meetodit nimega kaksik-kald-duaalne skaleerimine (ik *double skew-dual scaling*, *DSDS*) (Dzhafarov, 1999), rõhuga meetodi psühhomeetrilistel omadustel. Antakse ülevaade DSDS sisust ning rakendatakse meetodit juhu- ning reaalsetele – Eesti koolilaste Raveni testi – andmetele. Võrdluseks rakendatakse samadele andmetele ka Rasch'i analüüsi. Võrreldakse DSDS tulemusi traditsiooniliste meetodite omadega. Uuritakse dominantsus- ja subdominantsussuhete seoseid klasside ja sugude lõikes. Arutletakse erinevate DSDS metoodika rakendamise ja dominantsus- ning subdominantsusindeksite tõlgendusvõimaluste üle. Tuuakse programmikoodid DSDS rakendamiseks pakettides Statistica ja R.

Märksõnad: psühhomeetria metodoloogia; kaksik-kald-duaalne skaleerimine; Rasch'i ühemõõtmelised mudelid; Raven'i progresseeruvad maatriksid

Application of Double Skew-Dual Scaling (DSDS) to Estonian schoolchildren's Raven matrices' (RSPM) data

The main objective of the current work is to shed some new light onto an original method called Double Skew-Dual Scaling (DSDS) by E. Dzhafarov (1999) with an emphasis on investigation of its psychometric properties by example of an analysis of RSPM data of Estonian 6^{th}- to 12^{th}-graders. The present study is confessedly the first attempt to apply DSDS on real data.

To provide the reader with some comparative evidence, the same data are also fitted to the Rasch simple logistic model (SLM).

Introduction of DSDS

DSDS is a method allowing to disentangle the tendencies to dominate or be subdominated between row objects $\{\alpha_1,\ldots,\alpha_n\}$ and column objects $\{\beta_1,\ldots,\beta_m\}$ in a data matrix filled with boolean, probability or degree of relatedness values. By this, DSDS does not set any constraints on the relationship between matrices' elements – it is a quantitative descriptive language that provides a "reasonable" summary of the relationship between $\{\alpha_1,\ldots,\alpha_n\}$ and $\{\beta_1,\ldots,\beta_m\}$ and can be utilized in formulating falsifiable models for the structure of dominance matrices (Dzhafarov, 1999).

As a result, DSDS provides each row and column object in the matrix with two numerical values – index reflecting the object's overall tendency to a) dominate and b) be subdominated by column or row objects, respectively. The relationship between subdominance- and dominance-vectors is complex: intuitively, D_α (the dominance value for an α-object) characterizes the tendency of an α-object to dominate β-objects with large dominance values, whereas S_α characterizes the tendency of an α-object to fail to dominate β-objects with large subdominance values (Dzhafarov, 1999).

For example, one could think of α-objects as mountain-climbers, and β-objects as hills that the climbers are trying to surmount. The observable trait would be person's ability to climb a mountain. We would have a number of people climb a number of mountains and have as a result a matrix filled with boolean values – the result of one person climbing one hill would be scored 1 if the person reaches the top, or 0, if the person fails at any

point on the way up (or, for a polytomous case, we could define some ratio of the hill that the person has to reach in order to gain some score on a polytomous scale with a certain number of thresholds). Now, some of the causes of the possible failure are intrinsic to the properties of the hill, (e.g the profile of the hillside), whereas others are intrinsic to the properties of the mountaineer, (e.g endurance, preparedness), yet thirds are due to the interaction of the properties of the hill and these of the person (e.g, for some persons, the openness of the (if existent) bar on the foot of the hill would be decisive of failure to reach the top). The last group of influences on the outcome stem from random coincidents (e.g meeting an urgently troubled friend, injury) – which ultimately boil down to the main or interaction effect of the properties of the person and properties of the hill.
Now, a dominant α-object would be a mountaineer who has a tendency of succeeding to surmount hills that other dominant mountaineers have failed to surmount. If, at the same time, the specific mountaineer fails to surmount several hills that other subdominant mountaineers have surmounted, this person would simultaneously have a high dominance value and a high subdominance value.

The gist of the DSDS could be represented with the following formulas -

$$\begin{cases} A_{n\times m} \dfrac{D^{\beta}_{m\times 1}}{\sum D^{\beta}_{m\times 1}} = D^{\alpha}_{n\times 1} \\ \\ B^{T}_{m\times n} \dfrac{D^{\alpha}_{n\times 1}}{\sum D^{\alpha}_{n\times 1}} = D^{\beta}_{m\times 1} \end{cases} \qquad (1)$$

and

$$\begin{cases} B_{n\times m} \dfrac{S^{\beta}_{m\times 1}}{\sum S^{\beta}_{m\times 1}} = S^{\alpha}_{n\times 1} \\ \\ A^{T}_{m\times n} \dfrac{S^{\alpha}_{n\times 1}}{\sum S^{\alpha}_{n\times 1}} = S^{\beta}_{n\times 1} \end{cases} \qquad (2)$$

where

$A_{n\times m}$ is the original data matrix of relatedness values,

$B^T_{m\times n}$ is the transpose of the complementary matrix of $A_{n\times m}$,

$D^{\alpha}_{n\times 1}$ is the dominance vector for α-objects (i.e row-objects),

$D^{\beta}_{m\times 1}$ is the dominance vector for β-objects (i.e column-objects),

$S^{\alpha}_{n\times 1}$ is the subdominance vector for α-objects,

$S^{\beta}_{n\times 1}$ is the dominance vector for β-objects,

$B_{n\times m}$ is the complementary matrix of A,

$A^T_{m\times n}$ is the transpose of the original matrix $A_{n\times m}$.

Formulas (1) and (2) convey the idea that

(i) α-dominance values characterize the rows of matrix $A_{n\,x\,m}$ in such a way that the dominance value for α_i (i = 1, ..., n) be the mean of the ith row of $A_{n\,x\,m}$ weighted by the β-dominance values;

(ii) β-dominance values characterize the columns of matrix $B_{n\,x\,m}$ in such a way that the dominance values for β_j (j = 1, ..., m) be the mean of the jth column of $B_{n\,x\,m}$ weighted by the α-dominance values;

(iii) α-subdominance values characterize the rows of matrix $B_{n\,x\,m}$ in such a way that the subdominance value for α_i (i = 1, ..., n) be the mean of the ith row of $B_{n\,x\,m}$ weighted by the β-subdominance values;

(iv) β-subdominance values characterize the columns of matrix $A_{n\,x\,m}$ in such a way that the subdominance value for β_j (j = 1, ..., m) be the mean of the jth column of $A_{n\,x\,m}$, weighted by the α-subdominance values.

For the algorithm of calculating the dominance and subdominance values for α- and β-objects, refer to Appendix 1 or Appendix 2.

Note that DSDS has only superficial similarities with the classical dual scaling (Nishisato, 1980).

Rationale for using the Rasch Simple Logistic Model

The Rasch model (Rasch, 1960) is chosen to provide the background evidence for DSDS results for its objective measurement properties – i.e possibility to estimate item and person parameters separately – as well as potential to test for unidimensionality and differential item functioning. The algebraic representation of the model is –

$$P(X_{is} = 1|\theta_s,\beta_i) = \frac{e^{(\theta_s - \beta_i)}}{1 + e^{(\theta_s - \beta_i)}} \qquad (3)$$

Where

e = 2,7182 is the base of natural logarithm,

θ_s is the level of the measurable latent ability of person *s*,

β_i is the threshold (difficulty) of item *i*.

For more information on the Rasch model see, for example, Andrich (1988); Fischer and Molenaar (1995); Linacre, J. M. (1989).

Method

A double skew-dual scaling program was prepared in STATISTICA 6.0 (StatSoft, Inc., 2003) Visual Basic (for source code, refer to App. 1) after the algorithm presented in the original DSDS paper by E. Dzhafarov (1999) to find dominance and subdominance vectors for students and test items.

The data was fitted to the Rasch model using RUMM2020 (Rasch Unidimensional Measurement Models, RUMM Laboratory Pty Ltd., 1997-2004).

DSDS was applied to random data in R software environment (R Development Core Team, 2005) (for source code, refer to App. 2).

Factorial ANOVA and t-tests were performed in STATISTICA 6.0.

Application of DSDS to random data

DSDS algorithm was realized in R software environment[1] (R Development Core Team, 2005). (For source code, refer to App. 2.) and applied to square matrices (of dichotomous random numbers) of dimension from 1000 to 2000. It appears that for dichotomous random data the relationship between dominance and subdominance levels for row, as well as column objects is linear, – correlation between dominance and subdominance levels regresses to –1 as the dimension of the matrix increases (see Figure 1).

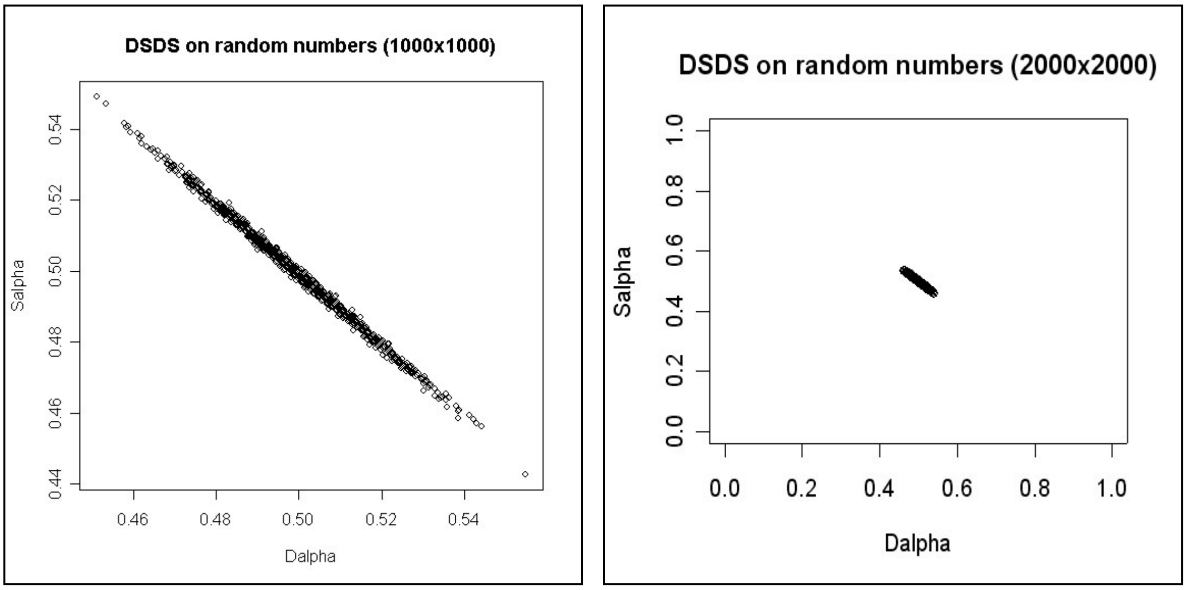

Figure 1. DSDS applied on matrices of random numbers of different dimension [2].

For square matrix of 1000^2 random numbers of set {0; 1}, correlations between dominance and subdominance levels for both row and column objects were –0.998, whereas for the matrix of size 2000^2 the same correlations were –0.999.

[1] I am indebted to Kenn Konstabel for the idea of testing DSDS on random numbers and for the help with programming in R software environment.

[2] Note the differences in scale between left and right panels. The left panel is to give better idea of the shape of the distribution, whereas the right panel is to provide comparison with the real data. (For column objects, the distributions are similar.)

Application of DSDS and Rasch analysis to real data

Sample

The data were collected in the framework of project *Academic Achievement as a Function of Mental Abilities, Self-Esteem and Personality* (grant of Estonian Science Foundation no 4519, grantholder: Jüri Allik, project manager: Helle Pullmann) by administering the Raven's Standardized Progressive Matrices (RSPM) to Estonian 6th-, 8th-, 10th- and 12th-graders in 2001 (with mean ages of 12.4, 14.4, 16.1 and 17.8 years respectively; *n* = 2739) and as a retest to 8th-, 10th- and 12th- graders in 2003 (with mean ages of 14.5, 16.4 and 18.1 years respectively; *n* = 911). The whole age range was from 11 to 21 in 2001; and from 13 to 20 in 2003[3].

The test was administered as a group-test and without time limit. RSPM consists of five subscales A, B, C, D and E, each of 12 items, with multiple choice questions of 6 alternatives in subscales A and B, and 8 alternatives in subscales C, D and E. For full details of this test refer to Raven (1981).

General overview of relations of dominance and subdominance in the real data

To give some preliminary overview, some relationships with classical testing theory results are shown on the sample collected in 2001.

The relationship between schoolchildren's (*n*=2739, data gathered in 2001) tendency to dominate or be subdominated by the test items appears nonlinear (see Figure 2).

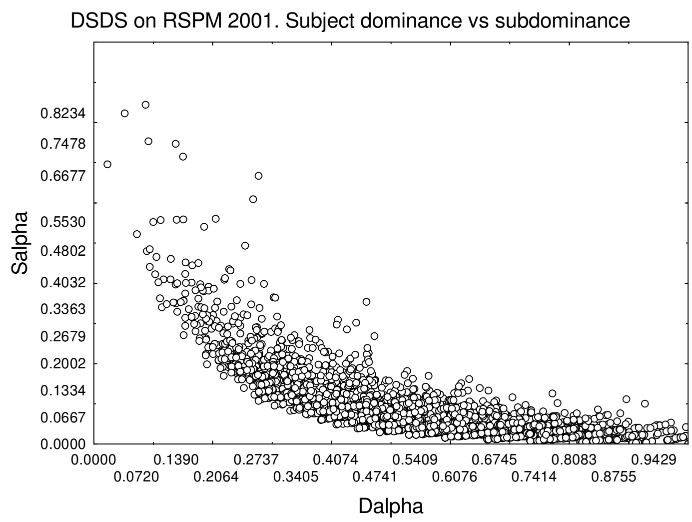

Figure 2. Distribution of subject dominance vs subdominance levels for the whole sample of 2001. Salpha: subdominance level of the student; Dalpha: dominance level of the student.

[3] The pupils of deviant age in their respective grade-group were included in the DSDS analyses, too, because the aim was not to fit the data to a model but rather to give way to discovering psychometric properties of the method.

As is evident from Figure 2, the distribution is clearly different from the respective relationship in random data, thus allowing to hypothesize a psychological principle reflected in DSDS results. Given the alternative-choice nature of the Raven matrices test, it is expected that the distribution asymptotically approaches but does not reach the y-axis. It might be expectable for the student with a high dominance level to be low in subdominance but it shows students exist whose dominance, as well as subdominance levels are comparatively high, and vice versa.

Correlation between dominance and subdominance values for students is –0.755. The general shape of the distribution is similar across sexes as well as classes. For boys (n=1273) the mean dominance value is 0.53, whereas for girls (n=1466) the value is 0.54; the mean subdominance values for boys and girls are 0.1 and 0.09, respectively. Expectedly, the variance of dominance and subdominance is greater for boys than girls: SDs for dominance and subdominance are 0.21 and 0.93 for boys compared with 0.19 and 0.86 for girls.

For the test items, the relationship between item dominance and subdominance values is nonlinear as well (see Figure 3). Correlation between dominance and subdominance values for test items is –0.957.

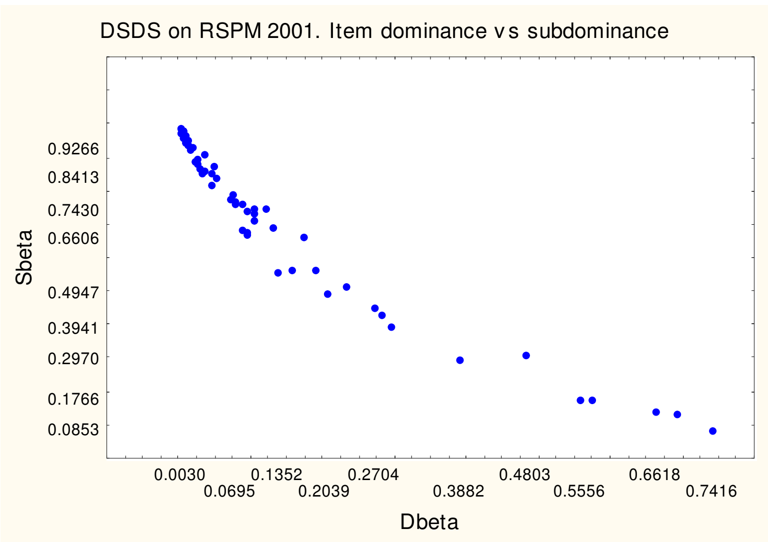

Figure 3. Distribution of item dominance vs subdominance levels for the whole sample of 2001. Sbeta: subdominance level of the item; Dbeta: dominance level of the item.

Relations of dominance and subdominance values to traditional parameters

Correlations between students’ dominance and subdominance values and proportions of correct responses are 0.885 and –0.972, respectively.

Classical item difficulty levels (in terms of percent of correct answers) correlate to item dominance values as highly as 0.978; correlation between item difficulty levels and subdominance values is –0.996 (see Figure 4).

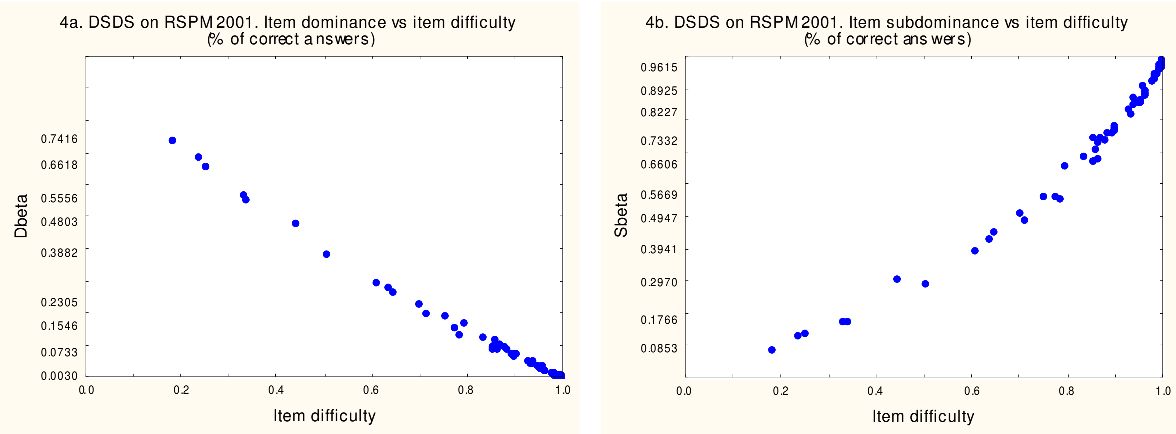

Figure 4. Item dominance level (left panel) and item subdominance level (right panel) vs classical item difficulty level (in terms of % of correct answers).

Left panel of Figure 4 shows item dominance levels to be relatively similar to linear transformation of classical item difficulty values, whereas the subdominance values might contain more new information (Figure 4, right panel) which is the case contrary to the respective relations in subjects (where the correlation of subdominance with traditional indices is higher than that of dominance).

The data was also fitted to the Rasch simple logistic model (SLM), the fit was excellent, and according to the principal component analysis (PCA) of the residuals the data can be considered unidimensional as the first principal component of the residuals accounted for 3.17% of the total variance. Correlation between item dominance and subdominance values and Rasch item thresholds are 0.880 and –0.960 respectively (see Figure 5) – note that these correlations closely repeat these with percentage of correct answers in students.

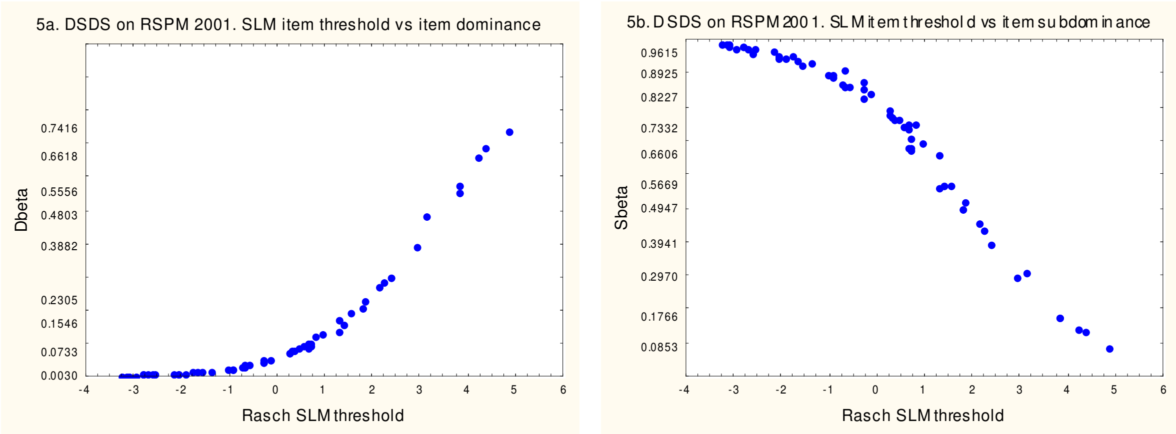

Figure 5. Item dominance level (left panel) and item subdominance level (right panel) vs Rasch SLM item thresholds.

Differences across grades and sexes in subjects' dominance and subdominance levels

The data was divided into 14 groups consisting of either boys or girls of x^{th} grade, year y. Factorial ANOVA revealed main as well as interaction effects of sex and grade on subjects' dominance levels in both 2001 and 2003.

The significances for the main effect of sex and grade in 2001 were $F(1, 2728) = 5.985$, $p = 0.0145$ and $F(3, 2728) = 138.13$, $p = 0.000$ respectively, and for the interaction effect $F(3, 2728) = 10.365$, $p = 0.000$. In 2003 the magnitude of the significances with all $p = 0.000$ was comparable with $F(1, 905) = 14.527$ and $F(2, 905) = 50.768$ for the main effect of sex and grade respectively, and $F(2, 905) = 13.682$ for the interaction effect (see Figure 6).

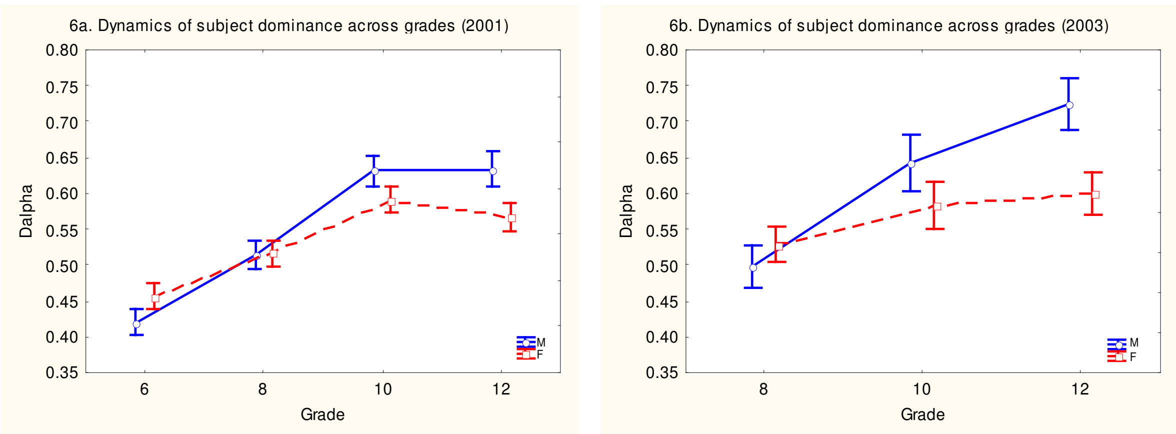

Figure 6. Differences in subjects' dominance levels in 2001 (left panel) and 2003 (right panel) across grades and sexes. Vertical bars denote 95% confidence intervals. M denotes boys, F denotes girls.

For subjects' subdominance levels, the main effect of grade was significant in both years with F(3, 2728) = 72.865 in 2001 and F(2, 905) = 23.413 in 2003, both p = 0.000. The interaction effect of grade with sex was significant as well with F(3, 2728) = 4.622, p = 0.003 in 2001 and F(2, 905) = 4.1026, p = 0.0168 in 2003.

The overall effect of sex was insignificant in both years (F(1, 2728) = 2.88, p = 0.09 in 2001 and F(1, 905) = 0.12, p = 0.73 in 2003) (see Figure 7).

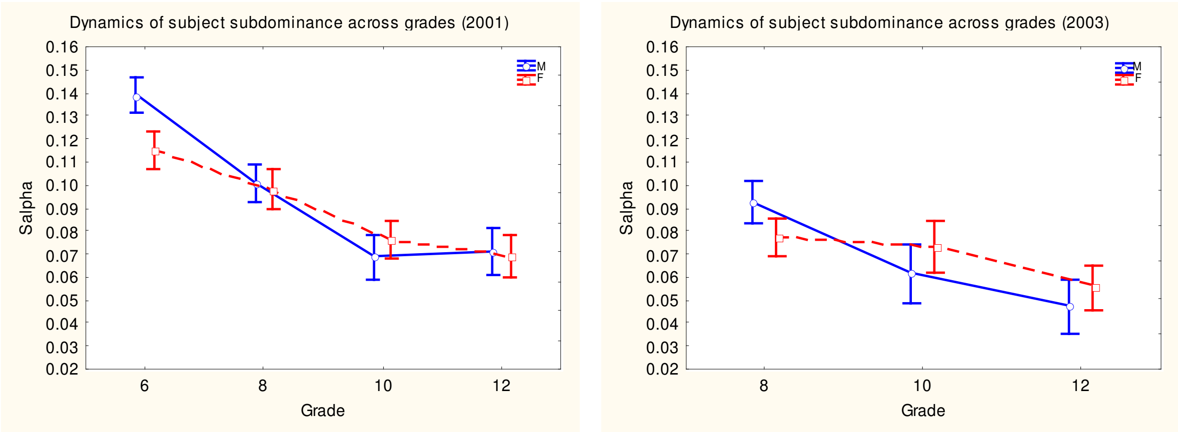

Figure 7. Differences in subjects' subdominance levels in 2001 (left panel) and 2003 (right panel) across grades and sexes. Vertical bars denote 95% confidence intervals. M denotes boys, F denotes girls.

The grade-wise comparisons between the means of dominance and subdominance levels of two sex-groups were performed with t-tests. Cross-sex differences in variances were compared with the Levene and Brown-Forsythe tests for homogeneity of variances.

In 2001, significant differences between boys' and girls' dominance and subdominance levels in 6^{th} grade emerged with [t(766) = -3.31, p = 0.001] and [t(766) = 3.20, p = 0.001] respectively.

In the 8^{th} grade of 2003 the boys were significantly more subdominant [t(395) = 1.978, p = 0.049].

In 10^{th} grades of both 2001 and 2003 there were significant cross-sex differences in dominance levels (with boys being more dominant in both years) with respective significances being [t(679) = 2.67, p = 0.008] and [t(214) = 2.20, p = 0.03].

Also, in 2001, girls' variance in subdominance was significantly higher.

Similarly, in 12^{th} grade of both 2001 and 2003, the dominance levels of boys were significantly higher than these of girls, [t(541) = -4.05, p = 0.0001] and [t(296) = 5.27, p < 0.0001], respectively.

For the total sample, in 2001 there appeared significant cross-sex differences in subdominance levels [$t(2734) = 3.04$, $p = 0.002$] as well as variances of subdominance levels ($F = 8.04$, $p<0.005$) and variances of dominance levels ($F = 14.40$, $p < 0.001$).
In the total sample of 2003, the groups differed significantly only in dominance levels [$t(909) = 2.55$, $p = 0.011$].
The dynamics of distribution of dominance and subdominance levels across grades can be seen from Figures 9 to 11 in Appendix 4.

Nevertheless, it would be unfounded to conclude anything solid about the cross-sex differences in students' cognitive abilities yet when differential item functioning (DIF) has not been excluded or accounted for. Hambleton et al. (1991) have provided the classic definition of DIF: *"An item shows DIF if individuals having the same ability, but from different groups, do not have the same probability of getting the item right."*
In the present analysis, Rasch measurement technique is used to assess dimensionality and DIF.

Dimensionality of and differential item functioning (DIF) in RSMP assessed by Rasch measurement

Method

In RUMM2020, two types of DIF can be identified – uniform and non-uniform DIF. With the former, there is a constant difference between groups in the probability of affirming an item (or category) across the trait (ANOVA main effect) and with the latter the difference varies across the trait (ANOVA interaction effect). The statistical test used for detecting DIF in RUMM is an ANOVA of the person-item deviation residuals with person factors and class intervals as factors (Tennant, 2004).

The fit of the data to SLM for the whole sample for both years was looked at. Otherwise the data were analyzed in 7 groups (all the grade-groups of both years separately).
The data of all the groups were fitted to the Rasch SLM using RUMM2020; the significance level of 0.05 was Bonferroni-corrected to 0.00083 (= $0.05/n$, where n is the number of items, 60). Weighted maximum likelihood method was used to assess person parameters. Pairwise conditional estimation procedure (Chopin, 1983; Zwinderman,

1995) was used to assess item parameters. Number of class intervals was 10 and residual criterion was 2.5 (Bonferroni corrected value 3.69, based on the formula: Bonferroni critical value = standard value + 2*(log(n)/log(1000)) , where n is the number of items). The power of test-of-fit and – since it has been implied (Tennant, 2004) that the fit statistics are not sensitive enough – principal component analyses (PCA) of the residuals were performed to assess dimensionality. (According to Tennant (2004) the variance of residuals in a pattern should be less than 30% to be acceptable). In all the seven grade-groups item characteristic curves for each item were splitted across sexes to screen for DIF for boys and girls.

Results of Rasch measurement of RSPM data

As stated previously, the data of whole sample of 2001 was unidimensional with the overall power of test-of-fit being excellent. Differently, the data of whole sample of 2003 did not fit the model at first, but after discarding the item 4 as extreme the overall power of test-of-fit was good and PCA of the residuals showed the first principal component to account for 3.2% of the total variance.

The fit of data of 6th grade, 2001 (n=768) to the model was excellent; PCA of the residuals showed the first principal component to account for 4.14% of the total variance. Unidimensionality can be concluded. None of the items displayed DIF.
In item 52 there was a tendency towards misfit ($p<0.01$). Items 46, 19, 50, 60, 53, 55 and 47 showed significant misfit ($p<0.0008$).

The data of 8th grade, 2001 (n=744; valid scores 741) did not fit the model at first, - items 3, 19, 28 and 54 had to be discarded as extreme. After that the fit was good. According to PCA of the residuals the first component accounted for 3.92% of the variance. Also, items 12 and 28 showed a tendency towards non-uniform DIF (at significance levels 0.008 and 0.0038 respectively) what according to Svend Kreinder (communication at the Rasch Methodological Workshop, September 2004) might imply multidimensionality. Items 24 and 34 showed a tendency towards uniform DIF, with item 24 rather favoring boys and item 34 rather favoring girls, (both at significance level of 0.004).
Item 60 showed uniform DIF (p=0.00056), by favoring boys.

Items 47, 58, 36, 37, 12, 53, 49 showed a tendency towards misfit ($p<0.01$), whereas the misfit was significant ($p<0.0008$) in items 54, 19, 28 and 3.

The data of 8th grade, 2003 ($n=397$; valid scores 393) did not fit the model, but after discarding items 4 and 6, the fit was good, PCA of the residuals showed that first component accounted for 4.49% of the variance. None of the items displayed DIF but there was a tendency towards items 19 and 20 to misfit the model ($p<0.01$). In items 53, 52, 5 and 12 the infit was significant ($p<0.0008$).

The data of 10th grade, 2001 ($n=682$; valid scores 664) fitted the model: overall power of test-of-fit was good and the first principal component of the residuals accounted for 4.61% of the variance. Unidimensionality can be concluded.
Items 7 and 40 showed a strong tendency towards total item DIF, which, according to the RUMM2020 manual (RUMM Laboratory, 2004) provides an overall estimate of the presence of DIF together with its Main Effect component [*DIF related to the person-factor*] and the Interaction value [*interaction of person-factor with Class-intervals*]. The tendencies' significance levels were $p=0.00095$ and $p=0.00098$ respectively. Item 60 showed significant uniform DIF in favor of boys ($p=0.000065$), as well as total item DIF ($p=0.000006$).
In items 35, 33, 53, 19, 36 and 37 there was a tendency towards misfit ($p<0.01$).
The data of 10th grade, 2003 ($n=216$; valid scores 205) did not fit the model before items 4, 7, 14, 16, 26 and 38 were discarded, after which the overall power of test-of-fit was good. The first principal component of the residuals accounted for 4.98% of the variance. None of the items displayed DIF but there was a tendency towards items 36, 37 and 6 not fitting the model ($p<0.01$).
The data of 12th grade, 2001 ($n=543$; valid scores 530) overall power of test-of-fit was good; the first principal component of the residuals accounted for 4. 31% of the variance. Unidimensionality can be concluded. In items 36 and 37 there was a significant uniform DIF ($p=0.0005$) favoring boys. In item 52 there was a strong tendency towards uniform DIF ($p=0.00091$), Item 20 showed significant nonuniform ($p=0.00002$), as well as total item DIF ($p=0.000005$).
In items 30, 36, 37, 5 and 18 there was a tendency towards misfit ($p<0.01$). In item 54 the misfit was significant ($p<0.0008$).

The data of 12th grade, 2003 did not fit the model before items 4, 6, 13, 14, 15 and 25 were discarded. Then the overall power of test-of-fit was good; first principal component of the residuals accounted for 6.24%. In items 21 and 59 there was uniform DIF favoring girls in item 21 (p=0.00063) and favoring boys in item 59 (p=0.000003). Also, in items 21 and 59 there was total item DIF at significance levels of 0.0001 and 0.00005 respectively.

Apparently, the RSPM data are not consistently unidimensional. At the same time, multidimensionality appears in some questions, rather than continuously along the whole scale. Also, the dimensionality of RSPM is largely dependent of the particular sample.

Investigation of sample-dependency of DSDS

To give some idea of sample-dependency of DSDS, the sample was split 1) by subjects into high-scorers and low-scorers, and 2) by items into easy and difficult in terms of percent of correct answers. Acquired four subsamples were scaled separately. The results are presented on Figure 8:

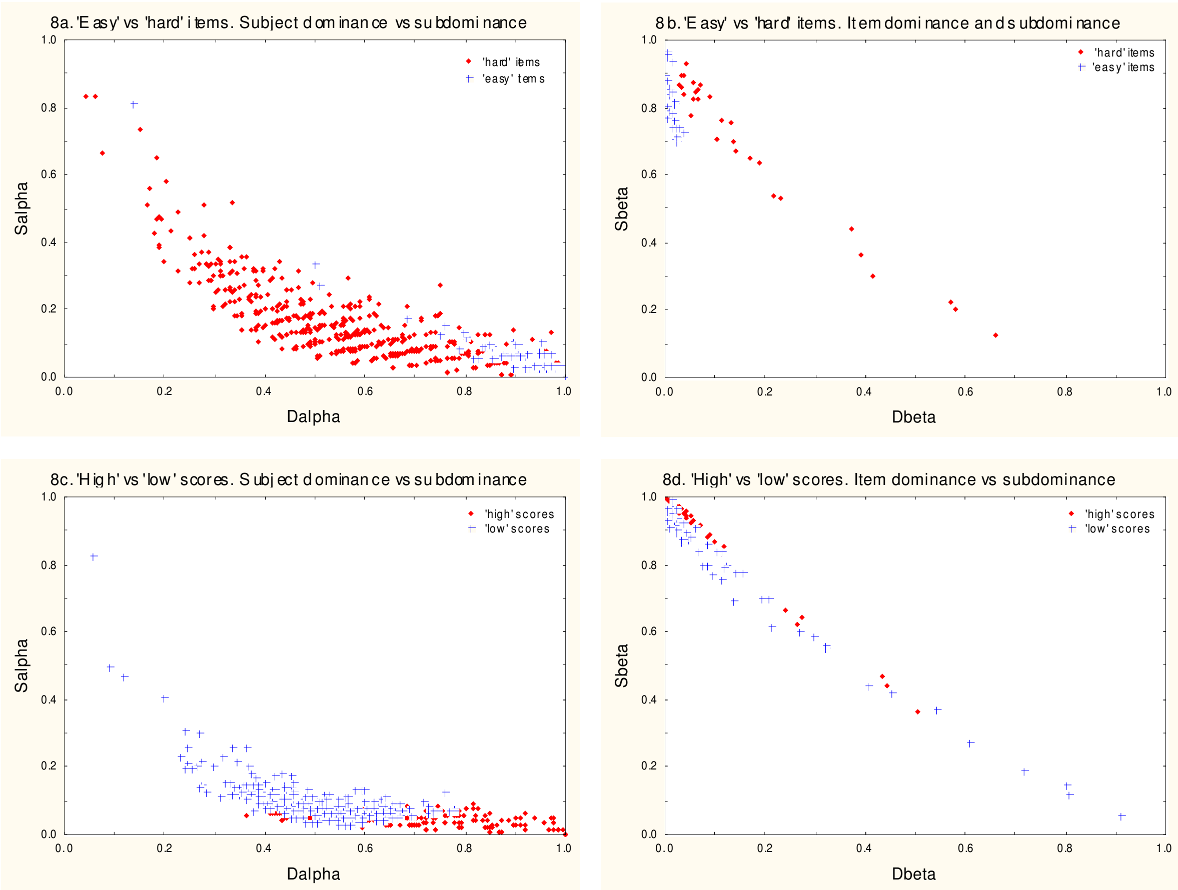

Figure 8. Subject dominance vs subdominance (upper left panel) and item dominance vs subdominance (upper right panel) for samples of easy and hard items.
Subject dominance vs subdominance (lower left panel) and item dominance vs subdominance (lower right panel) for subject-samples of high and low scores.

Figure 8 shows that, except for the items scaled in 'easy' and 'hard' subsamples, the rest of the subsamples are still more or less identifiable as belonging to the same population. Of course, deriving from DSDS algorithm, the particular estimates of and inidividual's dominanance and subdominance levels are still sample dependent.

Discussion

Probably the most valuable asset of DSDS for psychometrics is the possibility to disentangle the relationship between two sets of objects in an original way and without constraints on the data.

At the current state of knowledge the indices are not entirely interpretable, though.

In comparison to the results of DSDS of random data, it could be assumed that certain psychological phenomena are reflected in the relationships of dominance and subdominance values in subjects, whereas the respective relationship in items is more similar to that of random data.

It is interesting that the relationships of dominance and subdominance values with traditional parameters are asymmetrical with respect to subjects and items. In items, it seems that subdominance bears some more new information than item dominance, compared to the traditional indices. But these differences are minor, – the direction of this relation has to be controlled for in other sets of data as well.

In subjects, dominance seems to reflect a more distinct concept than the percentage of correct responses (or Rasch score) of a subject, compared to subdominance level which is highly indicative of a person’s test score (or Rasch score), whereas estimate of dominance conveys some additional information. Even though, as can be seen from Figure 6, the general trends in dominance follow the results of estimates of IQ scores (based on test scores) based on the analysis by Lynn, Allik, Pullmann, and Laidra (2004) of practically the same sample.

This result implies that DSDS presents a technique allowing to factorize students’ abilities as well as hindrances or lack of motivation that the subject was facing during the testing process. The subdominance levels might reflect the responding style of a student – if the subject was motivated and careful or hasty, – or hardships apart from lack of ability to perform well on the test, whereas the dominance level might imply the estimate of the potential of the student. Or, in other words, the dominance might reflect the strength of the strongest link of the chain whereas subdominance might reflect that of the weakest one (which in real life most probably determines the subject’s raw score). The maximal distance of these two strengths is probably limited by some function analogous, for example, to the function linking the coordinates of the points on the hyperbole or ellipse.

By estimating the relationship between dominance and subdominance values in the data matrix, dimensionality and differential item functioning come into play. The present data is roughly unidimensional (according to the fit to Rasch model) in most, but not all subsamples which might seem to contradict the findings of Lynn, Allik and Irwing (2004), who identified three factors in practically the same data-set. This fact can be reconciled by the notion of linear vs nonlinear analysis – in Rasch model, the data are fitted to a nonlinear dimension, which is expected to give different results from linear factor analysis. Also, estimation of dimensionality always contains a degree of arbitrariness in terms of amount of residuals not allowed to be unaccounted for. Not to mention that in real life data, it would be hard to imagine an inherently perfectly unidimensional construct whatsoever since unidimensionality in life reflects the more or less constant covariation of several subdimensions.

The comparative results of DSDS on random vs real data in the present analysis imply that in case of strict unidimensionality (that holds in random data of this analysis), relationship of dominance and subdominance values is linear. In this respect, DSDS could be used as a tool for assessing the deviation from unidimensionality. The respective quantifiers are to be developed, though.

Turning to the DSDS's sample-dependence, it is interesting that DSDS identifies the two split sub-samples as belonging to the same population – except for the case of analysis on item dominance and subdominance levels for 'easy' and 'hard' items. The latter is expected since it is not conceivable to estimate an item-scale based on a biased sub-scale since the lack of variation in items' "test-behaviour".

It remains yet to be investigated, whether the nature of the relationship between row and column objects could be derived for the whole population based on an analysis with a biased sample of subjects.

To conclude with, DSDS presents an interesting method with several unique and valuable properties. To investigate the possibilities of DSDS further, analyses of real data with DSDS provided with comparative methods and generated data of known parameters and dimensionality remain to be performed.

References

Andrich, D. (1988). Rasch models for measurement. Sage university paper series on quantitative measurement in the social sciences. Newberry Park, CA: Sage Publications.

Chopin, B. (1983). *A fully conditional estimation procedures for Rasch model parameters.* Report No 196. Los Angeles: University of California, Graduate School of education Center for the study of Evaluation.

Dzhafarov, E. N. (1999). Double Skew-Dual Scaling: A Conjoing Scaling of Two Sets of Objects Related by a Dominance Matrix. *Journal of Mathematical Psychology*, 43, 483-517.

Fischer, G. H., and Molenaar, I. W. (1995). Rasch models: Foundations, recent developments, and applications. New York: Springer Verlag.

Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage.

Laflin, S. (1988). *Numerical Methods of Linear Algebra.* Chartwell-Bratt (Publishing and Training) Ltd.

Linacre, J. M. (1989). Many-facet Rasch measurement. Chicago: MESA Press.

Lynn, R., Allik, J., Pullmann, H., Laidra, H. (2004). Sex differences on the progressive matrices among adolescents: some data from Estonia. *Personality and Individual Differences*, 36, 1249-1255.

Lynn, R., Allik, J., Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, 32, 411-424.

Nishisato, S. (1980). *Analysis of categorical data: Dual scaling and its applications.* Toronto: Univ. of Toronto Press.

R Development Core Team (2005). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Raven, J. (1981). *Manual for Raven's Progressive Matrices and Mill Hill vocabulary scales.* Oxford: Oxford Psychologists Press.

RUMM Laboratory Pty Ltd. (2004). *Displaying the RUMM2020 Analysis.*

RUMM Laboratory Pty Ltd. (1997-2004). *RUMM2020: Rasch Unidimensional Measurement Models.*

SPSS, Inc. (1989-2002). *SPSS, version 11.5.0.* www.spss.com.

StatSoft, Inc. (2003). *STATISTICA (data analysis software system), version 6.* www.statsoft.com.

Zwinderman, A, H. (1995). Pairwise estimation in the Rasch models. *Applied Psychological Measurement*, 19, 4, 369-375.

Tennant, A. (2004). Lecture: *Practice issues in Rasch analysis.* Academic Unit of Musculoskeletal and Rehabilitation Medicine, The University of Leeds, UK.

Appendix 1

Source code for DSDS program for use with STATISTICA v6 (StatSoft, Inc., 2003).

The algorithm makes use of the power-method extraction of eigenvalues (Laflin, 1988).

```
Sub EigenVectors (A() As Double, iterCount As Long, A_EVal() As Double, A_EVec() As
Double)
                                        '{subroutine for finding eigenvectors of
                                        'a matrix}

Dim X() As Double                       'variable declarations
Dim Y() As Double
Dim K() As Double

Dim I As Integer
Dim j As Integer
Dim iter As Integer
Dim kmax As Double

Dim n As Long

n = UBound(A,1)                 ' n  = the number of rows of A

ReDim X(n)                      ' set the dimensions of arrays
ReDim Y(n)
ReDim K(n)
ReDim A_EVal(n)
ReDim A_EVec(n)

For I = 1 To n                          ' initiation of the auxiliary vectors for
        X(I) = 1.0                      ' computing the eigenvectors
        K(I) = 0.0
Next I

iter = 0                                ‘cycle for computing the eigenvalues ->

start:
        For I = 1 To n
                Y(I) = 0.0
        Next I

                '{compute Y = A*X}
        For I = 1 To n
                For j = 1 To n
                        Y(I) = Y(I) + A(I,j)*X(j)
                Next j
        Next I

                '{estimate eigenvalues}
        For j = 1 To n
                If Abs(X(j)) > 0.0 Then
                        K(j) = Y(j)/X(j)
```

```
                Else
                        K(j) = 0.0
                End If
        Next j

        kmax = 0
        For j = 1 To n
                If Abs(K(j)) > kmax Then
                        kmax = Abs(K(j))
                End If
        Next j

        If kmax < 1.0 Then
                kmax = 1.0
        End If

        For j = 1 To n
                X(j) = Y(j)/kmax
        Next j

        iter = iter + 1

        If (iter<iterCount) Then              ' end of cycle for computing the eigenvalues
                GoTo start
        End If

MatrixCopy(X(),1,1,0,0,A_EVec(),1,1)
MatrixCopy(K(),1,1,0,0,A_EVal(),1,1)
End Sub


Sub NormVector(A() As Double) ' {subroutine for normalizing the vectors, i.e dividing
                                        ' each element by the sum of all elements of
                                        ' the vector}

Dim I As Integer                ' variable declarations
Dim n As Long
Dim Sum As Double

n = UBound(A,1)

Sum = 0
For I = 1 To n
 Sum = Sum + A(I)
Next I

For I = 1 To n
        A(I) = A(I)/(Sum)
Next I

End Sub
```

```
'$include: "*STB.svx" '{main program}
Option Base 1

Sub Main

Dim A() As Double                'variable declarations
Dim I() As Double
Dim B() As Double
Dim AT() As Double
Dim BT() As Double
Dim X() As Double
Dim Y() As Double
Dim Z() As Double
Dim U() As Double

Dim X_EVal() As Double
Dim Y_EVal() As Double
Dim Z_EVal() As Double
Dim U_EVal() As Double

Dim X_EVec() As Double
Dim Y_EVec() As Double
Dim Z_EVec() As Double
Dim U_Evec() As Double

Dim detX As Double
Dim detZ As Double
Dim detY As Double
Dim detU As Double

Dim Dalpha() As Double 'nm X m1
Dim Dbeta() As Double 'mn X n1
Dim Salpha() As Double 'nm X m1
Dim Sbeta() As Double 'mn X n1

Dim j As Integer
Dim K As Integer

Set theData = Spreadsheets.Open ("X:\...\*.*")   ' read the data from STATISTICA or Excel
                                                 ' spreadsheet

n = theData.NumberOfCases                        ' number of cases in the specified spreadsheet
m = theData.NumberOfVariables                    ' number of variables in the specified spreadsheet

A = theData.Data                 'copy data from spreadsheet to matrix A

ReDim I(n,m)            'set the dimensions of arrays
ReDim B(n,m)
ReDim AT(m,n)
ReDim BT(m,n)

ReDim X(n,n)
ReDim Y(m,m)
ReDim Z(n,n)
ReDim U(m,m)
```

```
ReDim X_EVal(n)
ReDim Y_EVal(m)
ReDim Z_EVal(n)
ReDim U_EVal(m)

ReDim X_EVec(n,1)
ReDim Y_EVec(m,1)
ReDim Z_EVec(n,1)
ReDim U_Evec(m,1)


MatrixFill(1,I,1,1,0,0) ' initiate all values in matrix I(n,m) to 1
MatrixSubtract(I, A, B) ' B = 1 – A
MatrixTranspose(A,AT) ' transpose matrix A
MatrixTranspose(B,BT) ' transpose matrix B

MatrixMultiply(A,BT,X)
MatrixMultiply(BT,A,Y)
MatrixMultiply(B,AT,Z)
MatrixMultiply(AT,B,U)

MatrixDet(X, detX)
MatrixDet(Y, detY)
MatrixDet(Z, detZ)
MatrixDet(U, detU)

EigenVectors X,30,X_EVal,X_Evec              'compute the eigenvectors
EigenVectors Y,30,Y_EVal,Y_EVec              'in 30 iterative steps
EigenVectors Z,30,Z_EVal,Z_EVec
EigenVectors U,30,U_EVal,U_Evec

NormVector(X_EVec) 'Valpha
NormVector(Y_EVec) 'Vbeta
NormVector(Z_EVec) 'Walpha
NormVector(U_Evec) 'Wbeta


ReDim Dalpha(n,1)                            'set the dimensions of arrays
ReDim Dbeta(m,1)
ReDim Salpha(n,1)
ReDim Sbeta(m,1)

MatrixMultiply(A,Y_EVec,Dalpha)
MatrixMultiply(BT,X_EVec,Dbeta)
MatrixMultiply(B,U_Evec,Salpha)
MatrixMultiply(AT,Z_EVec,Sbeta)

MatrixDisplay(Dalpha, "Dalpha")
MatrixDisplay(Dbeta, "Dbeta")
MatrixDisplay(Salpha, "Salpha")
MatrixDisplay(Sbeta, "Sbeta")

End Sub
```

Appendix 2

Source code for DSDS program for use with R 2.1.0 (R Development Core Team, 2005). I am indebted to Kenn Konstabel for the help with this implementation.

```
# {A is a square matrix that is filled with n*n random
# numbers of set {0; 1}, where n is the dimension of the matrix}
# B is a square matrix of equal dimension to A, all entries B[i,j] equal to (1 -  A[i,j])

A<-matrix(sample(c(0,1),n*n,replace=TRUE),n,n)     # {the desired size of the matrix
                                                   # should be substituted for n}
A<-as.matrix(A)                                    # error-check
B<-matrix(1, dim(A)[1], dim(A)[2]) - A

ABT <- A %*% t(B)          # ABT = A x B^T, where B^T is the matrix B transposed
BTA <- t(B) %*% A          # BTA = B^T x A
BAT <- B %*% t(A)          # BAT = B x A^T, where A^T is the matrix A transposed
ATB <- t(A) %*% B          # ATB = A^T x B

evs <- Mod(eigen(ABT)$values)
eABT<- abs(Mod(eigen(ABT)$vectors[,1]))  # eABT = eigenvectors of ABT
eBTA<- abs(Mod(eigen(BTA)$vectors[,1]))  # eBTA = eigenvectors of BTA
eBAT<- abs(Mod(eigen(BAT)$vectors[,1]))  # eBAT = eigenvectors of BAT
eATB<- abs(Mod(eigen(ATB)$vectors[,1]))  # eATB = eigenvectors of ATB

norm<-function(x) x/sum(x)    # {define a function for normalizing a vector,
                              # i.e dividing each element by the sum of the elements}

Valpha<-norm(eABT)         # Valpha = normalized eATB
Vbeta <-norm(eBTA)         # Vbeta = normalized eBTA

Dalpha<- A %*% Vbeta       # Dalpha = A x Vbeta   (= dominance vector for row objects)
Dbeta <- t(B) %*% Valpha   # Dbeta = B^T x Valpha   (= dominance vector for column objects)

Walpha<-norm(eBAT)         # Walpha = normalized eBAT
Wbeta <-norm(eATB)         # Wbeta = normalized eATB
Salpha<- B %*% Wbeta       # Salpha = B x Wbeta   (= subdominance vector for row objects)
Sbeta <- t(A) %*% Walpha   # Sbeta = B^T x Walpha   (= subdominance vector for column
                           # objects)
```

Appendix 3

Abbreviations

DIF	differential item functioning; [an item shows DIF if individuals having the same ability, but from different groups, do not have the same probability of getting the item right (Hambleton et al., 1991)].
DSDS	double skew-dual scaling: a conjoint scaling of two sets of objects related by a dominance matrix (Dzhafarov, 1999).
PCA	principal components analysis
R	a language and environment for statistical computing (R Development Core Team, 2005).
RSPM	Raven’s Standardized Progressive Matrices (Raven, 1981).
RUMM	Rasch Unidimensional Measurement Models (RUMM Laboratory Pty Ltd, 1997-2004).
SLM	Rasch simple logistic model (Rasch, 1960).

Appendix 4

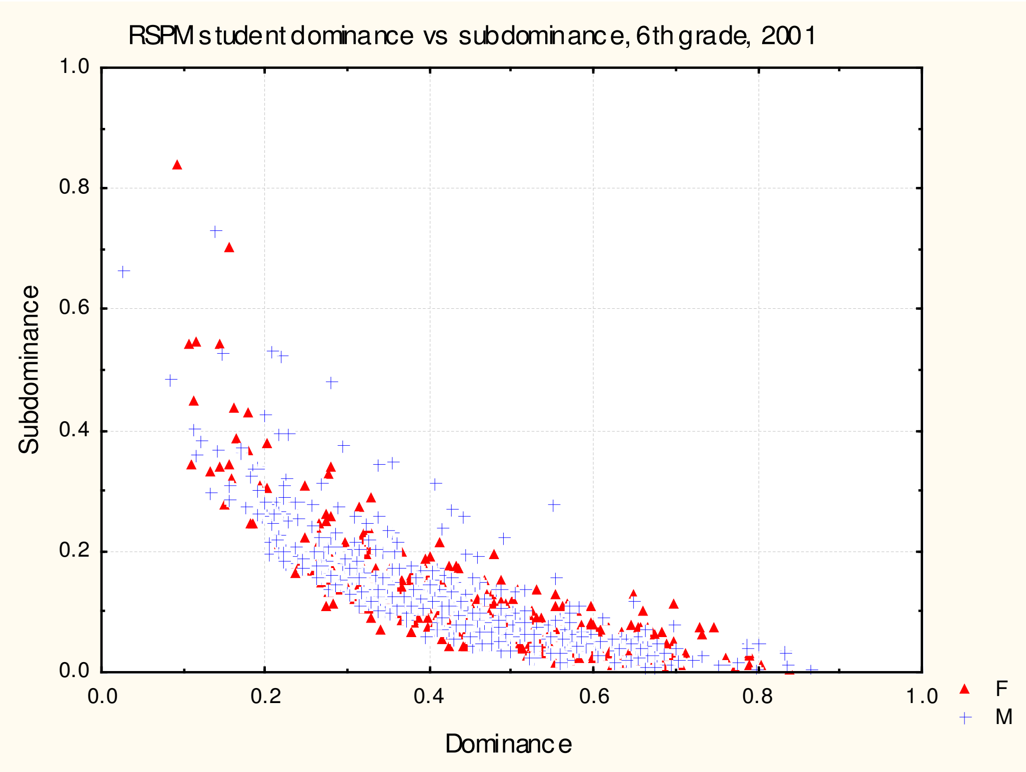

Figure 9. Student dominance and subdominance levels for 6th grade, 2001. S_F – subdominance for girls; S_M – subdominance for boys. Throughout all the Figures: F stands for girls; M stands for boys.

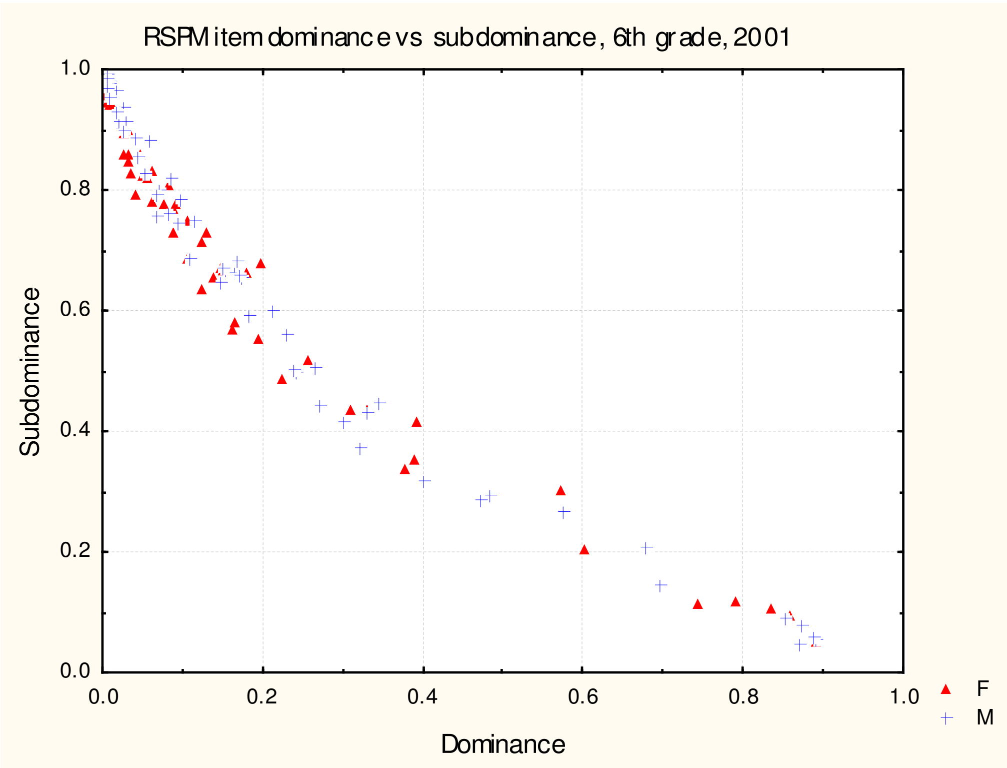

Figure 10. Item dominance and subdominance levels for 6th grade, 2001

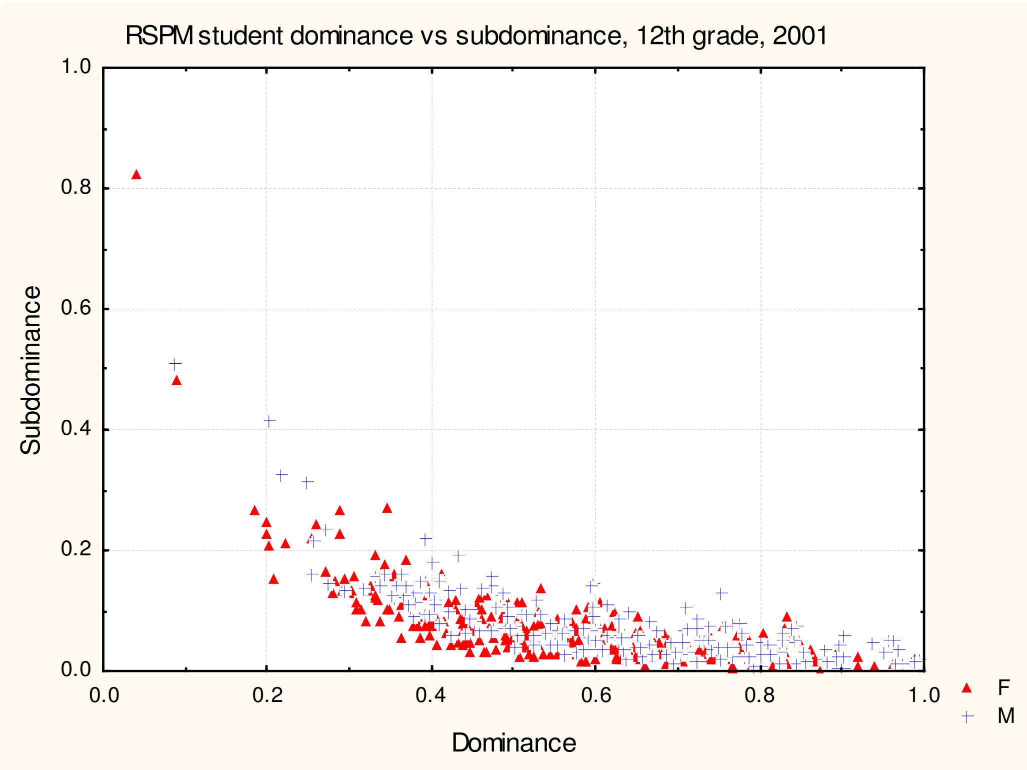

Figure 11. Student dominance and subdominance levels for 12th grade, 2001.

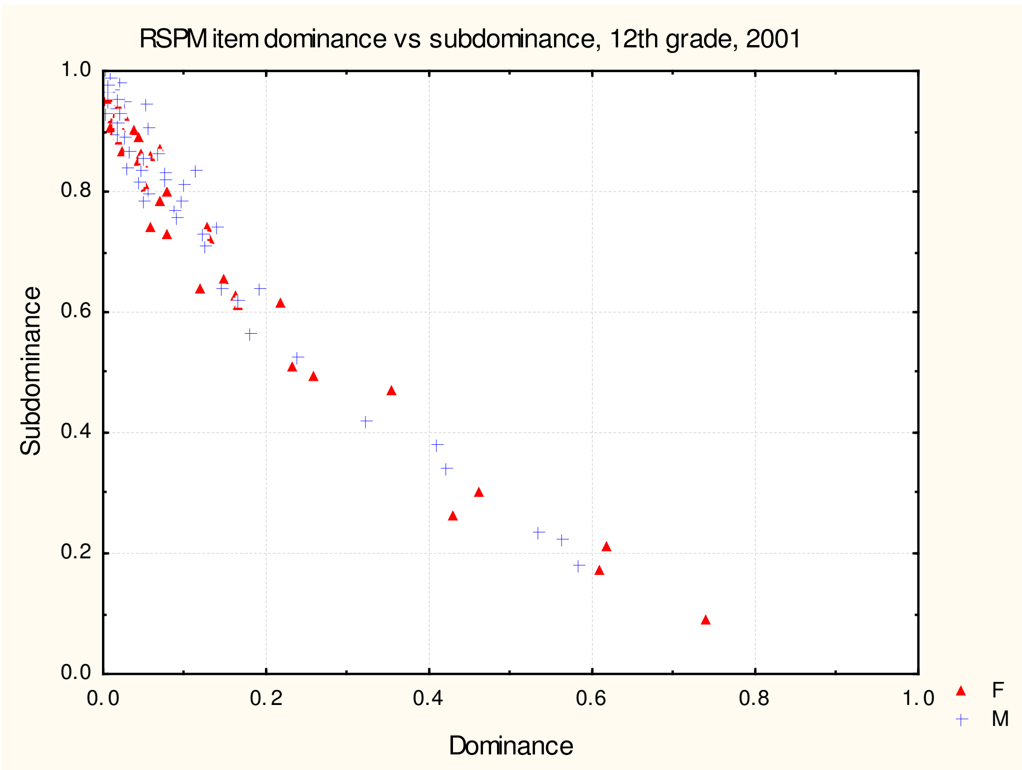

Figure 12. Item dominance and subdominance levels for 12th grade, 2001

Acknowledgements

I am grateful to my supervisors Aavo Luuk and Jüri Allik for providing inspiration and opportunities and showing the way; to Ehtibar Dzhafarov for teaching and support; Kenn Konstabel for discussion and help with R; Olev Must for good advice; Nele Kuldkepp and Carolina Murd for keeping up the spirits at night; Marika Paaver and Liisa Past for commenting on the draft; and, last but not least, all my dear colleagues on Tiigi Street not mentioned yet, for the atmosphere.