JOURNAL OF APPLIED QUANTITATIVE METHODS

The International Conference
"Innovation and Society 2011. Statistical Methods
for the Evaluation of Services (IES 2011)"

IMPACT OF EDUCATIONAL TEST FEATURES ON ITEM DIFFICULTIES BY THE LINEAR LOGISTIC TEST MODEL

Daniela MARELLA[1]
Researcher of Statistics,
Department of Educational Sciences,
University Roma Tre, Italy.

E-mail: dmarella@uniroma3.it

Carlo DI CHIACCHIO[2]
PhD, Data Analyst,
National Institute for the Evaluation
of the Educational System on the Project
for the International Student Assessment (PISA),
Italy

National Project Manager for PISA 2012,
Italy

E-mail: carlo.dichiacchio@invalsi.it

Giuseppe BOVE[3]
PhD, Professor of Statistics,
Department of Educational Sciences,
University Roma Tre, Italy

E-mail: bove@uniroma3.it

Abstract: *The aim of the paper is to investigate the effect of item and person properties on item difficulties using the Linear Logistic Test Model (LLTM) and its extensions. The data under investigation are the Italy mathematics data from the Program for International Student Assessment (PISA) 2006. The information regarding the geographical macro-area (North-Italy versus South-Italy) has been used in the application as person property. Furthermore, the comparison on item properties effects between North-Italy versus South-Italy is performed fitting the LLTM for each geographical macro-area.*

Key words: *Rasch Model, LLTM, Latent Regression LLTM*

1. Introduction

Educational testing studies focus on latent variables, usually named abilities. Remarkable examples are reading ability or mathematical ability. A primary goal of these studies is how much of such abilities persons possess and to this aim a test consisting of a number of items (questions) is developed. Item response theory (IRT) is essentially a theory of the relation between item responses and the underlying abilities (or traits). Mathematically, the relation is described by a function (named item characteristic curve, ICC) linking the probability of correct response to an item and the ability scale. Since test items are not necessarily equivalent in difficulty or validity in measuring the underlying trait, item parameters are included in ICC. When only one difficulty parameter for each item is considered and a logistic model is adopted for the ICC we obtain the famous Rasch model, whose attractive theoretical properties have been extensively studied (e.g. Fischer and Molenaar (1995)).

Formally, the Rasch model for dichotomously scored responses defines the probability of a correct answer to the *i*-th item (for $i = 1,...,I$) by examinee *p* as follows

$$\pi_{pi} = \frac{\exp(\theta_p - \beta_i)}{1 + \exp(\theta_p - \beta_i)} \qquad (1)$$

where β_i and θ_p are the item difficulty parameter and the person ability parameter, respectively. Equation (1) in logit form becomes

$$\eta_{pi} = \theta_p - \beta_i \qquad (2)$$

where $\eta_{pi} = \log(\pi_{pi}/(1-\pi_{pi}))$ is the logit link function. Person properties and/or item properties can be included in the model to explain person and item effects, respectively. The models we will consider in this paper can be conceived as specific instantiations of generalized linear mixed models with random intercept and fixed slopes (De Boeck and Wilson (2004)). The intercept θ_p is normally distributed with mean zero and variance σ_θ^2.

In order to explain the differences between students with respect to mathematical ability, person properties can be included in the Rasch model (2) as predictors obtaining the latent regression Rasch model. This model described by Zwinderman (1991) is especially helpful if subpopulations are represented in the sample. Formally, let M be the person predictor then the logit form of the latent regression Rasch model is

$$\eta_{pi} = \theta_M M_p + \varepsilon_p - \beta_i \qquad (3)$$

where M_p is the value of person *p* on person property, θ_M is the regression coefficient of person property M and ε_p follows a normal distribution with mean zero and variance σ_ε^2.

On the other hand, if the aim is to explain the differences between items in terms of item properties, a special case of the Rasch model named LLTM can be used. LLTM proposed by Fischer (1973) is a Rasch-family model that includes parameters for the impact of cognitive design variables and other test variables on item difficulty. Formally, LLTM breaks down the item difficulty parameter β_i in (2) into a linear combination of certain hypothesized elementary parameters as follows

$$\beta_i = \sum_{j=1}^{k} w_{ij}\alpha_j \qquad (4)$$

for $i = 1,...,I$, where α_j $(j = 1,...,k)$ are the so-called basic parameters representing the difficulty parameter for the item property *j* and w_{ij} are fixed and known weights. More specifically, $w_{ij} = 1$ if item *i* possesses the item property *j* and $w_{ij} = 0$ otherwise.In general, the weights can be constructed using all possible numbers, but in this case only dichotomous values have been assigned.The main advantage of LLTM model is that only $k(k < I)$ parameters needed to be estimated instead of *I*. The logit expression of the LLTM is given by

$$\eta_{pi} = \theta_p - \sum_{j=1}^{k} w_{ij}\alpha_j \qquad (5)$$

The idea motivating the LLTM was that the item parameters can be explained in terms of underlying cognitive operations involved in the solution process. In applications of this kind the basic parameters represent the difficulty of certain cognitive operations.

Clearly, to be applied the LLTM model requires the full specification of the weights w_{ij}. As shown in Baker (1993), misspecification of this matrix may lead to systematic errors in the estimates of α_j and furthermore be responsible for the misfit that is frequently observed in the applications of the LLTM (Fischer and Molenaar(1995)). This implies, the necessity of careful validation of the cognitive model that underlies an LLTM application.

As stressed in Kubinger (2008), there are several potential applications of the LLTM, all deal with measuring certain item administration effects (warming-up effects, effects of different item response formats, position effects of item presentation, and so on). Then the LLTM can also be used for measuring the effects of experimental condition of the test situation on item difficulty.

Combining (3) and (4) yields the latent regression LLTM

$$\eta_{pi} = \theta_M M_p + \varepsilon_p - \sum_{j=1}^{k} w_{ij}\alpha_j \qquad (6)$$

which take into account two parts: person contribution and item contribution. The LLTM (or equivalently the latent regression LLTM) also allows for interactions between the item properties, for instance if one is interested in the interaction between two item properties, their product can be added as additional term in (5) and (6), respectively.

The LLTM model is almost always rejected since it requires that the item difficulty can be perfectly predicted by the item properties (item parameters are regressed on item properties). More specifically, in (6) the person contribution has an error term ε_p while the item contribution does not include an error term. An interesting extension of this model is the LLTM plus error (Janssen et al. (1994)), which means that an error term is added in (5)

$$\eta_{pi} = \theta_p - \sum_{j=1}^{k} w_{ij}\alpha_j + \varepsilon_i \qquad (7)$$

where ε_i follows a normal distribution with mean zero and variance σ_ε^2. As a consequence, the latent LLTM plus error is given by

$$\eta_{pi} = \theta_M M_p + \varepsilon_p - \sum_{j=1}^{k} w_{ij}\alpha_j + \varepsilon_i \qquad (8)\,.$$

The aim of the paper is to investigate the effects of person and item properties on the item difficulties using the mathematics data of PISA 2006, reviewing recent developments and potentialities of the LLTM. The information regarding the geographical macro-area (North-Italy versus South-Italy) has been used in the application as person property. Furthermore, the comparison on item properties effects between North-Italy versus South-Italy is performed fitting the LLTM for each geographical macro-area.The paper is organized as follows. In Section 2 the mathematics data from PISA 2006 are described. In Section 3 the models described in Section 1 are applied to such data.

2. PISA 2006 Database

The data under investigation are from the Program for International Student Assessment (PISA) 2006, and we focused our study on Mathematics items (Science was the major domain in PISA 2006). The original PISA 2006 sample for Italy was partitioned in two subgroups: North-Italy and South-Italy. The North-Italy group is formed by the areas North-West and North-East while the South-Italy group is formed by the areas Center, South, South and Islands. A random sample of 1000 students was drawn from each macro-area.

In PISA 2006 were used 48 link items from the previous PISA 2003 cycle, where Mathematics was the major domain. For items construction, in PISA are followed both methodological and theoretical considerations (see OECD 2003 (2003),OECD 2006 (2006)).

Students' responses to PISA items are either dichotomously or polytomously scored. In dichotomous items correct responses are assigned score 1, while wrong answer as well as omitted responses are assigned score 0.

In polytomous items responses are graded respect their level of correctness. This means that a student's response, following the scoring guide for that particular question, will be regarded as completely right, partially right or incorrect. Completely right answers are assigned the highest score (usually score 2),partially correct answers are assigned a lower score (usually score 1) and wrong answers follow the same procedure of dichotomous items.

In PISA 2006 Math link items four items out of 48 are polytomously scored. These items were recoded dichotomously assigning score 1 to both completely and partially

corrected responses. Again, score 0 was assigned to incorrect responses and omitted responses as well.

PISA items are not curricular, they are intended to measure literacies providing students with situations and contexts as realistic as possible. It is possible to define situations as ``part of the student's world in which the tasks are placed. [they are] located at a certain distance from the students" (see OECD 2003 (2003), p. 81). For Mathematics, four situations have been considered: *personal(PE)*, *educational (ED)*,*public (PU)* and *scientific(SC)*.

PISA 2006 follows the Mathematics theoretical framework extensively developed in PISA 2003. Within the framework the theoretical dimensions and definitions of what Mathematical literacy is are clearly elaborated. Besides situations, other two components need to take into account: the mathematical content and the processes. The mathematical content refers to the main theme that a problem present to people as to be solved. Four content categories are identified: *space and shape(S)*, *change and relationship(C)*, *quantity (Q)* and *uncertainty(U)*. These categories are described as overarching areas and represent the reporting subscales. The processes refer to the mathematical competencies students bring into play when trying to solve a specific problem. Three groups of competencies are hypothesized, each involving different cognitive processes in varying levels during problem resolution: *reproduction(REP)*, *connection (CON)*and *reflection(REF)*.

Finally, different items response format are used in PISA in order to better assess the various aspects of Mathematical literacy. Broadly speaking they involve closed versus open response format. In the present study we considered the more detailed classification, that is: *closed constructed response(CCR)*, *complex multiple choice(CMC)*, *multiple choice(MC)*, *open constructed response(OCR)*and*short response (SR)*. Items distribution for each of the aspects described above are illustrated in Table 1.

Table 1. Distribution of items by the domains

Item Format	**Number of Items**
CCR	6
CMC	9
MC	12
OCR	11
SR	10
Context/Situation	**Number of Items**
PE	9
ED	8
PU	18
SC	13
Content	**Number of Items**
Q	13
S	11
C	13
U	11
Competency	**Number of Items**
CON	24
REP	11
REF	13

The framework dimensions item format, context, content and competency structure the conceptual grid underpinning items construction (OECD 2003 (2003)).

The LLTMweights w_{ij} were built relying on that conceptual grid: each item was assigned code 1 if it had a particular feature, otherwise code 0.

3. Application of LLTM model to the analysis of item difficulty sources

The analysis has been carried out by using lme4, an R packages for item response modeling(R Development Core Team (2009)).At first the Rasch model was fit to the dataset. The estimated person variance is 1.56. The standard deviation of the variance estimate is 1.25. The estimated item parameters vary from -3.36 to 3.85 with an average of 0.43.

In order to check the fit of the Rasch model a parametric Bootstrap goodness of fit test using Pearson's χ^2 statistic has been used, based on 200 data-sets.The non significant p-value is approximately equal to one then the Rasch model fits the data. Analogously, to assess the goodness of fit of test items a test based on χ^2 statistic has been used(see Reise (1990)). All the test items fit the Rasch model, the non significant p-values are approximately equal to one.

Table 2 reports the goodness of fit indices (AIC, BIC, Deviance) for the models described in Section 1 in order to compare their goodness of fit. Lower values of these indices indicates a better fit. One person property has been used in the analysis: the macro-area. A dummy coding is used, with 1 for North-Italy and 0 for South-Italy. From Table 2 it can be noted that the latent regression Rasch model (3) has a better fit than the Rasch model. Based on likelihood ratio test (LRT), the difference is significant ($\chi^2(1) = 165$, $p < 0.001$) meaning that the goodness of fit of the Rasch model is lower.

As expected, since the macro-area information explain part of the original person variance the estimated person variance is lower than the one estimated with the Rasch model. More specifically, its value is 1.40 with a standard error of 1.18. Furthermore, the estimated macro-area effect is 0.79 with a standard error of 0.06 and the effect is highly statistically significant ($p < 0.001$). This indicates that North-Italy students are more inclined to mathematics than South-Italy student overall.

Table 2. Goodness of fit

Model	No. of Parameters	AIC	BIC	Deviance
Rasch	**49**	**39610**	**40028**	**39512**
Latent Regression Rasch	**50**	**39447**	**39874**	**39347**
LLTM	**14**	**43090**	**43210**	**43062**
LLTM plus error	**15**	**39809**	**39937**	**39779**
Latent Regression LLTM plus error	**16**	**39646**	**39783**	**39614**
LLTM plus error with interactions	**27**	**39803**	**40033**	**39749**
Latent Regression LLTM plus error with interactions	**28**	**39639**	**39879**	**39583**

Since the Rasch model fits the data, we may go on further and compare it with more restrictive models.The third model we have considered is the LLTM in equation (5). In the LLTM the item properties in Table 1 are used to explain the differences between items in terms of the effect they have on π_{pi}. Clearly, instead of estimating individual item effects, the effects of item properties are estimated. The fit of LLTM has been tested in two ways. Firstly, a graphical model check comparing the estimated item parameters of the Rasch model β^{RM} to the parameters reproduced by the LLTM parameter estimates β^{LLTM} has been used. The logic of this method is that if the LLTM fits well, thepoints with coordinates representing estimates of item difficulty with the LLTM and the Rasch models, respectively,

should scatter around a 45° line through the origin. In Figure 1 we graphically plot β^{RM} against β^{LLTM}.

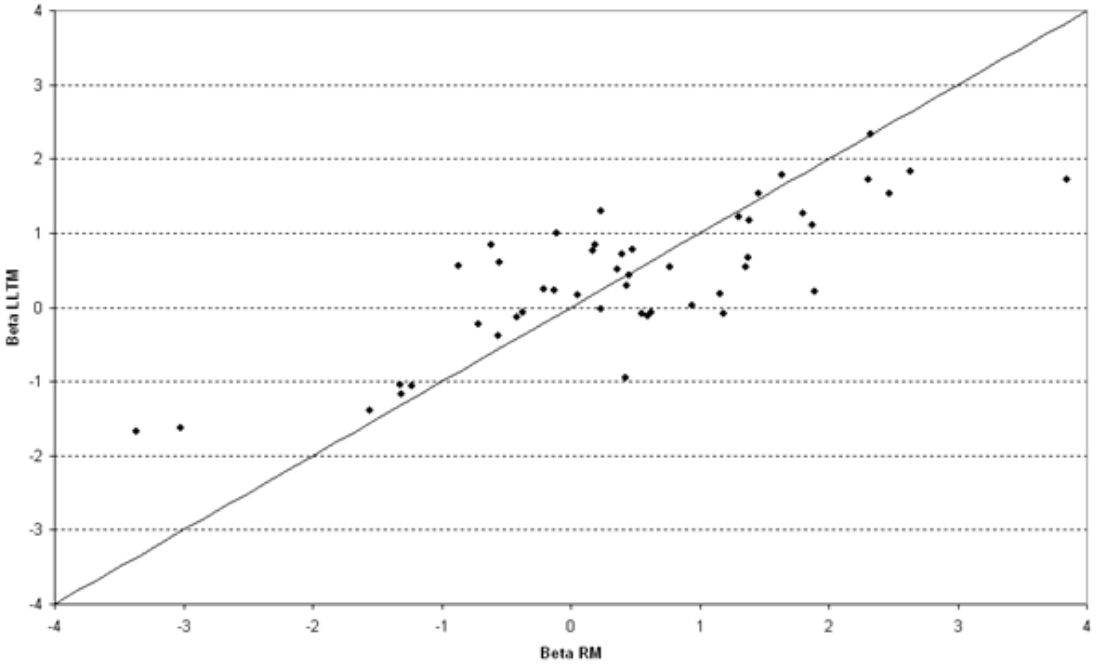

Figure 1. Graphical Goodness-of-Fit agreement between Rasch Items Difficulties and their Linear Logistic Test Model (LLTM)

An inspection of the graphical model test in Figure 1 discloses a good accordance between these item parameters estimates. The correlation between the Rasch and LLTM item difficulties estimates was 0.8. Such a correlation seems suggest a good fit of LLTM. However, since such an impression may be misleading the model fit has been tested using the LRT comparing the likelihood under the Rasch model with that under the LLTM. The LRT is significant $(\chi^2(35) = 3550, p < 0.001)$ leading to the rejection of the LLTM, in spite of the graphical method indicates a good match between the Rasch and LLTM item difficulties. Thus, the 48 item parameters of the Rasch model cannot be explained by only 16hypothesized basic parameters and the models must be rejected.

In spite of the significance of LRT statistic, the LLTM is still useful. As stressed by Hambletonand Van der Linden(1997), even if the explanation of item difficulty in terms of basic parameters α_j is not perfect, the LLTM still allows one to predict item difficulty of new items at least approximately. Furthermore, even though LLTM estimates are not wholly acceptable asestimates of difficulty parameters, then their ability to relate item performance to cognitive theory has proven useful in application such as assessing treatment effects and modeling item bias (Fischerand Formann(1982)).

In order to improve the LLTM fit an homoscedastic error term is added according to (7), so that the explanation of the item difficulties based on the item properties does not need to be perfect. As expected, from Table 2 the LLTM plus error fits the data significantly better than the LLTM. Finally, the latent regression LLTM plus error(8) taking into account both person properties and item properties has been considered andthe LRT comparing the

latent regression LLTM plus error to the LLTM plus error is significant $(\chi^2(1) = 165\,,\, p < 0.001)$.

As previously stressed, the LLTM also allows for interactions between the item properties. Let us consider the interactions betweenthe domain *competency* and the domains *context* and *content*, respectively. In Table 3 (see Appendix) the estimates for the LLTM plus error with interactions and the latent Regression LLTM plus error with interactions are reported, where significant coefficients are bolded-faced $(p < 0.05)$.

The estimated effect of macro-area is 0.78 with a standard error of 0.06,a similar result was obtained with the latent regression Rasch model. Furthermore, also the estimated item property effects are about the same as those obtained with the LLTM plus error with interactions. The differences in the estimates of Table 3 (see Appendix), checked by a Wald test, are not statistically significant.

Since in the model specification we have suppressed the overall intercept, the fixed effect of the item format is expressed as five means. More specifically, *open constructed response* (*OCR*) appears to be the most difficult item format whereas *multiple choice response* (*MC*) is the easiest. The *OCR*effect is statistically significant.

The interpretation is different for the domains situation, content and process. In fact, since each domain is considered as one facet, then for each facet with *l* categories *l-1* indicator variables are needed. Clearly, the interpretation of each indicator variable is always in reference to the base category. In our analysis, *education (ED)*, *change and relationship (C)* and *connection(CON)*were set as the reference categories for the three facets. Note that a positive estimate in LLTM means that the probability of getting the correct answer on the item belonging to the specified category of a facet is lower compared to that of the reference category.

With regard to the situation domain *scientific (SC)*, *public (PU)* and *personal (PE)* are easier than *educational (ED)*, but these coefficients are not statistically significant.

With regard to the content domain both *uncertainty (U)*, *space and shape (S)* and *quantity (Q)* appear to be more difficult than *change and relationship (C)*, whereas *quantity (Q)* is easier and *uncertainty (U)* is the most difficult. As expected, both *connection (CON)* and *reflection (REF)* are harder than *reproduction (REP)* with the *reflection (REF)* being the most difficult.

Finally, the effect on item difficulty of competency domain is larger than the effect of situation and content domains, respectively. This is confirmed by the analysis of interaction effects in Table 3.

In order to investigate possible differences in the item property effects between Italiangeographical macro-areas, the LLTM plus error with interactions has been separately fitted to South-Italy and North-Italy data. The estimated person variance for the two geographical macro-areas is approximately the same 1.40 with a standard error of 1.18.

The estimates of item effect properties are reported in Table 4 (see Appendix), where significant coefficients are bold-faced $(p < 0.05)$.

The results obtained from the analysis of Table 3 are confirmed when the LLTM plus error with interactions model has been fitted toSouth-Italy and North-Italy data, respectively. In order to identify different effects of item property we perform a Wald test on item property. Note that *open constructed response(OCR)* appears to be more difficult for the South-Italy than North-Italy $(p < 0.1)$.Such item format is regarded as most suitable for assessing items that would be associated with high-order cognitive activities, since a more

extended response is required from students, that must explain how the answer was reached.

4. Concluding remarks

In conclusion, *open constructed response* appears to be the most difficult for the domain item format. For the content domain *uncertainty*, *space and shape* and *quantity* appear to be more difficult than *change and relationship*, whereas *quantity*is easier and *uncertainty* is the most difficult. For the domain competency, both *connection* and *reflection* are harder than *reproduction*, with the *reflection* being the most difficult.

With regard to the North-Italy versus South-Italy comparison, the North-Italy students are more inclined to mathematics than South-Italy student overall, furthermore *open constructed response* appears to be more difficult for the South-Italy than North-Italy.

References

1. Baker,F.B. (1993). **Sensitivity of the linear logistic test model to misspecification of the weight matrix**. Applied Psychological Measurement 17, 201-210
2. De Boeck, P., Wilson, M. (2004).**Explanatory Item Response Models: A generalized Linear Approach**. Springer-Verlag, New York
3. Fischer, G. H. (1973). **The linear logistic test model as an instrument in educational research**. Acta Psychologica, 3, 359-374
4. Fischer,G.H., Formann,A.K. (1982). **Some applications of logistic latent trait model with linear constraints on the parameters**. Applied Psychological Measurement,6,396-416
5. Fischer,G.H., Molenaar,I.W. (1995). **Rasch Models: Foundations, Recent Developments and Application**. New-York, Springer
6. Hambleton R. K., Van der Linden W.J. (1997). **Handbook of Modern Item Response Theory**. Springer, New York
7. Janssen R., Schepers J., Peres D. (2004). **Models with Item and Item Group Predictors**. In P. De Boeck, M. Wilson (eds.). Explanatory Item Response Models. Springer-Verlag, New York
8. Kubinger, K.D. (2008). **On the revival of the Rasch model-based LLTM: From composing tests by item generating rules to measuring item administration effects**. Psychology Science Quarterly, 50, 311-327
9. OECD 2003. **The PISA 2003 Assessment Framework - Mathematics, Reading, Science and Problem Solving Knowledge and Skills**.http://www.oecd.org
10. OECD 2006. **Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006**.http://www.oecd.org
11. R Development Core Team (2009). **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
12. Reise, S. (1990). **A comparison of item- and person-fit methods of assessing model-data fit in IRT**. Applied Psychological Measurement, 14 , 127-137
13. Zwinderman, A.H. (1991). **A generalized Rasch model for manifest predictors**. Psychometrika,56, 589-600

Appendix

Table 3. LLTM and Latent LLTM estimates with interactions between item properties

Parameter	LLTM with interactions	Std. Error	Latent LLTM with interactions	Std. Error
CCR	-0.131	0.468	0.258	0.468
CMC	0.211	0.428	0.600	0.428
MC	-0.555	0.524	-0.166	0.525
OCR	**1.274**	0.495	**1.663**	0.497
SR	0.109	0.510	0.499	0.511
PE	-0.681	0.498	-0.678	0.498
PU	-0.690	0.381	-0.690	0.381
SC	-0.824	0.435	-0.823	0.435
REF	**1.932**	0.839	**1.934**	0.839
REP	**-4.939**	1.078	**-4.931**	1.077
Q	**0.911**	0.435	**0.910**	0.434
S	**1.209**	0.396	**1.209**	0.396
U	**1.531**	0.448	**1.533**	0.448
PE:REF	1.033	0.831	1.034	0.831
PU:REF	0.649	0.741	0.650	0.741
SC:REF	-0.461	0.857	-0.460	0.857
PE:REP	-0.540	0.798	-0.537	0.798
PU:REP	**-2.391**	0.783	**-2.390**	0.782
SC:REP	**-5.806**	1.193	**-5.801**	1.193
REF:Q	1.295	0.803	1.293	0.803
REP:Q	**-2.585**	0.858	**-2.581**	0.858
REF:S	**1.869**	0.677	**1.871**	0.677
REP:S	**-2.612**	0.941	**-2.607**	0.941
REF:U	1.324	0.718	1.327	0.717
REP:U	**-2.290**	1.050	**-2.282**	1.049

Table 4. LLTM estimates with interactions between item properties in Italy macro-area

Parameter	North-Italy estimates	Std. Error	South-Italy estimates	Std. Error
CCR	-0.727	0.458	0.518	0.490
CMC	-0.243	0.418	0.691	0.449
MC	-0.940	0.513	-0.147	0.550
OCR	0.714	0.484	**1.907**	0.520
SR	-0.409	0.499	0.619	0.535
PE	-0.699	0.487	-0.688	0.522
PU	-0.598	0.372	**-0.834**	0.399
SC	-0.698	0.425	**-0.998**	0.456
REF	**2.121**	0.821	1.689	0.884
REP	**-4.596**	1.070	**-5.364**	1.128
Q	**1.034**	0.425	0.827	0.454
S	**1.297**	0.388	**1.122**	0.414
U	**1.473**	0.437	**1.634**	0.468
PE:REF	1.074	0.813	0.944	0.875
PU:REF	0.803	0.724	0.415	0.779
SC:REF	-0.198	0.839	-0.773	0.904
PE:REP	-0.847	0.782	-0.352	0.836
PU:REP	**-2.230**	0.766	**-2.657**	0.818

SC:REP	**-5.274**	1.180	**-6.599**	1.252
REF:Q	**1.550**	0.785	1.074	0.841
REP:Q	**-2.337**	0.858	**-2.825**	0.899
REF:S	**1.834**	0.662	**1.908**	0.714
REP:S	**-2.098**	0.938	**-3.122**	0.986
REF:U	1.356	0.701	1.320	0.752
REP:U	-1.921	1.043	**-2.615**	1.099

[1] Daniela Marella is a researcher of Statistics at the Department of Educational Sciences of the University Roma Tre. Her main research interests include Statistical matching and Nonsampling errors.In those areas, she published several papers in refereed journals and in proceedings of national and international conferences.

[2] Carlo Di Chiacchio is an experimental psychologist graduated at the Faculty of Psychology "University of Rome LA SAPIENZA" and he has a PhD in Psychometrics at the Faculty of Psychology "University of Rome LA SAPIENZA". His main research interests are on the assessment of competencies and learning strategies, and on the development of emotion understanding during the childhood. He also worked at the University of Cassino for a national project concerning school guidance. Since 2001 he has been working as data analyst at the National Institute for the Evaluation of the Educational System on the Project for the International Student Assessment (PISA), and on several national projects as well. Currently he is the National Project Manager for PISA 2012.

[3] Giuseppe Bove is Professor of Statistics at the Department of Educational Sciences of the University Roma Tre. He was a member of the Board of the Italian Statistical Society (SIS) from 2002 to 2006. His main scientific interests concern multivariate statistical analysis and in particular Factor Analysis and Multidimensional Scaling, and their applications in Education. He published several papers on national and international journals, proposing new methods for multidimensional representation of two and three-way proximity data.