

# The Rasch Model Analysis for Understanding Mathematics Proficiency—A Case Study: Senior High School Sardinian Students

Carlo Andrea Pensavalle<sup>1</sup>, Giuliana Solinas<sup>2</sup>

<sup>1</sup>Department of Science for Nature and Environmental Resources, University of Sassari, Sassari, Italy

<sup>2</sup>Department of Biomedical Sciences, Laboratory of Epidemiology and Biostatistics, University of Sassari, Sassari, Italy

Email: [pensa@uniss.it](mailto:pensa@uniss.it)

Received November 5<sup>th</sup>, 2013; revised December 5<sup>th</sup>, 2013; accepted December 13<sup>th</sup>, 2013

Copyright © 2013 Carlo Andrea Pensavalle, Giuliana Solinas. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2013 are reserved for SCIRP and the owner of the intellectual property Carlo Andrea Pensavalle, Giuliana Solinas. All Copyright © 2013 are guarded by law and by SCIRP as a guardian.

Many students enrolled in Italian Universities don't have regular careers in their first level of university education mainly because of the obstacles encountered in their studies. As far as Mathematics proficiency assessment there are national and international important systematic studies which give evidence for positive relationships between achievement and varied classroom settings and provide a larger context for better understanding regional performance, extending and enriching the local picture. This paper presents some preliminary results aiming to evaluate specific Mathematics abilities in Senior High School Sardinian Students approaching university studies. For this purpose the Rasch model was applied. The information obtained by the application of this measurement approach provides clear indication for further analysis to ascertain the causes that influence Mathematics proficiency. The Rasch model was performed on 888 students coming from 28 High Schools located in the central-northern part of the Sardinia region using a questionnaire to evaluate the level of ability in procedural fluency and a second questionnaire to evaluate strategic competence. The study provides more evidence in favor of Rasch Model as an appropriate way for teachers and researchers to obtain richer interpretations on the relationship between students' proficiency and test items. Based on Infit and Outfit MNSQ, all items are within acceptable range between 0.7 - 1.3. In light of preliminary results there is a need for local schools and universities to become attuned to the full extent of the Mathematics problem as it affects Senior High School Sardinian Students.

*Keywords:* Rasch Analysis Model; Mathematics Proficiency; Procedural Fluency; Strategic Competence

## Introduction

This work is part of the wider project STUD.I.O.: "Students in Orientation", funded by the Sardinia Region with participation of the European Social Fund and the Italian Ministry of Labor and Social Policies. One of the main goals of the project is to promote a realignment of general skills (Reading/Writing, Mathematics and Science), aimed at improving the preparation of students attending the last year of High School (19 years old). There is evidence of a serious decline in students' mastery of basic mathematical skills and level of preparation for mathematics-based degree courses since the late nineties (Hawkes & Savage, 2000). This decline is well established and affects students at all levels. Students' difficulties in accessing universities are increasing and it is common to doubt if High Schools are adequately preparing students for university entrance. The admission test scores are becoming alarmingly low, anticipating a trend where students will not be regular with their course studies. This represents a serious problem in the Italian University system especially after the University reform of 2001

which introduced the "3 + 2" system for most of the curricula. The system is articulated in two parts: a first three year course degree corresponding to a first level of university education in a specific field of interest, followed by a two year specialized course degree in the same field. This system which was supposed to cope with the gap in students' preparation entering university education has generated instead a great level of distress in students who try to adapt effectively to the standards required by university studies and prompted a central role for universities to promote specific support to students in difficulty, creating instruments and opportunities to enable them to finish their courses within the allotted time-span. (Chiandotto et al., 2005; Cingano & Cipollone, 2007; Delvecchio & D'Ovidio, 2002; Solinas et al., 2012). Many students enrolled in Italian universities don't have regular careers in their first level of university education, mainly because of the obstacles encountered in their studies due to lack of preparation in general skills, or because of a wrong choice. The withdrawals from university generally occur in the first year of study, representing about

18% of the total students enrolled (MIUR/CNVSU, 2011). Reliable measures of students' performance are needed in order to identify risk factors and predictors of success, also to evaluate the quality and efficiency of curricula. The typical approach in analyzing students' career, centered on the grades and credits acquired is affected by the inherent characteristics of any higher education system setting (Mouw & Khanna, 1983; Rohde & Thompson, 2007). Therefore, it is necessary to provide standardized indicators to make instructional decisions and to evaluate individual student achievement in order to develop strategies to better understand causes and find solutions. As far as Mathematics proficiency assessment, there are national and international important systematic studies which give evidence for positive relationships between achievement and varied classroom settings (Dunne et al., 2012; Mossi et al., 2012) and provide a larger context for better understanding regional performance, extending and enriching the local picture (Samuelsson, 2010; Teddlie & Reynolds, 2000). Among the most important cognitive tests, we would like to mention the international one conducted by the Organization for Economic Cooperation and Development (OECD), started in 1997, known as Programme for International Student Assessment (PISA), and the national one conducted by the Italian Institute for the Educational Evaluation (INVALSI), started in 2007. Both studies generate theoretical reflections and empirical studies related to issues involved in the evaluation of educational programs, educational institutions, educational personnel and student assessment. They involve groups of students enrolled in the second and fifth year of primary school, in the first and third year of Junior High School and in the second year of High School. The statistics are available worldwide and allow different kinds of evaluations to identify and test promising approaches and curricula able to improve student outcomes. Recent indications show that the average level of Mathematic proficiency in a fifteen year old Italian student is lower (score 483) than the OECD (score 496) with a statistically significant difference (PISA, 2009). Based on these premises and under the circumstances of the project STUD.I.O., we decided to extend the study to Senior High School Students (19 years old), who voluntarily participated to the following two evaluations: the first one aimed at measuring if the average level in procedural fluency established by PISA regarding fifteen year old students, was well achieved by the sample under study. In order to do so, we administered a questionnaire composed of specific items similar in difficulty to the one administered by PISA surveys. The second one aimed at measuring the level of strategic competence required to freshman entering any Italian university scientific degree program by CINECA, a non-profit consortium established in 1969, composed of 54 Italian universities, two National Research Centers, and the Ministry of University and Research (MIUR).

The questionnaire administered composed of specific items was constructed in a similar way to the one used in the last past years. In this study, the choice of two questionnaires provide evidence to measure two distinct aspects of Mathematics proficiency in the same group of nineteen year old students graduating from High School, who aim to pursue scientific academic studies. For this purpose the Rasch model was applied. The information obtained by the application of this measurement approach provides clear indication for further analysis to ascertain the causes that influence Mathematics proficiency.

## Methodology

### Data Collection

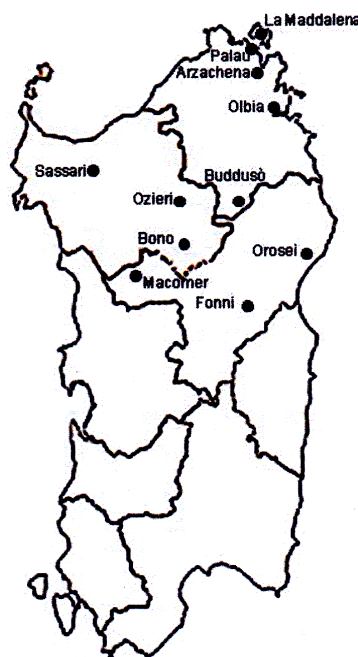
The sample consists of 1233 Senior High School Students coming from 28 High Schools located in the central-northern part of the Sardinia region (**Figure 1**), representing about 10% of the entire population of regional nineteen year old High School Students.

### Test Instruments

The first test instrument used in this study was a self-developed 20 item questionnaire, constructed according to guidelines provided by OECD curriculum specifications, aiming to evaluate the level of ability in procedural fluency, that is performing mathematical procedures appropriately and efficiently, including those that require decisions in sequence and interpreting and using representations based on different information sources (Kilpatrick et al., 2001). The second test instrument is a self-developed 25 item questionnaire, constructed according to guidelines provided by CINECA, aiming to evaluate the level of strategic competence, connected to operational skills and problem solving strategies. In particular, the ability to recognize and set a problem properly, to select appropriate information, to identify and organize the most appropriate tools to be used and to represent data and situations in an appropriate manner (Kilpatrick et al., 2001). The questionnaires were administered online through the Moodle platform. The first questionnaire was administered in February 2013, the second in April during scholastic activity.

### Methods of Statistical Analysis

In this paper instead of Classical Test Theory methods, we used the probabilistic approach of Item Response Theory (IRT)



**Figure 1.** Map of the Sardinian High Schools participating in the study.

named Rasch Model (Rasch, 1960). We defined Mathematics proficiency as a latent trait yielding a score that locates student ability and question difficulty on a common scale, measured in logits (Fisher, 1995). In this model the probability  $P_i$  of answering correctly to the  $i$ -th item is expressed in function of the difference between the respondent proficiency level  $\theta$  and the item difficulty level  $b_i$  as follows:

$$P_i = \frac{\exp(\theta - b_i)}{1 + \exp(\theta - b_i)}$$

This brings to the definition of logit:

$$\text{logit}(P_i) = \ln\left(\frac{P_i}{1 - P_i}\right) = \theta - b_i$$

Values are placed along a continuum enabling the item difficulty level to be compared with the respondent proficiency level. The Rasch model allows the estimation of these two parameters, with the advantage that despite different test items with different data collection periods, students' progress is estimated and reported on a common scale regardless of fluctuation in test difficulty. The model was implemented in the Acer ConQuest software, version 3.1 (Adams et al., 2012). The Marginal Maximum Likelihood Estimation method was considered in order to estimate the parameters above, through iterative calibration of both person and item (Bock, 1970). Information about the model quality was gained from the fit statistics on item level. These statistics show whether single test items correspond with the model assumptions. Conquest software generates Infit Mean Square (MNSQ) and Outfit MNSQ statistics, which provide indications on the differences between the data and model's expectations (Bond & Fox, 2007; Wright & Masters, 1982). In this study we adopted the range of acceptable fit between 0.7 and 1.3, for both fit indices (Bond & Fox, 2007), because Infit values greater than 1.30 and smaller than 0.70 are labeled problematic (Adams, 2002). Reliability and validity of measures meaning and interpretation are provided within the framework of Rasch measurement model. In addition, a Wright Map (Boone & Scantlebury, 2006) is used to visually display the simultaneous distributions (or "performances") of items and respondents related to the instruments used and the sample considered. Redundant items appear on the same or nearly the same point on the Wright map scale. The Wright map also illustrates regions of the scale in which items are absent, identifying where new items are needed.

## Results

The 72% of the total sample ( $n = 1233$ ) matched the specifications required to be part of the statistical analysis: respect of time administration and answer to all of the items of the questionnaire. **Table 1**, displays the main characteristics of the study sample: 55% of the students come from Lyceum, 28% Technical and 17% Vocational Schools.

As presented in **Table 2**, the Rasch Analysis applied to the first instrument finds both means of Infit MNSQ and Outfit MNSQ close to the expected value of 1.00.

Inspection with individual items shows Infit MNSQ values ranging from 0.82 to 1.21 while Outfit MNSQ values range from 0.93 to 1.15. Evaluation of the Infit statistics reveals that all cut off points of the items are compliant with the model. The unidimensionality of the construct was assessed by using prin-

**Table 1.**  
Sample characteristics and school typology.

Gender	Total	%
Males	414	47
Females	474	53
School Typology	Total	%
Scientific Studies	288	32
Classical Studies	205	23
Technical	246	28
Vocational	149	17

**Table 2.**  
First instrument: Item statistics.

Item Label	Difficulty (Logits)	Standard Error	Infit MNSQ	Outfit MNSQ
1	0.144	0.070	0.99	0.99
2	0.205	0.070	0.92	0.93
3	-0.433	0.072	0.99	0.99
4	0.810	0.072	1.06	1.04
5	0.444	0.070	1.01	1.01
6	0.661	0.071	1.21	1.15
7	1.142	0.075	0.99	0.99
8	0.762	0.072	1.10	1.06
9	0.920	0.073	1.06	1.04
10	-0.038	0.070	0.95	0.97
11	-0.219	0.071	0.91	0.93
12	-0.471	0.073	1.10	1.08
13	0.656	0.071	0.97	0.97
14	-0.819	0.076	0.99	1.00
15	0.757	0.072	1.01	0.99
16	-0.018	0.070	1.00	1.00
17	-1.585	0.091	1.01	0.99
18	-2.266	0.114	0.89	0.97
19	-1.724	0.095	0.82	0.94
20	1.072	0.074	0.93	0.94

cipal components analysis of Rasch residuals and item fit statistics (Brentari & Golia, 2010). Because the scores demonstrate little variation from model expectation there is evidence of consistency between students' responses and items on the scale and the model expectations.

Reliability of item difficulty measures is 0.99 suggesting that the ordering of item difficulty is replicable with other comparable sample of students. Consistency of student measures (KR20), equivalent to Cronbach's alpha (Cronbach, 1951), is 0.72 indicating that the ordering of the student proficiency can be likely replicated since most of the variance is attributed to the true variance of the Mathematics proficiency construct.

Based on the following **Figure 2**, where the student ability (to the left) and the item difficulty (to the right) are displayed graphically with the most able students and most difficult items

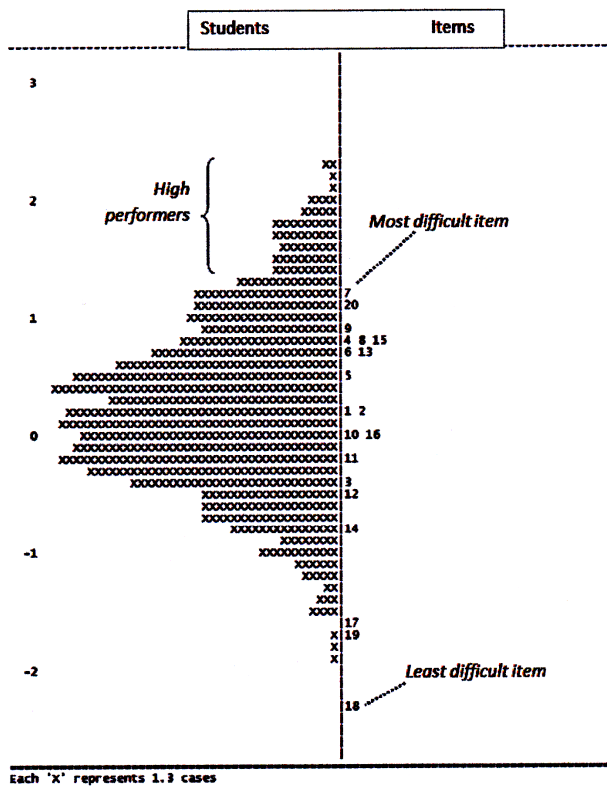


Figure 2. Wright map of the first instrument.

at the top, it is possible to note that the majority of the students (each  $x$  represent 1.3 students) are represented in the center of the graph (the area included between  $-1$  and  $+1$  logits) as most of the items (represented by their labels). The items and the students are quite well approximated by normal distribution. Therefore, the difficulty of the questionnaire is aligned with the average level of student ability, with Item 7 the most difficult and Item 18 the easiest (see Appendix 1). At the bottom and at the top of the chart there are two small groups of students, for which there are no items able to measure their ability. Alarming, it seems that the average level of Mathematics proficiency shown by our sample of Senior High School Students is comparable with that of fifteen-year-old students. Let's now proceed to the analysis of the data collected with the second questionnaire which was developed to measure in the sample of Senior High School Students the presence of the strategic competence level required to enter any Italian university scientific degree program. As presented in the following Table 3, the Rasch Analysis finds both means of Infit MNSQ and Outfit MNSQ close to the expected value of 1.00. Inspection with individual items shows Infit MNSQ values ranging from 0.92 to 1.11 while Outfit MNSQ values range from 0.94 to 1.02. As before having assessed the unidimensionality assumption of the construct and the acceptable variation of the scores from model expectation, there is evidence of consistency between students' responses and items on the scale and the model expectations. Reliability of item difficulty measures is 0.82 suggesting that the ordering of item difficulty is quite replicable with other comparable sample of students. Consistency of student measures is moderate ( $KR20 = 0.51$ ) and mainly this questionnaire resulted for the students' sample composed of items whose

Table 3. Second instrument: Item statistics.

Item Label	Difficulty (Logits)	Standard Error	Infit MNSQ	Outfit MNSQ
1	-0.126	0.189	1.03	1.02
2	-0.013	0.192	0.98	0.98
3	0.312	0.204	1.03	1.02
4	-0.856	0.181	0.96	0.96
5	-0.755	0.181	1.01	1.01
6	0.446	0.211	0.97	0.99
7	0.446	0.211	1.11	1.06
8	-0.015	0.192	1.02	1.01
9	-1.236	0.185	1.01	1.00
10	0.355	0.206	1.01	1.01
11	0.311	0.204	1.03	1.02
12	0.400	0.208	1.05	1.03
13	0.446	0.211	0.96	0.96
14	0.144	0.197	0.92	0.94
15	0.401	0.208	0.99	1.00
16	0.065	0.195	1.05	1.03
17	-0.342	0.184	0.96	0.96
18	-0.199	0.187	0.97	0.98
19	0.186	0.199	0.96	0.98
20	-0.412	0.183	0.99	0.99
21	0.145	0.197	1.03	1.02
22	-0.481	0.183	1.00	1.00
23	0.105	0.196	1.00	0.99
24	0.402	0.208	0.96	0.97
25	0.270	0.202	1.04	1.02

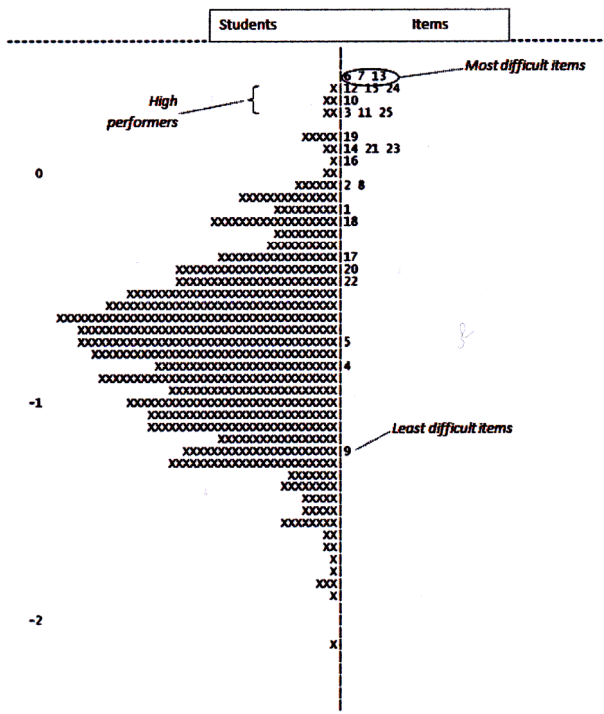
difficulty levels were higher than the students' ability levels, as shown in the following Figure 3.

Item 6, 7, 13 resulted to be the most difficult and Item 9 the easiest (see Appendix 2).

### Conclusion

This paper presents some preliminary results of the project STUD.I.O., aiming to evaluate specific Mathematics abilities in Senior High School Sardinian Students approaching university studies. The study provides more evidence in favor of Rasch Model as an appropriate way for teachers and researchers to obtain richer interpretations on the relationship between students' proficiency and test items. Based on Infit and Outfit MNSQ, all items are within acceptable range between 0.7 - 1.3.

Both statistics show that there is enough evidence that the data obtained fits the model expectations. With regard to students' procedural fluency investigated with the first instrument most of the students show responses that are within the expectation of the model. The results give suggestions that the items' questionnaire can discriminate students with different Mathematics proficiency levels. Since the measures are in interval scale, one important observation is that the most difficult item



**Figure 3.**  
Wright map of the second instrument.

of the first questionnaire, Item 7 (1.142 logits), was higher in difficulty as compared to Item 18 that was almost twice as easy (-2.266 logits). For the second questionnaire, it is difficult to define the students' strategic competence because of large gaps of the scale in which items are absent. These results provide evidence for construct validity of the second questionnaire. In particular, the item difficulty measures show that this questionnaire is composed of items whose level of difficulty did not correspond to the level of Mathematics proficiency of the students' sample. Therefore, in light of these preliminary results, there is a need for local schools and universities to become attuned to the full extent of the Mathematics problem as it affects Senior High School Sardinian Students. The diagnostic testing of new undergraduates is recommended as an effective means of investigating levels of Mathematics proficiency and identifying actions to improve quality, equity and efficiency of Sardinian educational systems.

### Acknowledgements

The authors would like to thank the Sardinia Region, the European Social Fund and the Italian Ministry of Labor and Social Policies whose funding made this study possible.

### REFERENCES

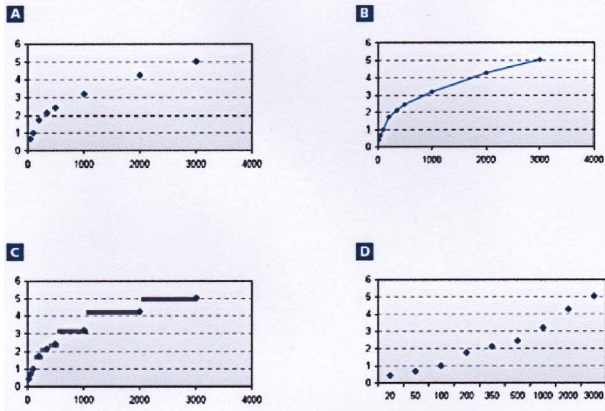
- Adams, R. (2002). Scaling PISA cognitive data. In R. Adams, & M. Wu (Eds.), *PISA 2000 technical report* (pp. 99-108). Paris: OECD.
- Adams, R., Wu, M., & Wilson, M. (2012). *ACER ConQuest 3.1*. Melbourne: ACER.
- Bock, R. D., & Lieberman, M. (1970). Fitting a response model for dichotomously scored items. *Psychometrika*, 35, 179-187. <http://dx.doi.org/10.1007/BF02291262>
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, 90, 253-269. <http://dx.doi.org/10.1002/sce.20106>
- Brentari, E., & Golia, S. (2007). Unidimensionality in the Rasch model: how to detect and interpret. *Statistica*, 67, 253-261.
- Chiandotto, B., & Giusti, C. (2005). L'abbandono degli studi universitari. In C. Crocetta (Ed.), *Modelli statistici per l'analisi della transizione università-lavoro* (pp. 1-22). Padova.
- Cingano, F., & Cipollone, P. (2007). University drop-out: The case of Italy. *Temi di discussione del Servizio Studi*, 626. Roma: Banca d'Italia.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. <http://dx.doi.org/10.1007/BF02310555>
- Delvecchio, F., & D'Ovidio, F. (2002). Alcuni fattori di rischio per l'abbandono degli studi universitari. In L. C. Sardi, & F. Delvecchio (Eds.), *Indicatori e metodi per l'analisi dei percorsi universitari e postuniversitari* (pp. 101-124). Padova: Cleup.
- Dunne, T., Long, C., Craig, T., & Venter, E. (2012). Meeting the requirements of both classroom-based and systematic assessment of mathematics proficiency: The potential of Rasch measurement theory. *Pythagoras*, 33, 16 p. <http://dx.doi.org/10.4102/pythagoras.v33i3.19>
- Fisher, G. (1995). Derivations of the Rasch model. In G. H. Fisher, & I. W. Molenaar (Eds.), *Rasch models: Foundations recent developments and applications* (pp. 15-38). Berlin: Springer-Verlag. [http://dx.doi.org/10.1007/978-1-4612-4230-7\\_2](http://dx.doi.org/10.1007/978-1-4612-4230-7_2)
- Hawkes, T., & Savage, M. (2000). *Measuring the mathematics problem*. London: Engineering Council.
- Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping how children learn Mathematics*. Washington DC: National Academic Press.
- MIUR/CNVSU (2011). Undicesimo rapporto sullo stato del sistema universitario. <http://www.cnvsu.it>
- Mossi, P., Venuleo, C., Tondo, P., & Salvatore, S. (2012). The measurement of students' performance. The use of an extended Rasch model for the analysis of predictors of high educational performance. *EJASA*, 5, 393-399.
- Mouw, J. T., & Khanna, R. K. (1983). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27, 328-336.
- PISA (2009). Results: What students know and can do. Student performance in reading, mathematics and science (Volume I). <http://dx.doi.org/10.1787/9789264091450-en>
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35, 83-92. <http://dx.doi.org/10.1016/j.intell.2006.05.004>
- Samuelsson, J. (2010). The impact of teaching approaches on students' mathematical proficiency in Sweden. *International Electronic Journal of Mathematics Education*, 5, 61-78.
- Solinas, G., Masia, M.D., Maida, G., & Muresu, E. (2012). What really affects student satisfaction? An assessment of quality through a university-wide student survey. *Creative Education*, 3, 37-40. <http://dx.doi.org/10.4236/ce.2012.31006>
- Teddle, C., & Reynolds, D. (2000). *International handbook of school effectiveness*. London: Falmer.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago: Mesa Press.

### Appendix 1. First Questionnaire

#### Item 7

Zedland's postal charges are based on the weight of the goods you want to mail, as shown in the table below (approximated to the nearest gram).

Which one of the following graphs is the best representation of the postal charges in Zedland?

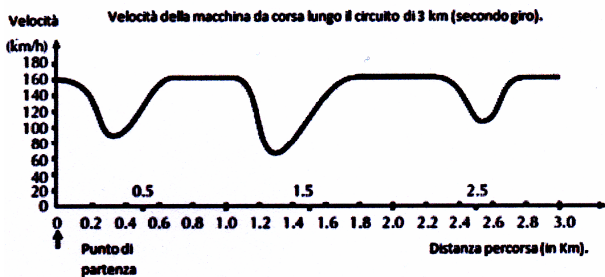


(The horizontal axis shows the weight in grams and the vertical axis the charge in zeds).

- A. Graph A
- B. Graph B
- C. Graph C
- D. Graph D

#### Item 18

The chart below shows the changes in speed of a racing car on the second lap of a circuit 3 kilometers long.



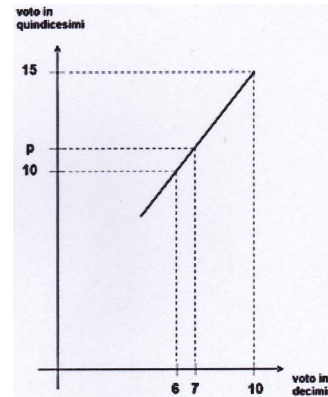
- Where is the slower speed during the second round?
- A. At the starting point
  - B. At about 0.8 km
  - C. At about 1.3 km
  - D. In the middle of the round

### Appendix 2. Second Questionnaire

#### Item 6

A student has obtained a grade of 7/10 in a Math test. You want to express this grade in fifteenths using the criteria suggested in the figure.

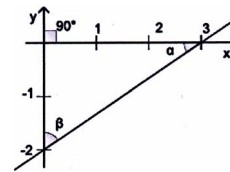
If  $p$  is the grade in fifteenths which corresponds to 7/10, then:



- A.  $11 < p < 11.1$
- B.  $11.1 < p < 11.2$
- C.  $11.2 < p < 11.3$
- D.  $11.3 < p < 11.4$

#### Item 7

Consider the angles  $\alpha$  e  $\beta$  as in the figure which of the following relationships is correct?



- A.  $\tan \beta < \cos \alpha$
- B.  $\sin \beta < \cos \alpha$
- C.  $\cos \beta > \cos \alpha$
- D.  $\tan \beta > \tan \alpha$

#### Item 13

A number sequence  $x_0, x_1, x_2, \dots$  is defined as:

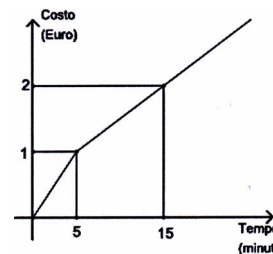
$$\begin{cases} x_0 = x_1 = 1 \\ x_i = x_{i-1} + 2x_{i-2} \end{cases} \text{ for any } i \geq 2.$$

- Then  $x_6$  is equal to
- A. 32
  - B. 43
  - C. 85
  - D. 61

#### Item 9

Shown in the following graph is the cost of a phone call in function of time.

What is the cost in euro of a 20 minute phone call?



- A. 2.00
- B. 2.25
- C. 2.50
- D. 2.75