Electronic Journal of Applied Statistical Analysis
EJASA (2012), Electron. J. App. Stat. Anal., Vol. 5, Issue 3, 393 – 399
e-ISSN 2070-5948, DOI 10.1285/i20705948v5n3p393

THE MEASUREMENT OF STUDENTS PERFORMANCE. THE USE OF AN EXTENDED RASCH MODEL FOR THE ANALYSIS OF PREDICTORS OF HIGH EDUCATIONAL PERFORMANCE

Piergiorgio Mossi*, Claudia Venuleo, Paola Tondo, Sergio Salvatore

Department of Hystory, Society and Human Studies, University of Salento, Italy

Received 01 August 2012; Accepted 26 September 2012
Available online 16 November 2012

Abstract: *Characteristics of the higher education programs (e.g. non systematic variability of course's difficulty among and within programs and across times) make observed data (e.g. number of credits acquired) poorly informative indexes of the students' performance. As an alternative, an extended version of Rasch model (the Three Facets Model, TFM) is proposed. TFM conceptualizes student's performance as the expression of a three-component latent variable to be esteemed. In so doing, TFM is able to take into account the non-systematic sources of variation characterizing higher education settings, thus avoiding limits entailed in the use of indexes based on observed data. An exemplificative longitudinal case study has been performed, aimed at detecting predictors of performance within an undergraduate program of psychology of an Italian university. Two regression models have been compared: one using a traditional index of performance based on observed data versus one using the TFM estimation.*

Keywords: *Rasch analysis, three facets model, higher educational performance, performance predictors.*

1. Introduction

The measurement of higher education students' performance is an important task. Reliable measures of students' performance are needed in order to evaluate the quality and efficacy of programs, as well as to identify risk factors and predictors of success. To this end, researchers

* E-mail: pmossi@gmail.com

usually base their analyses on observed data concerning the students' career – in particular: grades obtained at the various disciplines and/or credits acquired [4], [6]. Such data is assumed as being informative per se - as factual truths.
Yet, one should recognize that the reliability of such an approach is affected by some inherent characteristics of higher education settings. Firstly, students' abilities are distributed in a non-systematic way through educational contexts and over time. Huge numbers of statistics and many studies of sociology of education have systematically highlighted the differences both between and within macro socio-cultural contexts (e.g. between North and South Italy). Secondly, the degree of difficulty of academic courses varies both within and among programs – two courses associated with the same amount of credits may differ even dramatically as to levels of ability required for attending them; the same can be said even for the same subject as provided by two programs of different universities. Thirdly, the probability of success with courses may vary over time and among programs as a result of the variation of the standard of evaluation adopted by teachers who rate students' learning. Consequently one has to conclude that observed data as raw amount of credits and average grades does not lend itself to be seen in a single, invariant way: its use as indexes of performance is not very informative.
As an alternative, an extended version of Rasch model, the *Three Facets Model* (TFM) [2] is proposed, as a more reliable way of calculating students' performance. Such a method takes into account the non-systematic sources of variation mentioned above, thus avoiding limits entailed in traditional indexes based of observed data.
Rasch models are largely used in the educational field [9], [1]. This paper provides a less common usage: the estimation of a global index of students' performance, to be used for the detection of predictors of career failure/success. To this end, in what follows, after a synthetic illustration of the method, a longitudinal case study is presented. The case concerns the analysis of predictors of the performance within an undergraduate program of psychology provided by an Italian university.

2.The Three Facets Model

The Three Facets Model (TFM) is an expansion of the classic two-component Rasch model, obtained through the addition of a further component. The mathematical properties of the Rasch model are however kept [5].
The TFM is a logistic estimation. It conceives of ordinal observations as the expression of one latent variable. The latter is measured in terms of the linear composition of independent elements. As concerns the model of interest here, the student's performance is modeled as a function of three elements: student *ability*, course *difficulty* and rater *severity*. The more the ability, the less the difficulty as well as the severity, the greater the student's probability of obtaining credits associated with the course. The TFM can be formalized as follows:

$$P_{nsi} = \frac{\exp(\beta_n - \delta_s - \theta_i)}{1 + \exp(\beta_n - \delta_s - \theta_i)} \quad (1)$$

where P_{nsi} is the probability that the student n passes the examination related to the course s as judged by the rater i; β_n is the ability of the student n, δ_s is the difficulty of the course s and θi represents the severity of the rater i evaluating the student.
Being additive in its parameters, the model is consistent with the requirement for interval measurement. Like the original two-component Rasch model, raw scores provide all the information needed for the measurement of parameters (β_{sn}, δ_s e θ_i). This is so because TFM contains the Rasch model's property of measurement invariance: the measurement of parameters does not depend on the characteristics of the test adopted; for instance, the measurement of the student's ability is always the same, independently from the set of courses attended. As a result, the estimation of the parameters does not require the reference to a normative population.
The measurement of parameters is based on Fisher's principle of statistical sufficiency: for each parameter, the maximum-likelihood is given when the expected score equals the observed score. In doing so, the model assumes the randomness of data. This assumption is controlled by means of the fit statistics that can be performed in order to test the adequacy of data to the model.
Differently from observed data, the TFM esteemed scores can be seen as points on ratio scales. Moreover, the model is not affected by several sources of misfit. As result of this, it is not undermined by typical perturbations affecting dataset in field studies like the ones carried out in education settings. Among them, it is worth highlighting that TFM does not suffer from missing data: the fact that parameters are measured at the individual observation level means that estimates are obtained only from data that has been observed; consequently, there is no necessity of imputing missing data, or of making assumptions as to the global distribution of parameters.
Other relevant properties of the Rasch model contained in TFM are: conjoint ordering, transitivity, concatenation, and infinite divisibility [2].
In sum, due to its properties, TFM appears to be able to provide a more reliable way of defining an index of students' performance. The TFM index:
a) is not affected by contextual sources of variation (i.e. non-systematic differences as to students, courses and raters, among and within programs, universities and over time);
b) is not affected by missing data, therefore proving to be more suitable for field analyses like the ones performed in the higher education setting;
c) enables researcher to control the occurrence of biased results that are worth interpreting as due to the effect of not pertinent events/factors, however important (this is so as a result of the fact that the information is provided by the position of the subject on the latent variable rather than by the factual observed datum);
d) is able to take into account the failures (this is so as a result of the fact that such information is taken into account in the calculation of parameters β_n, δ_s e $\theta_{i.}$).

3. Case study

The longitudinal case study was designed to esteem predictors of the performance of students of an Italian undergraduate program of psychology. A cohort of students attending an undergraduate program of psychology was monitored for the entire course of their career and semesterly analyses of their career were performed. However, here only one point analysis is reported, the one concerning the impact of early predictors as spread over the student's career.

3.1 Sample

The study has involved the students comprising the whole cohort enrolled in the academic year 2007/08 at the Psychology undergraduate program of the University of Salento (situated in a middle-sized town of south-eastern Italy). More in particular, while the initial number of freshmen was 823, the analysis was limited to the 529 of them that were active at the beginning of the third year (the percentage of active students after the first two academic years - around 65% - is no different from the average of the undergraduate program in Psychology provided in similar socio-cultural contexts; for an analysis of predictors of drop-out in this cohort, see [7], [8]).

3.2 Data analysis

Two regression analyses (*standard method*) were performed, in order to esteem the impact of predictors on higher student performance. Both regressions analyses adopted the same student's variables as candidate predictors:

a) Gender;
b) Age (measured in term of year of birth);
c) The mark of high school leaving qualification;
d) Entry level of General knowledge (contemporary history, literature, constitutional structure, current politics);
e) The entry level of competence in basic English (lexicon, syntax and textual comprehension);
f) The entry level of competence in Reasoning;
g) The entry level of ability in Reading Comprehension.

The global career performance measured at the end of the third year's first semester (April, 2010) was adopted as dependent variable. In Regression analysis 1 the dependent variable was measured in terms of observed data, namely as the count of the examination overcome. In Regression Analysis 2 the dependent variable was esteemed by the TFM. The TFM estimation resulted presenting an acceptable level of fitness (Separation = 5,30; Reliability = 0,97)

3.3 Procedure

Independent variables were measured through an ad hoc paper and pencil questionnaire – the *Questionnaire for the Analysis of the Levels of Competences* (QUALC) – divided into 4 subscales: Reasoning, Reading Comprehension, Basic English, General Knowledge. The QUALC's final session allowed socio-demographic data and the mark of the high school leaving qualification to be collected.

Data concerning the independent variables was collected before the beginning of the program (September, 2007). The measurement was carried out collectively, in a single session taking about 2 hours, as part of the formal procedure freshmen had to used for enrolment.

3.4. Results

Table 1 and 2 report the parameters respectively of Regression analysis 1 (adopting observed data as index of performance), and Regression analysis 2 (using TFM estimation). As Durbin Watson indexes show, models have an acceptable level of adequacy. Both models are significant (both ANOVA shows $p <. 000$). As concerns the predictors, Regression analyses present result that are consistent with each other: from both analyses the Mark of high school leaving qualification, Age (in Year of birth) and Entry competence in basic English prove to be the most important predictors (the only ones in analysis 1). What distinguishes models clearly is the percentage of variance explained, which is much higher for Regression 2 (aR^2=0,415) than Regression 2 (aR^2=0,99). The different capability of explanations of two models has been compared by means of the standard method. It is largely significant (cf. table 3).

Table 1. Predictors of students' performance. Regression analysis 1 (Dependent variable: number of examination at the third year's first semester).

1a. Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	161,753	66,116		2,45	,015
Sex	-,976	,822	-,056	-1,19	,236
Year of birth	-,079	,033	-,110	-2,37	**,018**
Mark of high school leaving qualification	10,324	2,148	,232	4,80	**,000**
Entr. level of general knowledge	1,433	,874	,079	1,64	,102
Entr. level English skill	2,027	,838	,116	2,42	**,016**
Entr. level of Reasoning	-,042	,894	-,002	-,05	,963
Entr. level of Reading comprehensio	1,046	,843	,059	1,24	,216

1b. Model summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
,336	,113	,099	4,743	1,754

1c. ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	1232,238	7	176,034	7,824	,000
Residual	9674,386	430	22,499		
Total	10906,623	437			

Table 2. Predictors of students' performance. Regression analysis 2 (Dependent variable: TFM estimation based on number of examinations in the third year's first semester).

2a. Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	17,674	2,423		7,29	,000
Sex	,003	,030	,004	,11	,914
Year of birth	-,009	,001	-,284	-7,58	**,000**
Mark of high school leaving qualification	,893	,079	,441	11,34	**,000**
Entr. level of general knowledge	,174	,032	,212	5,42	**,000**
Entr. level English skill	,102	,031	,128	3,31	**,001**
Entr. level of Reasoning	,072	,033	,085	2,19	**,029**
Entr. level of Reading comprehensi	,068	,031	,084	2,21	**,027**

2b. Model summary

R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
,651	,424	,415	,174	1,796

2c. ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	9,574	7	1,368	45,268	,000
Residual	12,993	430	,030		
Total	22,567	437			

Table 3. Comparison between the two Regression models.

		R Square	F	R Square Change	F Change	df1	df2	Sig. F Change
Model	Regression 1. Dependent variable: number of exams	*0,113(a)*	*7,824*	*0,311*	*37,444*	*7*	*430*	*<0,001*
	Regression 2. Dependent variable: TSM estimation	*0,424(a)*	*45,268*					

4. Discussion and conclusions

Due to its relevance both for political aims and educational purposes, higher education performances need to be measured in a reliable way. To this end, a method of estimation, based on Rasch analysis, has been discussed, the *Three Facets Model* (TFM). In order to show the utility of TFM estimation in the context of the analysis of predictors of students' performance, preliminary results of a longitudinal case study have been reported. The study, focusing on an undergraduate program in psychology, shows that the TFM measurement of the dependent variable enables a much more powerful estimation of predictors of performance, if compared with the measurement based on observed data.

The basic difference between TFM and the classical approach lies in the fact that in TFM, observed data is not informative per se; rather it is used for esteeming the latent variable, taken as the "truth": the individual score is calculated in terms of her/his position on the latent variable.

> "Statisticians can find it difficult to adjust to Rasch methodology. They tend to believe that the data points tell the truth and that it is the task of statisticians to find models which explain them and to find the latent variables which underlie them. Rasch methodology takes an opposite position. It says that the latent variable is the truth, and when that latent variable is expressed in linear terms, it is the Rasch model that is necessary and sufficient to describe it. Consequently those data points which do not accord with the Rasch model are giving a distorted picture of the latent variable. They may be telling us very important things, e.g., "the students were uninterested", "the scoring key was wrong" - but those do not pertain to the central variable" ([3], p. 15).

References

[1]. Kamsuriah, A., Nazlena, M. A., Suhaila, Z. (2011). An Improved Course Assessment Measurement for Analyzing Learning Outcomes Performance Using Rasch Model. *Procedia Social and Behavioral Sciences*, 18, 442–449.

[2]. Linacre, J.M. (1994). *Many-Facet Rasch Measurement*. 2nd Ed. Chicago: MESA Press.

[3]. Linacre, J.M. (2012). *A User's Guide to FACETS Rasch-Model Computer Programs*. Program Manual 3.70.0. Beaverton, Oregon: Winsteps.com.

[4]. Mouw, J. T., Khanna, R. K. (1983). Prediction of Academic Success: A review of the literature and some recommendations. *College Student Journal*, 27(3), 328-336.

[5]. Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: MESA Press.

[6]. Rohde, T. E., Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, 35(1), 83-92.
[7]. Venuleo, C., Mossi, P. (2011). Modelli di interpretazione dell'ambiente sociale e continuità degli studi universitari. Un caso studio. [Interpretative models of the social environment and retention study. A case study]. *Psicologia scolastica* [School Psychology], 10(2), 171-193.
[8]. Venuleo, C., Mossi, P., Salvatore, S. Meaning and retention study at university. A case study. Submitted manuscript.
[9]. Waugh, R. (2003). Evaluation of quality of student experiences at university using a rasch measurement model. *Studies in Educational Evaluation*, 29, 145-168.