Reliability Analysis Of An Evaluation Rubric For University Accounting Students: A Learning Activity About Database Use

Adelaida Ciudad-Gómez, University Of Extremadura, Spain
Jesús Valverde-Berrocoso, University Of Extremadura, Spain

ABSTRACT

This study aims to analyze the reliability of the internal consistency of a rubric proposal to evaluate an undergraduate university course activity. The analysis is carried out within a model of training and competency assessment.

The study involved 47 second-year students enrolled in the course entitled Statements Business Information as part of their Degree in Finance and Accounting programmed.

To evaluate the internal consistency of their rubrics, the authors have used Cronbach's Alpha. In all, they worked with 72 cases for a total of four analyzed variables.

The results show that the rubric headings proposed have good internal consistency ($\alpha = 0.771$) so that the different items can be said to be interrelated and can therefore be combined into a single total score.

Taking these results into account, it can be concluded that the rubric is a reliable instrument for assessing the achievement attained by students through the use the criteria evaluated by the proposed activity.

Keywords: Rubrics; Competence; Reliability; Financial Accounting; Higher Education; Performance-Based Assessment

INTRODUCTION

Integrating the Spanish education system into the European Higher Education Area (EHEA) has involved a major change in the model of university instruction with a shift to an approach centred on learning and competences.

In this new educational model, summative assessment is giving way to formative assessment, a systematic process of continually obtaining evidence on the learning process (Heritage et al., 2009) and whose core element is feedback; i.e., any activity, information or process facilitating or accelerating learning, either by allowing students to achieve learning results of a high standard which they would not otherwise have achieved, or by allowing them to attain them sooner or more quickly (Hounsell, 2004).

Formative assessment is a process that provides information and support during learning so that lecturers and students can make the necessary adjustments in order to improve performance (Black & William, 1998). Formative assessment is an integral part of teaching and a major source of information for students and lecturers. A good use of feedback enhances students' ability to regulate their own learning (Nicol & Macfarlane-Dick, 2006).

Accordingly, and given the need for instruments to provide such feedback and to serve as tools for formative assessment, one measuring instrument which has proven to be of much use in such assessment is the rubric which, in this past decade, has shown to be a resource allowing integral formative assessment (Conde & Pozuelo, 2007) as well as serving as a guidance tool (Moskal & Leydens, 2000; Mertler, 2001; Hafner & Hafner, 2003; Tierney & Simon, 2004; Wamba et al., 2007).

Though there is no generally agreed definition, rubrics are normally defined as *"a scoring tool for qualitative rating of authentic or complex student work. It includes criteria for rating important dimensions of performance, as well as standards of attainment for those criteria"* (Jonsson & Svingby, 2007, pp. 131), and it is precisely this description of criteria which is credited with the potential of rubrics to help improve students' learning and performance, while facilitating feedback and self-assessment (Jonsson & Swingby, 2007). A rubric has three key features: assessment criteria, a rating scale, and a grading strategy (Popham, 1997).

In this pedagogical context, the authors have designed a competence-based instruction and assessment model for the university environment called *MANagement of COMpetence in the areas of Accounting* [MANCOMA] (Ciudad & Valverde, 2012), standardising the competences to be developed by students, providing rubrics and creating face-to-face and online activities for competence acquisition and assessment, implemented in an organised way in the *Moodle* platform (Ciudad et al., 2014).

One of these activities is ACT-3, "*Obtaining the annual accounts of a company in the database and analysing the annual accounts obtained*" - a group task offered to students and divided into two phases, so two rubrics were created - one for phase A and another for phase B.

The aim of this paper is thus to analyse the reliability in terms of internal consistency of the two proposed rubrics for assessing ACT-3 within this model of competence-based instruction and assessment.

METHOD

The reliability of the rubric was evaluated in terms of its internal consistency (IC) for which the authors used *Cronbach's alpha*, a statistic widely employed for quantifying the degree to which the various items of an instrument are correlated with each other and for which calculations were made using the SPSS statistical package (version 19.00 for Windows).

This index, which takes values between 0 and 1 and for which 0.70 is generally regarded as the minimum acceptable value (George & Mallery[1], 1995), allows us to ascertain whether the instrument being evaluated – in this case, the activity rubrics – compiles defective information, and so would lead us to erroneous conclusions, or is a reliable instrument making stable and consistent measurements.

Data for the study were gathered with the participation of 47 Finance and Accounting degree students enrolled at Extremadura University's Faculty of Business Studies and Tourism in the subject *"Statements Business Information* for the academic year 2013-2014.

For the group activity, the students were asked to organise themselves in a total of eight working groups, and the authors used self-assessment and peer assessment, so in calculating the rubrics' reliability, the authors considered the scorings given by the student groups in the rubric at the end of the activity assessment process and also the lecturer's scorings for each group.

In all, they worked with 72 cases (eight cases of self-assessments, 56 of peer-assessments and eight of assessments make by the teacher) for a total of four analyzed variables.

[1] George and Mallery (2003: 231) provide the following rules of thumb: *"α > .9 – Excellent, α > .8 – Good, α > .7 – Acceptable, α> .6 – Questionable, α > .5 – Poor, and α < .5 – Unacceptable".*

For the self-assessment and peer assessment, the authors used the GTEA *e-Rubric* tool (https://gteavirtual.org/rubric/).

RESULTS

In analysing their rubrics' degree of internal consistency, the authors obtained a *Cronbach's alpha* coefficient of α=0.774 (n of items: 2) for the rubric used in phase A of the activity and a *Cronbach's alpha* coefficient of α=0.818 (n of items: 2) for phase B, with a confidence level of 95% (p≤0.05).

The *Cronbach's alpha* value obtained for these two rubrics exceeds the limit set at 0.7, indicating that the authors' instrument has a good degree of reliability and validating its use for data collection.

To give the data-gathering instrument greater reliability, *Cronbach's alpha* was also calculated jointly for the phase A and phase B rubrics, and again the instrument showed satisfactory reliability with an *alpha* of 0.761 (n of items: 4) and a confidence level of 95% (p≤0.05) as shown in Table 1.

Table 1: Reliability Analysis Of Rubric: Phase A & B (Cronbach's Alpha)

Ítems	ID Indicator[i2]	Scale Mean If Item Deleted	Scale Variance If Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha If Item Deleted
ITEM3.1.	**CGI05.1.C**	**179,51**	**3456,451**	**,437**	**,779**
ITEM3.2.	CGS22.2.D	189,24	3271,591	,691	,637
ITEM3.3.	CGI05.2.A	185,42	3164,613	,536	,724
ITEM3.4.	CGI06.2.B	182,29	3722,491	,645	,679
Cronbach's Alpha (0,761)		Cronbach's Alpha Based on Standardized Items (0, 761)			N of Items (4)

To account for the number of items included in the rubric, the authors applied the *Cronbach's alpha* internal consistency method in the event of an item being deleted.

The *"Corrected Item-Total Correlation"* is the corrected homogeneity coefficient and if it is zero or negative, it should be deleted. In this case, none needed to be deleted. Moreover *"Cronbach's Alpha if Item deleted"* is equivalent to the *alpha* value if one of the items is deleted. Thus, the authors can make sure that once the item showing the weakest item-total correlation has been deleted, its deletion does not lead to obtaining significantly higher values than the *alpha* value for the rubric as a whole.

This indicator is CGI05.1.C (item 1.3.), which rates whether "*The student is able to make selective and advanced searches in databases and obtain quantitative data, selecting those which are key to the specific subject being addressed and the user's information needs*", and if it is deleted the *alpha* value would rise from 0.761 only to 0.779.

CONCLUSIONS

In their reliability study on the rubric proposed for the assessment of ACT-3 in the MANCOMA model, based on the *Cronbach's alpha* coefficient, the authors obtained an acceptable reliability coefficient allowing them to assert that their rubric has adequate internal consistency, meeting one of the conditions for reliability and thereby ensuring that scorings will not vary as a function of the instrument; i.e., that the grades given by the assessors are consistent and converge on a common construct to be evaluated.

[2] The « **ID indicator** » identifies the competence, capability and specific indicator (knowledge, skills and values) to develop. Each indicator is identified by an alphanumeric ID: AAABB.C.D, where:

- **AAA.BB** represents the competence number and type according to the verified degree report **(**CGI05: Instrumental competence**:** Elementary computing skills**;** CGS22: Systemic competences. Concern for quality**;** and CGI06: Instrumental competence: Ability to analyze and seek information from various sources).
- **C & D:** "C" represents the capability within the competence and "D" represents the indicator within the capability, a breakdown made in the competence standardisation phase.

Moreover, the authors see that if the item showing the weakest item-total correlation (item 3.1) is deleted, they would get a higher *alpha* value for the rubric as a whole, though not significantly higher, and as it is worth having this indicator for assessing the competences involved, they are going to rewrite it on the basis of these results so as to avoid any ambiguity which could impair assessment.

As to limitations found in their research, the authors should note the circumstance that students were encountering rubric assessment for the first time and so were not familiar with the concept, and this lack of experience gave rise to a certain degree of difficulty in peer assessment and self-assessment.

The authors should also note that their research is in a developmental phase and so here they are offering a set of interim results, restricted to a single subject, *Statements Business Information*, of a single degree, *Degree in Finance and Accounting*; so in future phases, the study should be widened to include other subjects and degrees, which would allow results to be determined with a higher degree of reliability.

Moreover, on a complementary basis, the authors propose obtaining and analysing information on students' perceptions of and attitudes on the use of rubrics as assessment instruments and their usefulness.

ACKNOWLEDGMENTS

This work has been funded by the Regional Government of Extremadura in Spain and the European FEDER program.

AUTHOR INFORMATION

Adelaida Ciudad-Gomez has a Degree in Business Administration by University of Sevilla and is a Professor in Accounting and Finance Department, Faculty of Business Studies and Tourism of Cáceres at the University of Extremadura (Spain). Research Interests include Accounting, educational, tourism and environment. E-mail: adelaida@unex.es.

Jesús Valverde-Berrocoso has a PhD in Pedagogy (University of Salamanca) and is a Professor in the Sciences Education Department, Faculty of Teacher Education at the University of Extremadura (Spain). He is currently Director of the Virtual Campus of the University of Extremadura (CVUEX) and Director of «Shared Virtual Campus» G9 Group of Universities (CVC-G9). Research interests include ICT in Education, E-learning in Higher Education, Instructional Design, and Technology. E-mail: jevabe@unex.es.

REFERENCES

1. Black, P. & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74. Retrieved from: http://wiki.biologyscholars.org/@api/deki/files/1278/=black_p_etal_1998b%255b1%255d.pdf [accessed January 2013].
2. Ciudad-Gomez, A. & Valverde-Berrocoso, J. (2012): "Design of a competency-based assessment model in the field of Accounting ". Contemporary Issues in Education Research (CIER), Vol 5, No 5, 343-348. Retrieved from: http://journals.cluteonline.com/index.php/CIER/article/view/7466/7532 [accessed December 2012].
3. Ciudad- Gómez, A., Valverde-Berrocoso, J., & Coca-Pérez, J. L. (2014). Integration of Virtual Teaching/Learning Environments in Higher Education for the Development of Formative Assessment in the Field of Accounting. En Peris-Ortiz, M., Garrigós-Simón, F. J., & Pechuán, I. G. (Coord.). *Innovation and Teaching Technologies*, Springer, pp 23-30.
4. Conde, A. & Pozuelo, F. (2007). Las plantillas de evaluación (rúbrica) como instrumento para la evaluación. Un estudio de caso en el marco de la reforma de la enseñanza universitaria en el EEES. *Investigación en la Escuela, 63*, 77-90. Retrieved from: http://www.investigacionenlaescuela.es/articulos/63/R63_6.pdf [accessed January 2013].
5. George, D. & Mallery, P. (1995). *SPSS/PC+step by step: A simple guide and reference.* Estados Unidos:

Wadsworth Publishing Company.

6. Hafner, J. C. & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *International Journal of Science Education* 25 (12), 1509-1528.
7. Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practices, 28(3)*, pp- 24-31.
8. Hounsell, D. (2004). Student feedback, learning and development. En Slowey, M. & Watson, D. (ed). *Higher Education and the Lifecourse*. Buckingham: SHRE & Open University Press, pp. 67-78.
9. Jonsson, A. & Svingby, S. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*, pp. 130–144 Retrieved from: https://facultycommons.macewan.ca/wp-content/uploads/Rubrics-Jonasson-and-Svingby-20071.pdf [accessed January 2013].
10. Mertler, C.A. (2001). Designing scoring rubrics for your classroom. Practical Assessment, Research & Evaluation, 7 (25). Retrieved from: http://pareonline.net/getvn.asp?v=7&n=25 [accessed January 2013].
11. Moskal, B. M. & Leydens, J .A. (2000). Scoring rubric development: validity and reliability. Practical Assessment, Research & Evaluation, 7(10). Retrieved from: http://pareonline.net/getvn.asp?v=7&n=10 [accessed January 2013].
12. Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. Retrieved from: http://www.lyderiulaikas.smm.lt/Atsisi%C5%B3sti%20failus:/article/1044/Formative%20assessment%20and%20self-regulated%20learning.pdf [accessed January 2013].
13. Popham, W. J. (1997). What's Wrong--and What's Right--with Rubrics. *Educational Leadership, 55*(2), 72-75.
14. Tierney, R. & Simon, M. (2004). What's still wrong with rubrics: focusing on the consistency of performance criteria across scale levels. Practical Assessment, Research & Evaluation, 9(2). Retrieved from: http://pareonline.net/getvn.asp?v=9&n=2 [accessed January 2013].
15. Wamba, A. M., Ruiz Aguaded, C., Climent, N., & Ferreras, M. (2007). Las rúbricas de evaluación de los Practicum como instrumento de reflexión para los estudiantes de Educación Primaria. En A. Cid *et* al (Coord.). *Buenas Prácticas en el Practicum.* AIDU (Asociación Iberoamericana de Didáctica Universitaria); U. de Vigo, U. de Santiago de Compostela. (Actas del IX Simposium Internacional sobre prácticas. Practicum y prácticas en empresas y en la formación universitaria. Poio, Pontevedra), 1251-1261.

NOTES

