

COMPARATIVE STUDY OF RUBRICS FOR ASSESSMENT OF LABORATORY PRACTICES OF BIOCHEMISTRY. A PERSPECTIVE OF STUDENTS FROM DIFFERENT UNIVERSITY DEGREES

M. D. Busto, N. Ortega, S. Ramos-Gómez, S. M. Albillos, M. C. Pilar-Izquierdo, D. Palacios

Area of Biochemistry and molecular Biology, Faculty of Science, University of Burgos (SPAIN)

Abstract

Definitions for the word “rubric” abound, but basically, the rubrics are documents that describe varying levels of performance from excellent to poor and show where on that scale a student is achieving a particular learning standard, goal or objective. The use of a rubric with specific criteria and proficiency levels is highly desirable for evaluating a multidimensional task. When used as teaching tools, rubrics not only make the instructor’s standards and resulting grading explicit, but they can give students a clear sense of what the expectations are for a high level of performance on a given assignment, and how they can be met. According to some authors the use of rubrics can be most important when the students are novices with respect to a particular task. Some authors also indicated that using rubrics may not improve the reliability or validity of assessment if raters are not well trained on how to design and employ them effectively.

The purpose of this study was to compare the effectivity of different rubrics designed as tools for assessment of students in laboratory practices of Biochemistry from diverse university degrees (Chemistry, Food Science and Technology, and Agricultural Engineering) during the academic year 2013-14. Another intention was also pretended to determine the style of rubric more adequate for the students depending on the level (first, second or third course). In this experience, three rubrics were analysed by means of students' satisfaction surveys. Moreover, students' perspective about the formative activity was also revised. Students were asked to make an assessment of both the formative activity carried out and the use of rubric for their evaluation. Each questionnaire contained several items that had to be rated from 1 to 5 (ranking the degree of satisfaction as 1 not all; 2, poor, 3 average, 4 fair and 5 high). Resulting data showed that students independently of the level or degree, considered positively the use of rubrics as a useful tool for the development and establishment of evaluating criteria for the formative activity. Furthermore, these students also valued globally the practices of laboratory as very satisfactory or satisfactory. Specific differences in the students' responses to the items of the survey between the different levels and studies were analysed and will help redesign the rubrics and improve the formative activity for the future.

Keywords: Assessment, Biochemistry, laboratory practice, rubric.

1 INTRODUCTION

Rubrics are documents that articulate the expectations of an assignment by listing the criteria for what is particularly important and by describing levels of quality on a scale from excellent to poor [1]. The use of a rubric with specific criteria and proficiency levels is highly desirable for evaluating a multidimensional task [2]. When used as teaching tools, rubrics not only make the instructor’s standards and resulting grading explicit, but they can give students a clear sense of what the expectations are for a high level of performance on a given assignment, and how they can be met [1].

A number of studies on the effects of rubrics on learning and performance have found a positive effect from their use when, for example, are applied for long enough [3]. The rubrics are also considered as an approach to implementing self-assessment in the classroom [4]. According to some authors the use of rubrics can be most important when the students are novices with respect to a particular task [5]. On the other hand, it has also been suggested that using rubrics may not improve the reliability or validity of assessment if raters are not well trained on how to design and employ them effectively [6].

An important aspect in the study of rubrics is the student's perception about their use. Many studies have found positive results, with rubric use decreasing anxiety for students and helping them feeling

more secure among other effects [7] and also being better when compared to another type of assessment tool [8].

The purpose of this study was to compare the effectivity of different rubrics designed as tools for assessment of students in laboratory practices of Biochemistry from diverse university degrees (Chemistry, Food Science and Technology and, Agricultural Engineering) during the academic year 2013-14. Another objective was to determine the style of rubric more adequate for the students depending on the level (first, second or third course). In this experience, three rubrics were analysed by means of students' satisfaction surveys. Moreover, students' perspective about the formative activity was also revised. Students were asked to make an assessment of both the formative activity carried out and the use of rubric for their evaluation.

2 METHODOLOGY

2.1 Participants and procedure

The research has focused on the training activity "laboratory practices" of Biochemistry during one academic year (2013-14). This subject is taught in first, second and third year in the degrees of Food Science and Technology, Agricultural Engineering and Chemistry, respectively.

A total of 129 students took part in this study, 27 from Chemistry, 25 from Agricultural Engineering and 77 from Food Science and Technology.

The study comprised the following steps:

- Design the rubric to assess the training.
- Explanation of the rubric by the teacher in the classroom at the beginning of the formative activity. The rubric was permanently available for students in the virtual platform based in Moodle (UBUVirtual).
- Design of the satisfaction surveys related to the formative activity (practices of laboratory) and the rubrics.
- Distribution of the surveys to the students once the activity was finalized.
- Statistical analysis of the data.

2.3 Instruments

2.3.1 Rubrics

Two kinds of rubrics were designed taking into account the different levels of students. Two of the rubrics were very similar, for first and second degree courses (Food Science and Technology and Agriculture Engineering), and another for the students of the third course (Chemistry).

The general criteria included in these rubrics to assess the laboratory training were:

- The experimental work in the laboratory (student's performance in the laboratory).
- Formal aspects of the report (report structure, formal writing, neat presentation...).
- Introduction, objectives and experimental procedure (brief description of objectives, theory, experimental protocol...).
- Data processing (explanation of calculations, data organized in tables and charts, correct calculations, graphs with units in the axis...).
- Interpretation of results (correct reasoning and analysis of the results, achievement of objectives, correct conclusions...).
- Bibliography (to cite the proper references and connect them with the practice).
- Additional questions (answer correctly any additional question asked related to the practice).

These criteria were evaluated at four gradations of quality: excellent, good, sufficient, and insufficient.

The fundamental difference between the rubrics laid on a briefer description and specification of each item for the students of the third course (Chemistry degree). In addition, as shown in Table 1, the score of the experimental work in laboratory had a greater weight for the students of the third year

Table 1. The score of criteria included in the rubrics to assess the laboratory training of Biochemistry, for the university degrees of Food Science and Technology (FST), Agricultural Engineering (AE) and Chemistry (CHM).

CRITERIA	SCORE (%)		
	FST (1st) ^a	AE (2nd)	CHM (3rd)
The experimental work in the laboratory	20	20	40
Formal aspect of the report	5	5	5
Introduction, objectives and experimental procedure	15	15	15
Bibliography	5	20	
Interpretation of results	10		40
Data processing	30	25	
Additional questions	15	15	

^a course

2.3.2 Surveys

Students were asked to make an assessment of both the use of a rubric for their evaluation and the formative activity itself. Each questionnaire contained several sentences (items) that had to be rated from 1 to 5 (ranking the degree of satisfaction as 1 not at all; 2 poor; 3 average; 4 fair and 5 high). The specific items presented for consideration were detailed in Table 2.

Table 2. Items included in the student satisfaction survey to assess the rubric and the formative activity (laboratory practices of Biochemistry)

RUBRIC ASSESSMENT
RA: The rubric is a useful tool
RB: The rubric may be simplified
RC: The rubric may be more detailed
RD: Rate the extent of support provided by the rubric in the elaboration of the final report
RE: Rate the degree of global satisfaction with the evaluation criteria for the formative activity
FORMATIVE ACTIVITY ASSESSMENT
FAA: The objectives and working methodology have been clearly defined
FAB: The documentation provided by the professor has been suitable
FAC: Rate the degree of difficulty in the development of the activity
FAD: The timing of the activity (schedule of deliverables) has been adequate
FAE: Rate the degree of knowledge acquired with the activity
FAF: Rate if this activity has helped you develop other skills.
FAG: Rate the degree of global satisfaction with this formative activity

2.3.3. Statistical

The surveys were first analysed by descriptive statistic, calculating the mean values, the standard deviation (SD), the asymmetry coefficient and kurtosis. Once the basic properties of the variables were described, comparison of weighted satisfaction percentages between rubrics was performed, calculated considering the ratings and the frequency of responses in each category and statistical significance was determined using the Wald test with continuity correction [9]. All analysis was conducted using Excel and p -value of less than 0.05 was considered as statistically significant unless otherwise stated.

3 RESULTS AND DISCUSSION

3.1 Descriptive statistics

All data obtained from the survey were analyzed in a global manner, firstly for the whole group of students and secondly taking into account the subgroups related to the different university degrees. The descriptive statistics used were the mean value, the standard deviation (SD), asymmetry and kurtosis for each of the posed questions in each of the assessment questionnaires related both to the rubric (Table 3) and the formative activity (laboratory practices) (Table 4).

Table 3. Descriptive statistics of the items for the assessment survey of rubrics for the laboratory practices of Biochemistry of different degrees and courses^a.

		Statistics				
		Valid N	Mean	S.D.	Asymmetry	Kurtosis
RA	T ^b	129	3.43	0.84	-0.40	0.11
	FST	77	3.51	0.79	-0.35	0.53
	AE	27	3.52	0.75	-0.07	-0.11
	CHM	25	3.12	1.01	-0.26	-0.89
RB	T	129	2.48	1.08	0.24	-0.65
	FST	77	2.43	1.06	0.16	0.89
	AE	27	2.74	1.06	-0.06	-0.41
	CHM	25	2.36	1.15	0.82	0.44
RC	T	128	3.27	1.05	-0.26	-0.45
	FST	77	3.32	1.02	-0.54	0.01
	AE	27	3.04	0.98	0.19	-0.22
	CHM	24	3.33	1.20	-0.06	-1.06
RD	T	128	3.15	0.87	0.00	-0.10
	FST	76	3.28	0.84	0.12	-0.59
	AE	27	3.00	0.92	-0.32	0.55
	CHM	25	2.92	0.86	0.16	0.66
RE	T	128	3.12	0.78	-0.31	0.08
	FST	76	3.11	0.79	-0.52	0.56
	AE	27	3.11	0.89	0.12	-1.04
	CHM	25	3.16	0.62	-0.11	-0.27

^aFST: Food Science and Technology (1st); AE: Agricultural Engineering (2nd); CHM: Chemistry (3rd)

^bT: Total

With respect to the mean satisfaction of the students with the rubrics designed for the three university degrees, the highest was for the item RA ("the rubric is a useful tool"), and the lowest was for item RB ("the rubric may be more detailed") (Table 3). The item RB showed the highest values of standard deviation with the RC, while the item RA was the second lowest. The negative asymmetry indicates that the punctuation in the assessments had a tendency towards the right, namely towards a higher degree of satisfaction. The results were slightly different depending on the university degree evaluated. Thus, asymmetry value was negative for the RA, RC and RE items in the university degrees of FST and CHM, and for RA, RB y RD in the AE degree. The positive kurtosis obtained for the different items indicate a leptokurtic distribution of data, meaning that most of the ratings were distributed surrounding the central values. This behavior was obtained for the RA, RB, RC and RE in FST degree, RB and RC in CHM degree and for the RD item in AE.

On the other hand, the negative asymmetry and positive kurtosis demonstrate a higher and more generalized satisfaction degree (e.g. RA, RC and RD items in FST degree, and RD in AE degree) and a positive asymmetry and a negative kurtosis indicate a shift of the ratings towards a lower degree of satisfaction (e.g. RD in FST and RC and RE in AE degree).

In general, the mean satisfaction degree of the students with the formative activity (laboratory practices of Biochemistry) (Table 4) was higher than for the rubric. The asymmetry and kurtosis values indicated a generalized and high level of satisfaction for total (T), except for question FAC. As in the previous study about the rubric (Table 3) the results also depend on the evaluated university degree, but with slight differences.

Table 4. Descriptive statistics of the items for the assessment survey of the formative activity (laboratory practices of Biochemistry) of different degrees^a.

		Statistics				
		Valid N	Mean	S.D.	Asymmetry	Kurtosis
FAA	T ^b	129	3.62	0.76	-0.63	0.63
	FST	77	3.61	0.73	-0.71	0.21
	AE	27	3.44	0.80	-1.02	2.23
	CHM	25	3.84	0.80	-0.22	-0.28
FAB	T	130	3.55	0.87	-0.56	0.15
	FST	78	3.53	0.80	-0.63	0.49
	AE	27	3.33	1.04	-0.30	-0.43
	CHM	25	3.88	0.83	-0.70	0.51
FAC	T	126	3.46	0.83	0.17	-0.05
	FST	74	3.55	0.78	0.44	-0.47
	AE	27	3.44	0.89	-0.35	1.14
	CHM	25	3.20	0.87	0.42	-0.19
FAD	T	129	3.67	0.91	-0.41	1.11
	FST	78	3.49	0.88	-0.67	0.43
	AE	26	3.58	0.64	0.67	-0.43
	CHM	25	4.32	0.99	-1.28	0.53
FAE	T	123	3.66	0.72	-0.71	1.11
	FST	73	3.73	0.58	-0.31	0.23
	AE	26	3.31	1.01	-0.43	-0.38
	CHM	24	3.83	0.64	0.14	-0.37
FAF	T	118	3.56	0.72	-0.49	0.72
	FST	68	3.62	0.67	-0.91	2.46
	AE	26	3.27	0.78	0.03	-0.34
	CHM	24	3.71	0.75	-0.12	-0.06
FAG	T	125	3.58	0.80	-0.45	0.76
	FST	75	3.61	0.71	-0.64	1.56
	AE	27	3.11	0.75	-0.78	1.10
	CHM	23	4.00	0.85	-0.48	-0.29

^aFST: Food Science and Technology (1st); AE: Agricultural Engineering (2nd); CHM: Chemistry (3rd)

^bT: Total.

3.2 Comparison of weighted satisfaction percentages. Wald test with continuity correction

Global percentages, calculated as the sum of students rating each item 3, 4 and 5 compared to the total number of students are described in Table 5. These results suggest that the students of the 1st and 2nd course valued the rubrics as a useful tool, 92.2 and 92.6%, respectively, in comparison with the students of 3rd course (68%).

The overall satisfaction, for the FST, AE and CHM degrees, with the proposed learning activity (laboratory practices of Biochemistry) was 96, 85.2 and 95.7%, respectively, and 82.9, 70.4 and 88.0% with the evaluation criteria, respectively. Nevertheless, these high values, obtained by global estimations, without any *post hoc* analysis, provided limited information.

Table 5. Global percentages of the items for the assessment survey of the formative activity (laboratory practices of Biochemistry) of different degrees^a

Global percentage (%) ^b							
Rubric assessment				Formative activity assessment			
Criteria	FST	AE	CHM	Criteria	FST	AE	CHM
RA	92.2	92.6	68.0	FAA	90.9	92.6	96.0
RB	48.1	63.0	40.0	FAB	98.7	77.8	92.0
RC	81.8	70.4	70.8	FAC	95.9	92.6	80.0
RD	81.6	77.8	72.0	FAD	87.2	100.0	92.0
RE	82.9	70.4	88.0	FAE	98.6	76.9	100.0
				FAF	97.1	84.6	95.8
				FAG	96.0	85.2	95.7

^aFST: Food Science and Technology (1st); AE: Agricultural Engineering (2nd); CHM: Chemistry (3rd). ^bSum of students rating each item 3, 4 and 5 compared to the total students.

The outcome of the previous section suggested that it could be desirable to perform a comparison of the weighted satisfaction percentages for each of the items of the survey, taking also into account the different degrees, in order to be able to find significant differences amongst the higher levels of satisfaction. Composite satisfaction percentages calculated as weighted percentages of students that rated a degree of satisfaction of 3 or above (highly satisfied) for each of the items posed and differentiated by university degrees are presented in Fig. 1 and 2. Specifically, the Wald test was applied to the results of surveys of first and second year, grouped together, corresponding to the degrees of SFT and AE, compared to the superior course, 3rd of the degree of CHM.

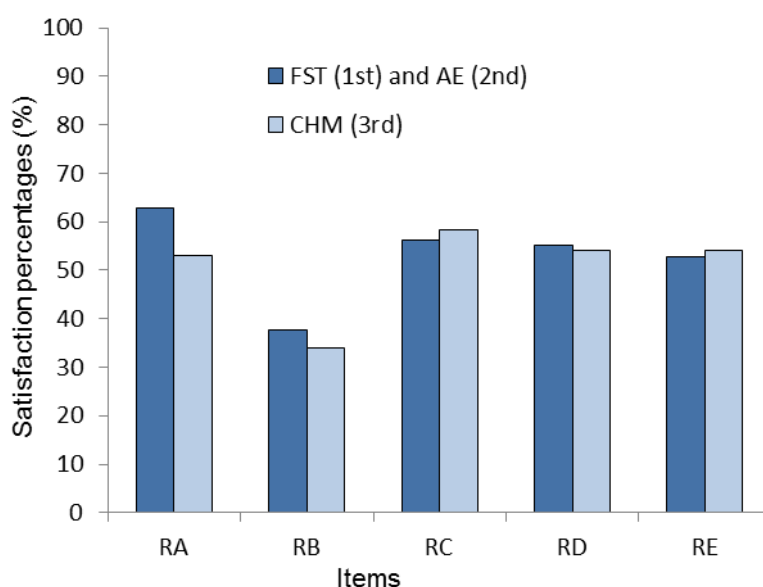


Fig. 1. Weighted percentages of students, differentiated by university degrees (FST- Food Science and Technology; AE- Agricultural Engineering and CHM- Chemistry) and course (1st, 2nd and 3rd), rating 3 or above (highly satisfied) for each sentence related to the rubric.

It can be seen in Fig. 1, that item RB had the lower rating in terms of satisfaction. Between 50 and 60% of students expressed a high or average level of satisfaction for the rest of the items related to the evaluation of the rubric, with RA scoring the highest percentage for students in first and second year, and RC for the students of third year.

With regard to the formative activity, Fig. 2 shows that between 55 to 83% of the students were "above average satisfied" or "very satisfied" when rating the laboratory practices of Biochemistry. The degree of satisfaction with the training activity was greater for students of the third course of Chemistry. Moreover, these students indicate a lower degree of difficulty in the development of the activity" (item FAC).

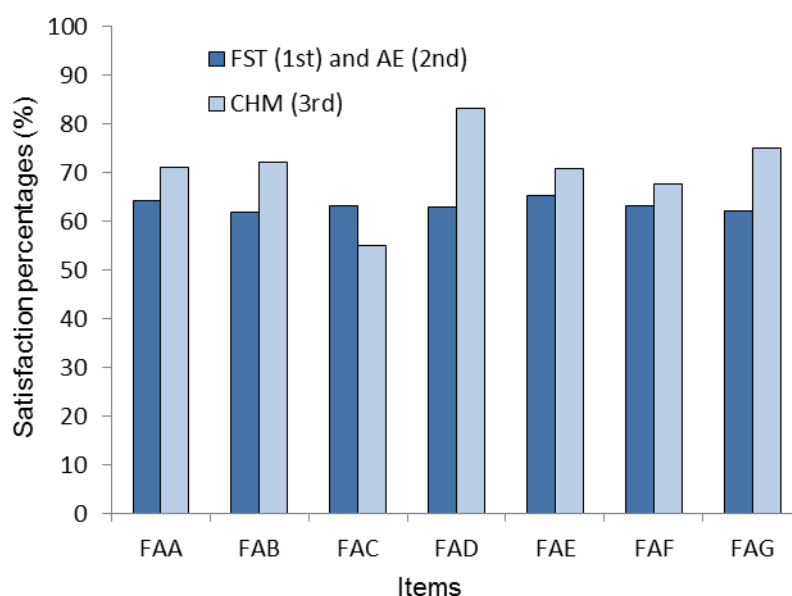


Fig. 2. Weighted percentages of students, differentiated by university degrees (FST- Food Science and Technology; AE- Agricultural Engineering and CHM- Chemistry) and course (1st, 2nd and 3rd), rating 3 or above (highly satisfied) for each sentence related to the formative activity (laboratory practices of Biochemistry).

For determining if there was a statistical significance between the weighted percentages obtained for each item comparing the responses obtained for both groups of level, a corrected Wald test (with continuity correction) was applied taking into account the responses of those with higher degrees of satisfaction (rating 3 or above). Results of the corrected Wald test (not shown) concluded that there were no significant differences for the responses obtained for each of the items, except for the item FAD ("the timing of the activity (schedule of deliverables) has been adequate").

4 CONCLUSIONS

Resulting data showed that students valued globally the formative activity, laboratory practices of Biochemistry, as very satisfactory or satisfactory, independently of the level or degree. Furthermore, these students also considered positively the rubric as a useful tool for the development and establishment of evaluating criteria for the formative activity, especially those of first and second year. On the other hand, the main difficulty observed in this study was related to the understanding of the surveys in relation to some aspects/details in some items (e.g RC) and, in the meaning of a sentence with a negative adjective (e.g. FAC). This indicated that the surveys should be modified, with these items removed or rephrased before addressing changes of the rubrics.

Acknowledgements: We thank financial support by Vicerrectorado de Profesorado y de Personal de Administración y Servicios de la Universidad de Burgos (Proyecto de Innovación Educativa y Mejora Docente 2012/2015).

REFERENCES

- [1] Panadero, E. and Romero, M. (2014). To rubric o not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), pp. 133-148.
- [2] Allen, D. and Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE-Life Sciences Education*, 5, pp. 197-203.
- [3] Panadero, E. and Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review, *Educational Research Review*, 9, pp. 129-144.
- [4] Taras, M. (2012). Student self-assessment: Processes and consequences. *Teaching in Higher Education*, 15, pp. 199-209.
- [5] Bresciani, M. J., Zelna, C. I. and Anderson, J. A. (2004). Criteria and rubrics. In: *Assessing Student Learning and Development: A Han-book for Practitioners*, Washington, DC: National Association of Student Personnel Administrations, pp. 29-37.
- [6] Rezaei, A. L. and Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15, pp. 18-39.
- [7] Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment Research & Evaluation*, 15(8). Retrieved from <http://pareonline.net/getvn.asp?v=15&n=8>.
- [8] Panadero, E., Alonso-Tapia, J. and Huertas, J. A. (2014). Rubrics vs-self-assessment scripts. Effects on first year university students' self-regulation and performance. *Infancia y Aprendizaje*, 37(1), 125-132.
- [9] Correa, J. C. and Sierra, E. (2003). Intervalos de confianza para la comparación de dos proporciones. *Revista Colombiana de Estadística*, 26(1), pp.61-75.