



<Research article>

Rubrics and in-class feedback on a higher education course with statistical content

Roser Bono¹ , María Isabel Núñez-Peña² 

Sent: 27/07/2018. Accepted: 02/11/2018. Published in press: 11/12/2018. Published: 07/01/2019

//Abstract

INTRODUCTION: This study analyzed the influence of two types of feedback (via rubrics and in-class) on students' academic achievement in a higher education course with statistical content. Students' views regarding the usefulness of these types of feedback were also examined.

METHODS: After validating the rubrics in a sample of 100 students, a sample of 135 students was used to address the main objective of the study. The samples comprised undergraduates from the University of Barcelona who were taking a course called Research Designs as part of the University of Barcelona's Degree in Psychology.

RESULTS: The majority of students regarded both types of feedback as being useful. Academic achievement did not differ according to whether students used rubrics, attended feedback classes, or made use of both types of feedback. However, there was a positive association between the frequency of attendance at feedback classes and academic performance. Finally, the students who reported feeling less anxious about exams as a result of using rubrics or attending feedback classes obtained better exam grades.

DISCUSSION: This study shows that teaching approaches which include feedback may lead students to have a more favourable view of their learning, especially when the feedback is more personalized and given in class.

//Keywords

Rubrics; In-class feedback; Academic achievement; Test anxiety; Higher education.

//Authors' information

¹ Universitat de Barcelona, Facultat de Psicologia. Corresponding author: rbono@ub.edu

² Universitat de Barcelona.

//Recommended reference

Bono, R., y Núñez-Peña, M. I. (2019). Rubrics and in-class feedback on a higher education course with statistical content. *REIRE Revista d'Innovació i Recerca en Educació*, 12(1), 1–14. <http://doi.org/10.1344/reire2019.12.122560>

© 2018 Roser Bono *et al.* This is an open access article distributed under the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>





//Títol

Rúbriques i *feedback* a classe en un curs amb contingut estadístic en educació superior

//Resum

INTRODUCCIÓ: Amb aquest estudi es va analitzar la influència de dos tipus de *feedback* (a través de rúbriques i a classe) sobre el rendiment acadèmic en una assignatura amb contingut estadístic. També es va examinar l'opinió dels estudiants sobre la utilitat d'aquests dos tipus de *feedback*.

MÈTODE: Després de validar les rúbriques amb una mostra de 100 estudiants, es va utilitzar una altra mostra de 135 estudiants per a l'objectiu principal de l'estudi. Les mostres estaven formades per estudiants de Disseny de Recerca del grau de Psicologia de la Universitat de Barcelona.

RESULTATS: La majoria d'estudiants va valorar positivament la utilitat dels dos tipus de *feedback*. No es van observar diferències en el rendiment acadèmic entre els estudiants que van utilitzar les rúbriques, els que van assistir a les classes de *feedback* i els que van utilitzar els dos tipus de *feedback*. No obstant això, es va observar una associació positiva entre la freqüència d'assistència a les classes de *feedback* i el rendiment. Finalment, els estudiants que van percebre que l'ús de les rúbriques o de les classes de *feedback* ajudaven a reduir la seva ansietat davant els exàmens van obtenir millors notes.

DISCUSSIÓ: Aquest estudi mostra que els processos d'ensenyament amb *feedback* poden portar els estudiants a tenir una visió més favorable del seu aprenentatge, especialment quan el *feedback* es dona a classe de manera personalitzada.

//Paraules clau

Rúbriques; *Feedback* a classe; Rendiment acadèmic; Ansietat als exàmens; Educació superior.

//Título

Rúbricas y *feedback* en clase en un curso con contenido estadístico en educación superior

//Resumen

INTRODUCCIÓN: Este estudio analiza la influencia de dos tipos de *feedback* (a través de rúbricas y en clase) sobre el rendimiento académico en una asignatura con contenido estadístico. También se examinó la opinión de los estudiantes sobre su utilidad.

MÉTODO: Después de validar las rúbricas con una muestra de 100 estudiantes, se utilizó otra muestra de 135 estudiantes para el objetivo principal del estudio. Las muestras estaban formadas por estudiantes de Diseños de Investigación del grado de Psicología de la Universidad de Barcelona.

RESULTADOS: La mayoría de estudiantes valoró positivamente la utilidad de ambos tipos de *feedback*. No se observaron diferencias en el rendimiento académico entre los estudiantes que utilizaron las rúbricas, los que asistieron a las clases de *feedback* y los que usaron ambos tipos de *feedback*. Sin embargo, se observó una asociación positiva entre la frecuencia de asistencia a las clases de *feedback* y el rendimiento. Finalmente, los estudiantes que percibieron que el uso de las rúbricas o las clases de *feedback* ayudaban a reducir su ansiedad ante los exámenes obtuvieron mejores notas.

DISCUSIÓN: Este estudio muestra que los procesos de enseñanza que incluyen *feedback* pueden llevar a los estudiantes a tener una visión más favorable de su aprendizaje, especialmente cuando el *feedback* se da de manera más personalizada, en clase.

//Palabras clave

Rúbricas; *Feedback* en clase; Rendimiento académico; Ansiedad ante los exámenes; Educación superior.

1. Introduction

In recent years, formative assessment has become an important part of the teaching and learning process. A defining characteristic of formative assessment is the timely correction of errors, meaning that the feedback which students receive about their learning activities is given early enough to allow them to improve (Brown, Peterson & Yao, 2016; Tang & Harrison, 2011). When this is the case, students have the opportunity to revise and improve their work prior to sitting exams, and this can have a notable impact on their academic achievement. Recent research has not only highlighted the importance of feedback for improving learning in general (Brown *et al.*, 2016; Evans, 2013; Hodgson & Pang, 2012; Núñez-Peña, Bono & Suárez-Pellicioni, 2015; Panadero & Jonsson, 2013; Pereira, Flores, Veiga-Simão & Barros, 2016) but has shown that in-class feedback is especially useful for students who experience high levels of anxiety when confronted with statistical course content (Núñez-Peña *et al.*, 2015).

Over the last three decades a number of meta-analyses and reviews have aimed to determine which types of feedback are most effective for improving learning (Black & William, 1998; Contreras-Pérez & Zuñiga-González, 2017; Fraser, Walberg, Welch & Hattie, 1987; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Li & de Luca, 2014; Lysakowski & Walberg, 1982; Penny & Coe, 2004; Tenenbaum & Goldring, 1989). Alongside the various kinds of feedback analysed in these studies, the use of rubrics has recently begun to attract considerable attention. A number of definitions of rubrics can be found in the literature. For example, Jonsson and Svingby (2007) defined rubrics as an assessment tool that provides feedback about students' performance of tasks. Others, like Reddy and Andrade (2010), identified three key features as constituents of analytical rubrics: a) evaluation criteria (factors that an assessor considers when determining the quality of a student's work), b) quality definitions (detailed descriptions of what a student must do to demonstrate a skill, proficiency, or criterion in order to attain a particular level of achievement), and c) scoring strategy (scale for interpreting judgements of a product or process). Andrade (2000) and Panadero and Jonsson (2013) claimed that rubrics promote learning by helping students to make more accurate appraisals of their own work. Support for this idea comes from studies showing that students' learning is enhanced through approaches to instruction which help them to develop their self-assessment skills (Efklides, 2012; Hodgson & Pang, 2012; Panadero, Alonso-Tapia & Huertas, 2012; Sáiz, Montero, Bol & Carbonero, 2012). In a similar vein, it has been argued that rubrics can help students to monitor and assess their progress, not only while performing a task but also once it is completed (Panadero & Jonsson, 2013). Interestingly, the use of rubrics has also been associated with a reduction in negative emotions such as anxiety, leading in turn to improved academic achievement (Andrade & Du, 2005; Panadero *et al.*, 2012; Reynolds-Keefer, 2010). Reviews regarding the reliability, validity, and usefulness of rubrics and the effects of their use on student learning have been carried out by Brookhart and Chen (2015), Jonsson and Svingby (2007), Panadero and Jonsson (2013), and Reddy and Andrade (2010).

Among the various ways in which students may receive feedback, this study focuses on two: directly from the tutor in class and through rubrics. Although research supports the use of the former as a way of improving students' learning (Hattie, 2013; Mulliner & Tucker, 2017; Panadero & Romero, 2014), there is evidence that the use of rubrics is equally effective (Sáiz-Manzanares & Bol-Arreba, 2014). This suggests that high-quality learning may best be achieved through a combination of detailed in-class feedback from tutors and the frequent use of tasks that can be assessed by rubrics. In accordance with this view, Ene and Kosobucki (2016) proposed combining corrective feedback with rubric use in order to improve learning,

while Brookhart and Chen (2015) argued that future research should examine the use of rubrics with various instructional and formative assessment strategies.

In the present study a formative assessment system was developed and implemented in the context of a course entitled Research Designs, a second-year core subject in the University of Barcelona's Degree in Psychology. The Research Designs course has a considerable amount of statistical content and it requires numerical reasoning. It is therefore difficult and unpleasant for many students, who either fail the exams or drop out because they feel incapable of passing them. Núñez-Peña, Suárez-Pellicioni and Bono (2013) found that most students who failed to complete a Research Designs course showed a high level of math anxiety and negative attitudes towards this subject. In other studies, math anxiety has been associated with negative attitudes towards maths, poor maths performance, and low confidence in learning maths (Jasen *et al.*, 2013). The formative assessment system described here was implemented with the aim of minimising these negative effects and lowering the level of difficulty of the Research Designs course.

The nature and content of the formative assessment was determined by taking into account previous studies that have investigated the best strategies for achieving learning objectives in courses with mathematical and statistical content (for a review, see Garfield & Ben-Zvi, 2007). Based on this research we chose to include the following series of activities: seven problem-solving activities, five computer-based activities, and 22 independent study tasks (see the section Materials). We began by designing the various activities and the corresponding solution sheets that would support the in-class feedback. In parallel, and for each activity, we also wrote the corresponding rubrics. These rubrics were produced through consensus among all the course tutors and were validated prior to evaluating their usefulness. Because our aim was to determine which kind of feedback (i.e., rubrics, personalized in-class feedback, or a combination of the two) was most likely to enhance students' learning on the Research Designs course, we wrote a questionnaire to gather students' views about the usefulness of these different kinds of feedback. A further objective was to explore the influence that the two types of feedback had on students' academic achievement.

2. Methods

Participants

The validity and reliability of the rubrics were assessed using a sample of 100 volunteer undergraduates from the University of Barcelona, all of whom were enrolled during the 2014-2015 academic year in the Research Designs course as part of the University of Barcelona's Degree in Psychology (35 students to evaluate the rubric for the problem-solving activities, 36 for the rubric corresponding to the computer-based activities, and 29 for the independent study rubric). All these students participated anonymously.

We then assessed the usefulness of the two types of feedback using a sample of 135 undergraduates from the same university who were taking the Research Designs course during the 2015-2016 academic year. This sample consisted of 99 women (73.3%) and 36 men (26.7%) with a mean age of 21.9 years (SD 5.1, range 18-50). Of these participants, 106 stated that they had used the rubrics, 129 said they had attended feedback classes and 100 stated that they had both consulted the rubrics

and attended feedback classes. In accordance with the University of Barcelona's Code of Good Research Practice, all 135 students signed informed consent prior to their participation in the study.

Materials

The Research Designs course was structured around seven problem-solving activities, five computer-based activities (Núñez-Peña, 2012), and 22 independent study tasks (Núñez-Peña & Bono, 2012; Núñez-Peña, Bono & Suárez-Pellicioni, 2013), each accompanied by the corresponding rubric and, in the case of the problem-solving and computer-based activities, by a solutions sheet.

A 16-item questionnaire (Bono, Núñez-Peña & Suárez-Pellicioni, 2017) was used to explore the perceived usefulness of the two types of feedback with regard to the following aspects (Table 1): a) course preparation (questions 1 and 9), b) improved learning (questions 2 and 10), c) understanding the assessment criteria (questions 3 and 11), d) self-assessment (questions 4 and 12), e) reduction in test anxiety (questions 5, 6, 13 and 14), f) overall usefulness (questions 7 and 15) and g) value of using the two types of feedback in other courses in the degree (questions 8 and 16). All the questions were answered using a five-point Likert-type scale: (1) strongly disagree, (2) disagree, (3) neither agree nor disagree, (4) agree, and (5) strongly agree. In addition to completing the questionnaire, students had to indicate how often they had used the rubrics and/or attended the feedback classes, using a four-point scale: (1) never, (2) occasionally, (3) regularly, and (4) always.

Table 1. Questions regarding the perceived usefulness of rubrics and in-class feedback

| Questions about rubrics | Questions about in-class feedback |
|--|--|
| 1. The rubrics helped me to prepare for the course. | 9. The feedback classes helped me to prepare for the course. |
| 2. My learning improved as a result of the rubrics. | 10. My learning improved as a result of the feedback classes. |
| 3. The rubrics gave me a better understanding of the assessment criteria. | 11. The feedback classes gave me a better understanding of the assessment criteria. |
| 4. Through the use of the rubrics I learned to self-assess my performance. | 12. Through the feedback classes I learned to self-assess my performance. |
| 5. The rubrics helped me to feel less anxious when preparing for the exam involving open questions. | 13. The feedback classes helped me to feel less anxious when preparing for the exam involving open questions. |
| 6. The rubrics helped me to feel less anxious when preparing for the multiple-choice exam. | 14. The feedback classes helped me to feel less anxious when preparing for the multiple-choice exam. |
| 7. Overall, I think the rubrics were really useful. | 15. Overall, I think the feedback classes were really useful. |
| 8. I would like it if there were rubrics for the practical and project work of all the other courses on my degree. | 16. I would like it if there were feedback classes for the practical and project work of all the other courses on my degree. |

Source: Bono *et al.* (2017)

Finally, the Research Designs course was assessed by two exams. In one, students had to answer 12 open questions (both theoretical and practical) related to four case studies. The other was a multiple choice exam involving 30 questions (each with four response options) related to two independent study tasks (15 questions per task).

Procedure

The rubrics for all the activities (seven problem-solving activities, five computer-based activities, and 22 independent study tasks) were developed during the 2013-2014 academic year, with course tutors agreeing as to their content. During the 2014-2015 academic year the rubrics were validated by comparing students' self-assessment scores with those awarded by peers and tutors. Finally, during the 2015-2016 academic year the formative assessment system was implemented in the Research Designs course.

At the beginning of the 2015-2016 Research Designs course the activities that students had to carry out and their respective rubrics were uploaded to the course website (www.ub.edu/disin). Each type of research design featured in the course syllabus was worked on in three sessions spread across one week: the first was a lecture, the second was a practical session (problem-solving or computer-based activity) and in the third the tutor offered specific feedback on the previous practical session. As a complement to the information provided in feedback classes, solution sheets were prepared for all the problem-solving and computer-based activities. All these sheets were available for download from the course website once the in-class feedback had been given. Regarding the independent study tasks, each student was set two tasks to work on outside class. In the problem-solving sessions, students were set methodological questions related to published research in different areas of psychology. In the computer-based sessions, students explored SPSS, a software package that is widely used for data analysis in psychology, and they also learnt to interpret the results of statistical analyses in practical examples. Finally, in the independent study tasks, students answered methodological questions about scientific research articles and carried out the corresponding statistical analyses. For each of these tasks, students were given a data matrix, simulated on the basis of the results of the original study. In this way they learnt about the standard format and structure of research articles, and also how to obtain statistical information from a data matrix in a more or less real-life situation.

Before the end of the course, students completed the questionnaire about the perceived usefulness of rubrics and in-class feedback. Then, at the end of the course, they sat two exams: one involved open questions about both theoretical and practical aspects of the course (80% of the final grade), while the other was a four-option multiple-choice exam about two of the independent study tasks (20% of the final grade). In order to discourage guessing in the multiple-choice exam, students were told that 0.25 marks would be deducted for each incorrect answer. The open-question exam used practical examples similar to those that students had worked on in the problem-solving and computer-based activities and about which they had received error feedback in class. The multiple-choice exam for each student concerned the two independent study tasks they had previously worked on, and the statistical analyses that had been carried out when they completed these tasks could be consulted during the exam.

Data analysis

The rubrics were validated by calculating the level of association (correlations) and consistency (degree of agreement). Using the scores obtained when applying the rubrics for each kind of activity (problem-solving, computer-based work, and independent study) we calculated Kendall's tau-b

correlation coefficients between students' self-assessment ratings and the ratings of their peers and tutors. The consistency of the rubrics was examined by calculating the degree of agreement between students' self-assessments, assessment by peers, and tutor assessments.

Regarding responses to the questionnaire shown in Table 1, we calculated percentages for the perceived usefulness of rubrics and in-class feedback. For the responses of the students who had used rubrics and attended feedback classes, we applied the Wilcoxon signed-rank test for related samples in order to determine which kind of feedback was rated more highly. We also calculated Spearman correlations between (1) responses to the questionnaire items and overall achievement in the Research Designs course (i.e., mark in the open-question and multiple-choice exams), (2) frequency of attendance at feedback classes and achievement, and (3) frequency of use of rubrics and achievement. Note that the mark for the multiple-choice exam was broken down into final mark, number of hits, number of errors, and number of unanswered questions. Finally, we conducted a one-way variance analysis with Bonferroni correction to compare the marks in the final exam of students who only used rubrics, students who only attended feedback classes, and students who did both.

3. Results

Analysis of rubrics

Analysis of the scores obtained with the rubric for the problem-solving activities revealed statistically significant correlations between students' self-assessments and peer assessments ($\tau_b = .863$; $p < .01$), between self-assessments and tutor assessments ($\tau_b = .820$; $p < .01$), and between peer and tutor assessments ($\tau_b = .884$; $p < .01$). Correlations between the scores obtained with the rubric for the computer-based activities were likewise statistically significant: $\tau_b = .954$ ($p < .01$) between self-assessments and peer assessments, $\tau_b = .786$ ($p < .01$) between self-assessments and tutor assessments, and $\tau_b = .802$ ($p < .01$) between peer and tutor assessments. Finally, for scores obtained with the independent study rubric the correlations were as follows: $\tau_b = .852$ ($p < .01$) between self-assessments and peer assessments, $\tau_b = .808$ ($p < .01$) between self-assessments and tutor assessments, and $\tau_b = .863$ ($p < .01$) between peer and tutor assessments.

Percentages of agreement between students' self-assessments, assessment by peers, and tutor assessments were above 66%. For all the activities the highest degree of agreement was that between students' self-assessment and the assessment by peers: 91.4% for the problem-solving activities, 97.2% for the computer-based work, and 79.3% for the independent study tasks. A high degree of agreement was also observed between students' self-assessments and tutor assessments: 80% for the problem-solving activities, 72.2% for the computer-based work, and 69% for the independent study tasks. The lowest level of agreement was that between peer and tutor assessments: 80% for the problem-solving activities, 66.7% for the computer-based work, and 69% for the independent study tasks. In summary, agreement was greater between students and their peers than between students and tutor or between peers and tutor.

Analysis of questionnaire answers and exam grades

Table 2 shows that both types of feedback were perceived as useful by a high percentage of students, although this was especially the case for in-class feedback (questions 9-16). Lower percentages were only observed for the questions referring to a reduction in test anxiety (questions 5, 6, 13, and 14).

Table 2. Percentage of responses by response option for each question about the perceived usefulness of rubrics and in-class feedback

| Questions | Strongly disagree | Disagree | Neither agree nor disagree | Agree | Strongly agree |
|-----------|-------------------|----------|----------------------------|-------|----------------|
| 1 | 0.9% | 7.5% | 33.0% | 45.4% | 13.2% |
| 2 | 0.9% | 12.3% | 36.8% | 42.5% | 7.5% |
| 3 | 1.9% | 1.9% | 7.5% | 51.9% | 36.8% |
| 4 | 3.8% | 6.6% | 21.7% | 52.8% | 15.1% |
| 5 | 10.4% | 24.5% | 33.0% | 27.4% | 4.7% |
| 6 | 7.4% | 20.8% | 32.1% | 34.0% | 5.7% |
| 7 | 0.9% | 7.5% | 24.5% | 51.1% | 16.0% |
| 8 | 2.8% | 1.9% | 17.9% | 34.0% | 43.4% |
| 9 | 0.0% | 2.3% | 12.4% | 48.9% | 36.4% |
| 10 | 0.0% | 2.3% | 17.8% | 45.0% | 34.9% |
| 11 | 0.7% | 3.9% | 13.2% | 43.4% | 38.8% |
| 12 | 0.0% | 4.6% | 21.7% | 41.9% | 31.8% |
| 13 | 4.6% | 21.7% | 29.5% | 27.1% | 17.1% |
| 14 | 6.1% | 18.6% | 26.4% | 35.7% | 13.2% |
| 15 | 0.0% | 2.2% | 17.1% | 48.1% | 32.6% |
| 16 | 0.7% | 0.8% | 10.9% | 41.1% | 46.5% |

We applied the Wilcoxon signed-rank test for related samples to the responses of the 100 students who said they had used rubrics and attended feedback classes. This analysis showed that in-class feedback was viewed more positively than rubrics for all the analysed aspects (all p -values $\leq .01$), with the exception of questions 3 and 11, which refer to an understanding of the assessment criteria ($p = .833$).

Table 3 shows that there was a positive correlation between the final grades obtained by students and a reduction in perceived test anxiety (i.e., less perceived test anxiety, higher mark), whether as a result of using rubrics (questions 5 and 6) or through feedback classes (questions 13 and 14). This significant correlation applied to both the open-question and multiple-choice exams. The negative correlation between a reduction in perceived test anxiety and the number of questions left unanswered in the multiple-choice exam (i.e., less perceived test anxiety, fewer questions left unanswered) was also significant. Moreover, for the students who used rubrics a negative correlation was observed between a reduction in perceived test anxiety and the number of errors in the multiple-choice exam (i.e., less perceived test anxiety, fewer errors).

Responses to Question 7, which asked students for their overall view on the usefulness of rubrics, were positively correlated with performance in the open-question exam and negatively associated with the number of unanswered questions in the multiple-choice exam. Responses to questions 10 and 11, which asked students to consider whether the feedback classes had improved their learning

and their understanding of assessment criteria, were positively associated with grades in the open-question exam. Question 12 asked students whether the feedback classes had helped them to self-assess, and responses here were positively correlated with overall performance in both exams, and also with the number of correct answers in the multiple-choice exam. Responses to Question 12 were negatively associated with the number of errors students made in the multiple-choice exam. Finally, the perceived need to incorporate in-class feedback into other degree courses (Question 16) was significantly associated with students' performance in the open-question exam.

Table 3. Spearman correlations between performance in the two exams and responses to the questions about the perceived usefulness of rubrics and in-class feedback

| Questions | Open questions exam | Multiple choice exam | | | Mark |
|-----------|---------------------|----------------------|---------|----------------------|--------|
| | Mark | Hits | Errors | Unanswered questions | |
| 1 | .138 | .070 | -.039 | -.133 | .050 |
| 2 | .186 | .097 | -.055 | -.075 | .093 |
| 3 | .095 | -.064 | .047 | .036 | -.067 |
| 4 | .157 | .126 | -.050 | -.124 | .118 |
| 5 | .416** | .414** | -.290** | -.278** | .410** |
| 6 | .351** | .404** | -.259** | -.295** | .399** |
| 7 | .206* | .175 | -.033 | -.209* | .158 |
| 8 | .145 | .015 | -.003 | -.021 | .010 |
| 9 | .190 | .052 | -.024 | -.047 | .050 |
| 10 | .234* | .109 | -.135 | -.005 | .118 |
| 11 | .201* | .034 | -.065 | .021 | .038 |
| 12 | .408** | .296** | -.296** | -.101 | .309** |
| 13 | .334** | .329** | -.120 | -.326** | .309** |
| 14 | .289** | .331** | -.125 | -.322** | .311** |
| 15 | .138 | .095 | -.070 | -.054 | .093 |
| 16 | .235* | .086 | -.142 | .028 | .098 |

* $p < .05$ ** $p < .01$

Variance analyses performed in order to compare the academic achievement of students who only used rubrics, students who only attended feedback classes, and students who did both revealed no significant differences for either the open-question exam ($F(2,130) = 0.654, p = .522$) or the multiple-choice exam ($F(2,128) = 2.191, p = .116$). However, results from the Bonferroni correction showed that students who had received both kinds of feedback left fewer questions unanswered in the multiple-choice exam than students who only used rubrics ($p = .015$). A positive correlation was also observed between the frequency of attendance at feedback classes and both the final mark in the multiple-choice exam ($r = .178; p = .042$) and the number of hits in this exam ($r = .179; p = .041$). By contrast, the frequency with which students used any of the rubrics (whether for the problem-solving activities, the computer-based activities, or the independent study tasks) was not related to their achievement in either the open-question or the multiple-choice exam.

4. Discussion

Our goal was to implement a formative assessment system that involved two ways of giving students feedback: rubrics and personalized in-class feedback from the tutor. The analysis showed that the rubrics

were particularly useful in helping students to understand the assessment criteria, while feedback classes with the tutor allowed them to correct their mistakes and learn from them. Our expectation was that by using rubrics and attending feedback classes, students would become more aware of their errors and see that what they learnt in this process would enable them to perform better in subsequent exams.

In the feedback classes, which focused exclusively on giving students feedback about the activity they had carried out in the previous practical session (problem-solving or computer-based activity), our aim was to encourage students to reflect on their own learning process. When feedback is given almost immediately after a task, it can lead to improved learning (Brown *et al.*, 2016; Tang & Harrison, 2011). Like the rubrics, these feedback classes enabled students to have a clearer idea of what was expected of them, including the level of achievement required to obtain a good mark in the end-of-year exams.

The first step in implementing the formative assessment system involved analysing the validity and consistency of all the rubrics that would be used in the Research Designs course. Subsequently, we examined the usefulness of these rubrics, in comparison with in-class feedback about errors.

Together, the positive correlations and the high degree of agreement between students' self-assessment and the corresponding peer and tutor assessments support the validity and consistency of the rubrics written in this study. Although there are other, more precise indicators of reliability, we chose these measures because the use of rubrics for formative assessment leads to a constant dialogue between student and tutor that enables them to discuss and revise their respective views of the student's work (Marin-Garcia & Santandreu-Mascarell, 2015). Thus, the rubrics used here can be regarded as valid and reliable instruments for evaluating the acquisition of competences in the activities that are part of the Research Designs course.

The students considered both types of feedback to be useful for their learning, although they gave higher ratings to direct feedback from tutors. More specifically, the students who indicated in the questionnaire that the feedback classes had helped them to improve their learning and to have a better understanding of the assessment criteria achieved a higher mark in the open-question exam; students who considered that rubrics in general were useful also did better in this exam. Students who felt that the feedback classes had helped them to become better at self-assessment achieved higher marks in both exams (i.e., open questions and multiple choice), and they also answered more questions correctly and made fewer errors in the multiple-choice test. Exam performance was also better among the students who reported feeling less anxious about exams as a result of using rubrics or attending feedback classes. This finding is consistent with previous studies showing that the use of rubrics helps to reduce negative emotions such as anxiety, leading to better academic achievement (Andrade & Du, 2005; Panadero *et al.*, 2012; Reynolds-Keefer, 2010).

Panadero and Jonsson (2013) and Reddy and Andrade (2010) concluded that rubrics alone did not improve students' performance. However, other authors have claimed that rubrics are more useful when they contain examples of different achievement levels (Jonsson & Svingby, 2007). The results of the present study suggest that personalized in-class feedback can make students more confident about their grasp of statistical course content, helping them to feel better prepared and more capable of succeeding in exams. This is reflected in the results of a recent study by Núñez-Peña *et al.* (2015), who found a positive relationship between academic achievement in a course with mathematical content and the perceived usefulness of in-class feedback. These authors concluded that providing students with in-class feedback

about errors helped them to feel more confident about their ability to learn and reduced the negative impact of math anxiety on performance. In another recent study, Sesé, Jiménez, Montañó and Palmer (2015) found that the students who performed better in statistics courses were those who had a more positive attitude towards this subject, and that attitudes were negatively affected by anxiety. Thus, attitudes play a mediating role on the relationship between anxiety and performance. In relation to our study, this could explain why the students who experienced a change in attitude (i.e., those who felt less anxious about exams as a result of the rubrics or in-class feedback) subsequently did better in their exams.

To conclude, the results of this study show that feedback can help students to develop a more positive view of their ability to learn, especially when it is personalized and given in class. Both types of feedback employed here were able to reduce perceived test anxiety in 35.9% of the students following the use of rubrics and in 46.5% following in-class feedback. Importantly, the students who reported feeling less anxious as a result of either type of feedback did better in their exams. Finally, note that students' academic achievement did not differ significantly according to the type of feedback they had received (rubrics, in-class feedback, or both). However, performance in the multiple-choice exam was related to the frequency of attendance at feedback classes.

A limitation of this study is that we did not use a quasi-experimental design involving three groups of students: one that used rubrics, another that received in-class feedback, and the other with access to both types of feedback. Instead, the results were obtained by grouping students a posteriori according to the type of feedback they had made use of. Ideally, we would have used a quasi-experimental design with a control group and pre/post-test measures, incorporating appropriate covariables so as to isolate their effects (e.g., students' learning styles, previous academic achievement, number of hours of study, motivation, etc.). A further issue to consider is that the study groups comprised our own students. Despite these limitations, and given both the difficulty of using a quasi-experimental design and the small number of studies that use such a design with convenience samples in higher education (Brookhart & Chen, 2015), we consider that the present results will be of interest both to researchers working on teaching approaches in higher education and to university tutors themselves. In any case, the results could be bolstered by replicating the study in future courses.

Acknowledgements / Supporting agencies

This research was carried out with funding from the University of Barcelona, Spain (projects 2014PID-UB/069 and GINDOC-UB/099 of the Consolidated Group for Innovation in Teaching).

<References>

- Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57, 13–18.
- Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation*, 10, 1–11.
- Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80, 139–148.



- Bono, R., Núñez-Peña, M. I., & Suárez-Pellicioni, M. (2017). Rubrics use and in-class feedback in higher education: Students' perceptions and their effect on academic achievement. *Proceedings of the HEAd'17 - 3rd International Conference on Higher Education Advances*, pp. 338–346. Valencia: Editorial Universidad Politècnica de València. <https://doi.org/10.4995/HEAd17.2017.5198>
- Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, 67, 343–368. <https://doi.org/10.1080/00131911.2014.929565>
- Brown, G. T. L., Peterson, E. R., & Yao, E. S. (2016). Student conceptions of feedback: Impact on self-regulation, self-efficacy, and academic achievement. *British Journal of Educational Psychology*, 86, 606–629. <https://doi.org/10.1111/bjep.12126>
- Contreras-Pérez, G., & Zuñiga-González, C. G. (2017). Concepciones de profesores sobre retroalimentación: Una revisión de la literatura. *Magis, Revista Internacional de Investigación en Educación*, 9, 69–90. <https://doi.org/10.11144/Javeriana.m9-19.cpsr>
- Efklides, A. (2012). Commentary: How readily can findings from basic cognitive psychology research be applied in the classroom? *Learning and Instruction*, 22, 290–295. <https://doi.org/10.1016/j.learninstruc.2012.01.001>
- Ene, E., & Kosobucki, V. (2016). Rubrics and corrective feedback in ESL writing: A longitudinal case study of an L2 writer. *Assessing writing*, 30, 3–20. <https://doi.org/10.1016/j.asw.2016.06.003>
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83, 70–120. <https://doi.org/10.3102/0034654312474350>
- Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Identifying the salient facets of a model of student learning: A synthesis of meta-analyses. *International Journal of Educational Research*, 11, 187–212.
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75, 372–396. <https://doi.org/10.1111/j.1751-5823.2007.00029.x>
- Hattie, J. (2013). Calibration and confidence: Where to next? *Learning and Instruction*, 24, 62–66. <https://doi.org/10.1016/j.learninstruc.2012.05.009>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112. <https://doi.org/10.3102/003465430298487>
- Hodgson, P., & Pang, M. Y. C. (2012). Effective formative e-assessment of student learning: A study on a statistics course. *Assessment and Evaluation in Higher Education*, 37, 215–225. <https://doi.org/10.1080/02602938.2010.523818>



- Jasen, B. R. J., Louwese, J., Straatemeier, M., Van der Ven, S. H. G., Klinkenberg, S., & Van der Maas, H. L. J. (2013). The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. *Learning and Individual Differences, 24*, 190–197. <https://doi.org/10.1016/j.lindif.2012.12.014>
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*, 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284.
- Li, J., & de Luca, R. (2014). Review of Assessment Feedback. *Studies in Higher Education, 39*, 378–393. <https://doi.org/10.1080/03075079.2012.709494>
- Lysakowski, R. S., & Walberg, H. J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal, 19*, 559–578. <https://doi.org/10.3102/00028312019004559>
- Marin-Garcia, J. A., & Santandreu-Mascarell, C. (2015). ¿Qué sabemos sobre el uso de rúbricas en la evaluación de asignaturas universitarias? *Intangible Capital, 11*, 118–145.
- Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment and Evaluation in Higher Education, 42*, 266–288. <https://doi.org/10.1080/02602938.2015.1103365>
- Núñez-Peña, M. I. (2012). *Diseños de investigación en psicología: Problemas y prácticas de ordenador*. In Colección OMADO del Depósito Digital de la Universidad de Barcelona. Available at <http://hdl.handle.net/2445/21822>
- Núñez-Peña, M. I., & Bono, R. (2012). *Diseños de investigación en psicología: Prácticas para el trabajo autónomo del alumno – Actividades 1-10*. In Colección OMADO del Depósito Digital de la Universidad de Barcelona. Available at <http://hdl.handle.net/2445/22002>
- Núñez-Peña, M. I., Bono, R., & Suárez-Pellicioni, M. (2013). *Diseños de investigación en psicología: Trabajo autónomo del alumno – Actividades 11-22*. In Colección OMADO del Depósito Digital de la Universidad de Barcelona. Available at <http://hdl.handle.net/2445/45192>
- Núñez-Peña, M. I., Bono, R., & Suárez-Pellicioni, M. (2015). Feedback on students' performance: A possible way of reducing the negative effect of math anxiety in higher education. *International Journal of Educational Research, 70*, 80–87. <https://doi.org/10.1016/j.ijer.2015.02.005>
- Núñez-Peña, M. I., Suárez-Pellicioni, M., & Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research, 58*, 36–43. <https://doi.org/10.1016/j.ijer.2012.12.004>



- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*, 22, 806–813. <https://doi.org/10.1016/j.lindif.2012.04.007>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. <https://doi.org/10.1016/j.edurev.2013.01.002>
- Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy and Practice*, 21, 133–148. <https://doi.org/10.1080/0969594X.2013.877872>
- Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74, 215–253. <https://doi.org/10.3102/00346543074002215>
- Pereira, D., Flores, M. A., Veiga-Simão, A. M., & Barros, A. (2016). Effectiveness and relevance of feedback in higher education: A study of undergraduate students. *Studies in Educational Evaluation*, 49, 7–14. <https://doi.org/10.1016/j.stueduc.2016.03.004>
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment and Evaluation in Higher Education*, 35, 435–448. <https://doi.org/10.1080/02602930902862859>
- Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research and Evaluation*, 15, 1–9.
- Sáiz-Manzanares, M. C., & Bol-Arreba, A. (2014). Aprendizaje basado en la evaluación mediante rúbricas en educación superior. *Suma Psicológica*, 21, 28–35. [https://doi.org/10.1016/S0121-4381\(14\)70004-9](https://doi.org/10.1016/S0121-4381(14)70004-9)
- Sáiz, M. C., Montero, E., Bol, A., & Carbonero, M. A. (2012). An analysis of learning competences at the university. *Electronic Journal of Research in Educational Psychology*, 10, 253–270.
- Sesé, A., Jiménez, R., Montaña, J. J., & Palmer, A. (2015). Can attitudes toward statistics and statistics anxiety explain students' performance? *Revista de Psicodidáctica*, 20, 285–304. <https://doi.org/10.1387/RevPsicodidact.13080>
- Tang, J., & Harrison, C. (2011). Investigating university tutor perceptions of assessment feedback: Three types of tutor beliefs. *Assessment and Evaluation in Higher Education*, 36, 583–604. <https://doi.org/10.1080/02602931003632340>
- Tenenbaum, G., & Goldring, E. (1989). A meta-analysis of the effect of enhanced instruction: Cues, participation, reinforcement and feedback and correctives on motor skill learning. *Journal of Research and Development in Education*, 22, 53–64.