

Formative assessment and academic achievement in pre-graduate students of health sciences

María T. Carrillo-de-la-Peña · Eva Baillès · Xavier Caseras · Àlvar Martínez · Generós Ortet · Jorge Pérez

Received: 16 October 2006 / Accepted: 8 October 2007 / Published online: 31 October 2007
© Springer Science+Business Media B.V. 2007

Abstract Although educational experts recommend the use of formative assessment, there is a dearth of empirical studies on its impact on academic achievement. In this research the authors analyse to what extent participation and performance in formative assessment are associated with positive academic outcomes of pre-graduate students of health sciences. A total of 548 students from three health science degrees (Medicine, Psychology and Biology) from four Spanish universities were involved in this study. The students who carried out mid-term formative assessment got better marks and had higher success rates in final summative assessment than the students who did not participate. In addition, success in formative assessment tests was associated with better summative marks. Interestingly, participation in formative assessment was a better predictor of final outcome than success in formative assessment, a result that supports the key role of feedback in formative assessment. Students who took the mid-term examination, irrespective of their success, obtained feedback about their achievement and probably this determined their greater involvement in the learning process. Although causal relationships between formative and summative assessment cannot be established from this research, the generalized benefits of formative assessments found here encourage the practice of them in health sciences education.

Keywords Formative assessment · Higher education · Academic achievement

M. T. Carrillo-de-la-Peña (✉)
Faculty of Psychology, Universidade de Santiago de Compostela, Campus universitario s.n., 15702
Santiago de Compostela, Spain
e-mail: pcpbmtc@usc.es

E. Baillès · J. Pérez
Faculty of Health and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain

X. Caseras · À. Martínez
Faculty of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

G. Ortet
Faculty of Human and Social Sciences, Universitat Jaume I, Castellon, Spain

Overview

- There is a dearth of empirical data on the benefits of formative assessment on academic achievement.
- This study shows that participation in formative assessment was a better predictor of final outcome than success in it, and thus supports previous conceptualisations which consider feedback the key element in formative assessment.
- From our results we cannot extract causal relationships, but formative assessment was associated positively to summative achievement. Future studies should be designed so as to control for possible initial academic differences between participants and non participants in the formative assessments.

Introduction

Assessment of students' academic achievement is a basic step in any educational project since it provides information about the success in the attainment of our specific teaching objectives (Guilbert 1992; Wass et al. 2001). It is well known that students cope with learning processes according to the particular type of assessment used (Rolfe and McPherson 1995; Cohen-Schotanus 1999; Wass et al. 2001), and consequently assessment must closely match the learning objectives. The choice of the most suitable type of assessment is a crucial question.

Educational experts recommend earnestly the use of formative assessment in addition to summative assessment (Rolfe and McPherson 1995; Friedman 1999; Nendaz and Tekian 1999; Wass et al. 2001; Liaison Committee on Medical Education 2002; Rushton 2005). The purpose of formative assessment is not accreditation; it provides direct feedback about the learning and teaching processes and may have beneficial effects for both students and teachers (Rolfe and McPherson 1995; Wass et al. 2001). There are some Medical schools where formative assessments are institutionalised and carried out systematically (Neufeld et al. 1989; Rolfe and McPherson 1995; Nendaz and Tekian 1999; Nennin and Kalishman 1999), but unfortunately this kind of assessment is not very frequent in Mediterranean countries, especially in Spain. This may be due to the lack of empirical evidence on their advantages. It is not clear whether student outcomes are better when formal formative assessment is conducted (Nendaz and Tekian 1999) and thus it has been proposed that this issue needs to be explored more extensively (Rolfe and McPherson 1995).

Formative assessment has been specially recommended in our country (Editorial 1998) and recently it has been implemented in some educational processes in different Spanish universities. Nevertheless, no studies on the impact of formative assessment have been undertaken so far, probably because of the difficulty for experimental research in this field. It would be desirable to compare the academic outcome of students randomly assigned to an experimental group (formative assessment), or to a control group (only final summative examination), but this procedure is not possible due to evident ethical reasons. To overcome this problem, in this study we let students choose voluntarily their participation in the formative mid-term examination, and compared their achievement in final written exams with students who did not participate.

Some time ago, we presented a pilot study reporting different experiences about formative assessment in health sciences studies that showed a positive relationship between formative and summative assessments (Caseras et al. 1999). The present paper represents

an extension of the previous report since it involves three different health science studies (Medicine, Psychology and Biology) in four Spanish universities.

Methods

Participants

A total of 548 students belonging to the following courses and universities were included in the study:

- (a) One hundred and sixty-eight second-year Medicine students from a course of Medical Psychology at the Faculty of Medicine in the Autonomous University of Barcelona (MP-UAB).
- (b) One hundred and forty-one first-year Psychology students from a course of Personality Psychology at the Faculty of Human and Social Sciences in the Jaime I University of Castellon (PP-UJI).
- (c) Sixty-five fourth-year Psychology students from a course of Psychophysiology at the Faculty of Psychology in the University of Santiago de Compostela (PPh-USC).
- (d) One hundred and seventy-four students of Biology at Pompeu Fabra University, enrolled in Zoology, a first year course (ZOO-UPF), Genetics (GEN-UPF), in second year, or Ecology (ECO-UPF), in third year.

All of them are mandatory courses in their respective degrees. The first three cases represent isolated teaching experiences developed by the lecturers responsible for the courses, all of them authors of the present study. However, the last three cases are part of a global educational project in which formative assessments are institutionalised and carried out in all the subjects of the Faculty of Health and Life Sciences at Pompeu Fabra University.

Data collection

Although the contexts were different (various universities, different university degrees, different academic characteristics of students...), the experiences about formative assessment were very similar. Formative assessments were carried out at the middle of the academic term, using mostly short answer question tests and also Multiple Choice Questions (MCQ). After the examinations, students had immediate feedback on their performance. Once out of the classroom, they found the correct responses to all the items in the notice board, and thus had an immediate knowledge of their results. Lately, the day after the exam was corrected, the main results and most frequent errors were commented on in the classroom. All the students, regardless of whether or not they had participated in the formative assessment, usually took part in these feedback sessions.

In all cases, participation in the formative assessment was completely voluntary. Not participating in or failing the tests had no negative impact on the final summative mark. On the other hand, and in order to improve motivation, pupils who passed the mid-term tests (with marks equal or higher than 5 in a 0–10 scale) obtained a small increase in the final summative mark (proportional to the achievement, to a maximum of a 5% of the final mark).

In each case, summative assessment of knowledge at the end of the academic term was carried out using multiple choice question exams. In two subjects, ZOO-UPF and

ECO-UPF, short answer question exams were also used. We would like to point out that only the first summative examination period was used in the collection of the data. In Spain, students who do not pass their first exam may re-sit. This study presents the results of the students who completed the first summative examination.

In spite of the differences among the subjects, all the formative and summative tests were constructed using specification tables to guarantee their content validity, and had good reliability indices. The teaching staff involved is expert in assessment. Although the tests tapped mainly theoretical contents, they were not limited to factual knowledge but also included items on comprehension, reasoning and critical analysis of the main contents. The MCQ used in the summative evaluations presented adequate internal consistency, with Crombach's alpha indices superior to 0.80.

Data analyses

Firstly, we compared the mean summative examination results obtained by the students who took part in formative assessment with the mean results of the students who did not. To avoid biased results, we subtracted the small amount obtained by the students who were successful in the formative assessment test from the final summative mark.

Secondly, we compared the percentage of students that passed the established standard in the group that participated and in the group that did not participate in formative assessment. The pass mark (equal or higher than 5 in a 0–10 scale) was the same for all cases, since we followed the standard used in the Spanish university system.

Finally, and only within the group that participated in formative assessment, we compared the final summative marks between the students who were and who were not successful in the formative test.

The comparison of means was carried out using Student's *t*-tests. For the comparison of percentages we used χ^2 tests.

Results

The percentages of participation in formative assessment among the students enrolled in the courses were the following: MP-UAB, 91.1%; PP-UJI 39.7%; PPh-USC 50.0%; ZOO-UPF 92.1%; GEN-UPF 60.3% and ECO-UPF 66.1%.

Table 1 shows the mean of the final summative examination results of the students who took part together with the ones who did not participate in formative assessment. *T*-tests were significant for all the comparisons between the groups, i.e., for all the academic experiences, students who completed formative tests obtained higher marks in the final examination.

In Table 2, we present the percentages of success rates (passing the established standards) for the two groups considered (participants and non participants in formative assessment). χ^2 tests showed that there was a significantly higher percentage of students who passed the standard (5 or higher out of 10 points) within the group who participated in formative assessment.

The percentages of success rates in the final summative exams were also analysed with a view to compare the subjects successfully passed the formative examination with those that failed the mid-term test (see Table 3). χ^2 tests revealed that students with higher examination results in formative tests also obtained higher results in summative exams, although in some cases the differences were non-significant.

Table 1 Means and standard deviations (SD) of the summative results (0–10) obtained by the participants and non-participants in formative assessment (FA)

Subject and university	FA participants		FA non-participants		
	<i>N</i>	Mean (SD)	<i>N</i>	Mean (SD)	<i>t</i> -Test
MP-UAB	153	6.41 (1.1)	15	5.41 (1.1)	2.99**
PP-UJI	56	5.33 (1.5)	85	4.21 (1.6)	4.04***
PPh-USC	40	6.61 (2.0)	25	5.41 (1.8)	2.41*
ZOO-UPF ^a					
GEN-UPF	38	6.63 (1.0)	19	5.60 (1.1)	3.49***
ECO-UPF	39	7.74 (0.9)	20	6.86 (0.9)	3.50***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

MP-UAB, Medical Psychology-Autonomous University of Barcelona; PP-UJI, Personality Psychology-Jaume I University of Castellón; PPh-USC, Psychophysiology-University of Santiago de Compostela; ZOO-UPF, Zoology-Pompeu Fabra University; GEN-UPF, Genetics-Pompeu Fabra University; ECO-UPF, Ecology-Pompeu Fabra University

^a No comparisons were made because practically all of the students completed the formative test in this subject

Table 2 Number of participants and non-participants in formative assessment (FA) tests, and number and percentage of them who passed summative assessment (SA) exams

Subject and university	FA participants/passed SA (%)	FA non-participants/passed SA (%)	χ^2
MP-UAB ^a	153/138 (90.2)	15/9 (60.0)	13.39**
PP-UJI	56/40 (71.4)	85/36 (42.4)	11.48***
PPh-USC	40/32 (80.0)	25/14 (56.0)	4.28*
ZOO-UPF ^b			
GEN-UPF	38/36 (94.7)	19/14 (73.7)	5.21*
ECO-UPF ^c			

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

MP-UAB, Medical Psychology-Autonomous University of Barcelona; PP-UJI, Personality Psychology-Jaume I University of Castellón; PPh-USC, Psychophysiology-University of Santiago de Compostela; ZOO-UPF, Zoology-Pompeu Fabra University; GEN-UPF, Genetics-Pompeu Fabra University; ECO-UPF, Ecology-Pompeu Fabra University

^a In this subject, two formative tests were carried out; only those students who completed both tests and those who did not complete any of them were included

^b No comparisons were made because practically all of the students completed the formative test in this subject

^c No comparisons were made because all the students passed summative assessment in this subject

Discussion

In relation to the main aim of this research, the results of all the experiences were very similar: the students who participated in formative assessment obtained better examination results in summative assessment than the ones who did not. This was observed both quantitatively (better mean marks, Table 1) and qualitatively (higher percentage of

Table 3 Number of successful and unsuccessful students in formative assessment (FA), and the number and percentage who passed summative assessment (SA)

Subject and university	Success FA/passed SA (%)	Failure FA/passed SA (%)	χ^2
MP-UAB	138/127 (92.0)	15/11 (73.3)	5.35*
PP-UJI	32/27 (84.4)	24/13 (54.2)	6.13*
PPh-USC	21/18 (85.7)	19/14 (73.7)	0.90
ZOO-UPF	12/12 (100)	46/36 (71.7)	4.37*
GEN-UPF	18/18 (100)	20/18 (90.0)	1.79
ECO-UPF ^a			

* $p < 0.05$

MP-UAB, Medical Psychology-Autonomous University of Barcelona; PP-UJI, Personality Psychology-Jaume I University of Castellon; PPh-USC, Psychophysiology-University of Santiago de Compostela; ZOO-UPF, Zoology-Pompeu Fabra University; GEN-UPF, Genetics-Pompeu Fabra University; ECO-UPF, Ecology-Pompeu Fabra University

^a No comparisons were made because all the students passed summative assessment in this subject

students who passed the summative tests, Table 2). Success in formative assessment also predicted success in summative assessment (Table 3), but interestingly, participation was a better predictor of final outcome than success. The results showed that most of the students who got good academic marks in formative assessment also obtained good results in the summative exams. However, failing in formative assessment was not a predictor of examination failure, since the majority of the students who participated in formative assessment, but failed the formative test, were successful in the summative exams. In summary, the act of taking out formative assessment, with or without success, allows the students to receive positive feedback about their learning processes, which in turn may favour the final academic performance. This result is in line with relevant conceptualisations of formative assessment which identify feedback as its core element (Sadler 1989; Harlen and James 1997; Black and William 1998; Rushton 2005).

It is noteworthy that, although the academic experiences were developed in very different educational situations (type of university, university degrees, and academic characteristics of students) the results were practically identical in all the cases. Thus, this study provides consistent evidence on the beneficial effects of mid-term assessments. Remember that students had immediate information about their results in the formative exams. This procedure promotes learning since it allows, on the one hand, to detect which learning objectives were well or badly acquired, and on the other, to discuss the results with other students. Also, the informative session in the classroom after the formative exam would provide additional feedback for students.

It may be argued that these assessments were not strictly formative, since they could represent a small increase in the final summative score. Nevertheless, the impact of this small increase was corrected in the analyses. We maintain this small reward (note that very few students summed the maximum 0.5 points) because of its motivational value; without it, some of the students would probably not have participated in the formative assessments.

Interest has been shown in whether student participation and performance in formative exams predict later summative assessment results (Rolfe and McPherson 1995). From our results we cannot extract causal relationships nor confirm that formative assessment was the cause of higher summative achievement. Probably, the more brilliant and motivated

students are the more they participate and obtain better results in summative assessment. Assuming this important drawback, in some of the cases (i.e. medical students at UAB or biological students at UPF) we may assume that participants and non participants in formative assessments would present similar academic characteristics, since there are very few differences in motivation and previous academic history among the students enrolled in those studies. All of them were very motivated and had very high previous marks.

In order to establish causal relationships it is necessary to design educational experiments with random groups, but this is difficult to put into practice for obvious ethical reasons. Alternatively, future studies should use previous academic achievement as a covariate in order to control for possible initial differences between participants and non participants in the formative assessments.

In spite of its limitations, the present study suggests that the mere participation in formative assessment, irrespective of success, may determine greater involvement of students in the learning process and therefore greater success. Therefore, formative assessment seems to have an important motivational role and thus nothing goes against considering it as beneficial and worthy in any educational project.

Finally, and considering the necessity of a best evidence education (Hart 1999), we believe that our results may promote the use of formative assessment in health science studies, especially in Spain. Formative assessment could be used in specific teaching processes developed by university lecturers or, even better, developed within joint educational projects co-ordinated by the faculty government team. In this report, we used only mid-term formative assessments but in the future it would be desirable to perform them several times within the term. Also, it would be interesting to explore if the beneficial effects of formative assessments extend to practical skills and not only the acquisition of theoretical knowledge.

References

- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Caseras, X., Martínez, A., Ortet, G., & Pérez, J. (1999). Evaluaciones formativas en estudios de ciencias de la salud [Formative assessment in health sciences studies]. *Educación Médica*, 2, 58–9.
- Cohen-Schonatus, J. (1999). Student assessment and examination rules. *Medical Teacher*, 21, 318–321.
- Editorial. (1998). La evaluación formativa [Formative assessment]. *Educación Médica*, 1, 89.
- Friedman, M. (1999). Assessment in outcome-based education. *Medical Teacher*, 21, 23–25.
- Guilbert, J. J. (Ed.). (1992). *Education handbook for health personnel* (6th ed.). Geneva: World Health Organization.
- Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education*, 4, 365–369.
- Hart, I. (1999). Best evidence medical education (BEME). *Medical Teacher*, 21, 453–454.
- Liaison Committee on Medical Education. (2002). *Guide to the institutional self-study*. Washington DC: Association of American Medical Colleges.
- Nendaz, M. R., & Tekian, A. (1999). Assessment in problem-based learning medical schools: A literature review. *Teaching and Learning in Medicine*, 11, 232–243.
- Nennin, S. P., & Kalishman, S. (1999). Student assessment. *Academic Medicine*, 73, s46–s54.
- Neufeld, V. R., Woodward, C. A., & McLeod, S. M. (1989). The McMaster MD program: A case study of renewal in medical education. *Academic Medicine*, 64, 423–432.
- Rolfe, I., & McPherson, J. (1995). Formative assessment: How am I doing? *Lancet*, 345, 837–839.
- Rushton, A. (2005). Formative assessment: A key to deep learning? *Medical Teacher*, 27, 509–513.
- Sadler, D. R. (1989). Formative assessment in the design of instructional systems. *Instructional Science*, 18, 119–144.
- Wass, V., Van der Vloten, C., Shatzer, J., & Jones, R. (2001). Assessment of clinical competence. *Lancet*, 357, 945–949.