

# Enhancing students' performance: Alternative assessment in an ESP course for journalism students

TERESA RIESTRA CARRIÓN

Escuela de Lenguas Modernas

Universidad de Costa Rica

GLORIELA CHACÓN CHAVARRÍA

Sede de Occidente

Universidad de Costa Rica

## Abstract

The present study is based on an English for Specific Purposes (ESP) course for journalism students at University Costa Rica (UCR). To design the course assessment as recommended for ESP courses (Dudley Evans & St. John, 1998), the instructors based their decisions on the needs analysis and the assessment type used in UCR journalism major, which consists of projects. Thus, this course assessed students' performance by means of both traditional and alternative assessment, based on students' assessment preferences identified in the needs analysis. The aim of this research study was to find out which type of assessment--traditional or alternative--enhanced students' performance the most. As presumed by the researchers, the results clearly showed that alternative assessment was the type of evaluation that enhanced students' performance the most, as it is authentic and mirrors students' real professional contexts.

**Key words:** alternative assessment, traditional assessment, enhancing performance, ESP assessment

## Resumen

Esta ponencia está basada en un curso de Inglés para Fines Específicos (IFE) para estudiantes de periodismo de la Universidad de Costa Rica (UCR). Para diseñar la evaluación del curso como es sugerido para los cursos de IFE (Dudley Evans & St. John, 1998), las docentes basaron sus decisiones en el análisis de necesidades de los estudiantes, así como en el tipo de evaluación empleado en los cursos de la carrera de periodismo de la UCR, la cual consiste en la elaboración de proyectos. Se evaluó

el desempeño de los estudiantes por medio de pruebas tradicionales y evaluación alternativa, debido a que esta fue la forma en la que los estudiantes escogieron ser evaluados en el curso de acuerdo con el análisis de necesidades. El objetivo de este estudio fue determinar qué tipo de evaluación, tradicional o alternativa, enriqueció más el desempeño de los estudiantes. Como las investigadoras lo habían supuesto, los resultados del estudio mostraron claramente que la evaluación alternativa fue el modo de evaluar que enriqueció más el desempeño de los estudiantes debido al principio de autenticidad, pues refleja el contexto profesional.

**Palabras claves:** evaluación alternativa, pruebas tradicionales, enriquecimiento del desempeño, evaluación en IFE

In alternative assessment, “students are tested as they actually perform the behavior itself” (Brown, 2004, p.405). This type of assessment possesses a high degree of simulation of real-life contexts. Students feel empowered for they feel they can really show what they know. Since tasks are “consistent with course goals and curriculum, students and teachers are likely to be more motivated to perform them, as opposed to a set of multiple-choice questions (...)” (Brown, 2004, p.255). The present study is based on the News Broadcasters’ and Journalists’ (NBJ) English course taught as part of the Practicum compliance from the Masters’ in Teaching English as a Foreign Language at University of Costa Rica (UCR). The course was an English for Specific Purposes (ESP) for journalism students designed by student-teachers in the Master’s degree (MA). After conducting an analysis to determine the specific needs of journalism students in June 2010, the NBJ English course was team-taught by Gloriela Chacón and Teresa Riestra during the second semester in 2010. The research study will show how different types of alternative assessment could enhance students’ performance in an ESP course.

### Research questions

This research study aims to answer the following questions:

- Which type of assessment enhanced the students’ performance better: alternative or traditional testing?
- Which variables affected these results?

The objective addressing this area is as follows:

- To determine which type of assessment -traditional or alternative- enhanced the students’ performance more.

The following section reviews the most pertinent arguments on the topic.

## Literature review

Academic courses usually associate their immediate effectiveness to a crucial point for both teachers and learners: assessment. Students are concerned about negative indicators of their production, while teachers worry about course validity drawn from the assessment data collected from their students. Overall, many believe that evaluation and assessment indicate the same concept. Although divergent opinions about the term may be found, assessment is usually defined as the “process of gathering information about students” (Hart, 1994, p. 1). Teachers observe students’ production, skills and knowledge to certify that the course content is being learned. The concept of assessment is thus seen as an umbrella term in which evaluation is only one way of assessing students (Brown, 2004; Hart, 1994). Evaluation would then be described as a way of processing this information and translating it into specific judgment criteria (Hart, 1994, p. 1). Probably the most common method of evaluation is by means of written tests.

Students tend to believe that test scores reveal effective or ineffective teaching and learning of course content. In fact, Hutchinson and Waters (2000) claimed that “evaluation of learners reflects not just the learners’ performance but, to some extent, the effectiveness or otherwise of the course [itself]” (p.145). ESP courses are intended to be successful for they address learners’ specified objectives based on a previous needs analysis of the target students’ lacks and expectations. Prior to this discussion, the review will examine an underlying aspect which is a main focus in the present study regarding assessment: student performance.

## Performance

Chomsky (as cited in Shohamy, 1996) described performance as the actual language use in concrete situations. Students are a clear sign of the results of any course, just by observing their performance in the classroom and through assessment. Brown (2001) explained the difference between competence and performance, by stating that, “when you propose to assess someone’s ability in one or a combination of the four skills, you assess the person’s *competence*, but you observe the person’s *performance*” (p. 117). Owing to external or internal factors that can greatly influence students—a bad night sleep, test anxiety, an illness, or limited memory—it is important to recognize the “fallibility” of the results of a single performance. Hence, the observation of many performances will make the teacher capable of drawing a more accurate conclusion about students’ competence. As performance samples, Brown (2001) mentioned “several tests that are combined to form an assessment; in-class and extra-class graded work [are all] alternative forms of assessment (journal, observation, self-assessment, peer-assessment)” (p. 117). As a result, we “must rely on observable performances in our assessment of students as much as possible” (Brown, 2001, p. 117). Not all the skills are observable, though. Types of observable performance comprise speaking and writing products, where teachers perceive “the process [of internalizing

meaning] as it is performed” (Brown, 2001, p. 118). The next section of the literature review examines recommended types of assessment in ESP courses.

### **Assessment in ESP**

There is no one sole way to assess in English for Specific Purposes language courses. ESP is an approach; as a result, there is no “particular kind of language, teaching material or methodology” which defines these language courses (Hutchinson & Waters as cited in Dudley-Evans & St. John, 1998, p.2). Furthermore Dudley-Evans & St. John (1998) stated that ESP courses are characterized by “mak[ing] use of the underlying methodology and activities of the discipline it serves” (p. 4). Dudley-Evans & St. John (1998) described this particular way of assessing in ESP –labeled *continuous assessment*– as tasks which are accomplished over a period of time; students have the advantage of consulting a variety of resources to complete the task, such as books, other people and, currently, the Internet. Allegedly, this type of evaluation –commonly known as alternative assessment-- is more “flexible and formative” (p. 211). Classroom tests, on the other hand, have a negative connotation, at times seen by learners with “obsessive and fearful attitudes engendered by viewing tests exclusively as determiners of grades” (Hutchinson & Waters, 2000, p. 152). These written in-class tests are called traditional tests. The differences between traditional and alternative testing are discussed in the following section in order to highlight their advantages and disadvantages.

### **Traditional testing**

Traditional tests identify strengths and weaknesses in students’ abilities by evaluating content in a given time limit in a classroom setting. Some features which characterize this type of testing are the following: (a) students work individually; (b) tests are indirect or de-contextualized; (c) they are norm-referenced; and (d) traditional tests foster extrinsic motivation (Brown, 2004). Clapham and Wall (as cited in Coombe, Folse & Hubley, 2007) sorted traditional testing into the following categories: placement, progress or progressive, achievement, proficiency and diagnostic tests. Every one of these modes has a different objective. Hutchinson and Waters (2000) believed that progressive tests are “the most likely to be constructed in ESP” (p. 147). These tests are internal and show the nature and content of the course itself; this type of test assesses particular language skills and measures mastery of class work (Dudley-Evans & St. John, 1998; Coombe et al., 2007). The purpose of progressive tests is measuring “what the learner has learned with respect to different points in the course” (Graves, 2000, p. 212). One of the principles of progress test, thus, is the notion of “assess[ing] only what has been taught” (Graves, 2000, p. 212). Additionally, students are assessed on the basis of certain expected criteria during a written test—accurate grammar structures, reading or listening skills, for instance. These criteria are in turn the guidelines which teachers place in a continuum of met goals. As

students' errors are marked, feedback on the "overall progress in a language course" is then given in the form of a score (Coombe et al., 2007, p. 44). Logically, a word of caution is given by Dudley-Evans and St. John (1998) for ESP teachers to be aware of their responsibility "when [they] set, administer and mark [written] tests (...) [to] enable students to perform as well as possible" (p. 221).

Traditional testing –also called standardized– overtook the educational system as a result of the need for objectivity (Hart, 1994, p. 2), replacing individualized instruction and assessment. Hart reported that thousands of standardized tests have proven their validity and reliability since 1928 and have produced massive testing of young learners. However, she asserted that "in the great push for efficiency and accountability, American education became all but addicted to testing" (p.4). Critics who show strong disapproval of standardized tests point out three main faults: (a) flawed tests; (b) poor measurement of students' real abilities; and (c) a corruption of the teaching and learning process. The latter denotes trivialized content and skill development as well as overuse of rote learning as opposed to understanding and reflection (Hart, 1994). Regarding the infallibility of traditional testing, Bailey (as cited in Brown, 2004) stated that "one of the disturbing things about tests is the extent to which many people accept the results uncritically, while others believe that all testing is invidious. But tests are simply measurement tools: It is the use to which we put their results that can be appropriate or inappropriate" (p. 251). Thus, it is not surprising that educators have been investigating other types of testing to replace over formal classroom tests.

### **Alternative assessment**

Alternative assessment differs from traditional assessment mostly in the sense that it "asks students to show what they can do" (Hart, 1994; Coombe et al., 2007). In other words, this type of assessment provides authenticity because "it reflects the curriculum that is actually being implemented in the classroom" (Huerta-Macias as cited in Coombe et al., 2007, p.44). If ESP courses, as indicated above, use tasks and activities which reflect the learners' "specialist world" (Dudley-Evans & St. John, 1998), then these learners' evaluation must coincide with these types of specialized tasks. Moreover, alternative assessment is said to be "more multi culturally sensitive, free of norm, linguistic, and cultural biases found in traditional testing" (Huerta-Macias as cited in Coombe et al., 2007, p.44). This is probably due to untimed constraints and continuous, longitudinal assessment. In alternative assessment, grading highlights students' strengths based on clearly stated performance standards rather than counted errors (Hart, 1994). Alternative assessment can be considered as performance assessment since it emphasizes processes and competencies, while it de-emphasizes theoretical scoring outcomes. There are many ways to face challenges; hence, students personalize tasks when they discover their own way to accomplish them. Short assessment tasks done in the classroom will mirror event tasks, performed outside of the classroom setting (Hart, 1994). Event tasks are designed "to reveal

not only what students know but how well they can put their knowledge into use” (Hart, 1994, p.50). Average event tasks include projects, in which students work around a topic, using authentic information sources and evaluating data to attain a final product (Coombe et al., 2007). Projects can be carried out in groups or individually, in either case profiting from “a higher degree of involvement and ownership” (Dudley-Evans & St. John, 1998, p. 195). ESP practices encourage the integration of language skills in real-life context through “mini-research projects (...) formulated around interviews and questionnaires (...) engaging individuals in extensive purposeful reading followed by a written report” (p. 196). These writing assignments are tasks with specified real-life purposes. As Cushing Weigle (2002) perceptively stated, “any writing test that involves actual writing, as opposed to completing multiple-choice items, can be considered a performance test, since the written product represents a performance of writing” (p.46).

These alternative assessment tasks seem to benefit students more because they are assigned with a new role: students are active participants who take advantage of the variety and flexibility of an assessment activity to unveil their own learning styles and strategies, simultaneously increasing their self-esteem in real context situations. Huerta-Macías (as cited in Coombe et al., 2007) claimed that in this type of assessment, “students are evaluated on what they integrate and produce rather than on what they are able to recall and reproduce” (p. 44), meaning that traditional testing tends to centralize outcomes based on memorization. However, the final decision for effective testing will not depend on the type of testing in itself but on the correlation of the test and the guiding principles of good testing.

### **Principles of Testing: Authenticity and Validity**

Following Bachmand and Palmer (as cited in Dudley-Evans & St. John, 1998) as well as Coombe, Folse and Hubley (2007), it can be inferred that the chief characteristics for testing are authenticity and validity. Authenticity, as aforementioned, is achieved when tasks are “worthwhile, significant and meaningful” (Hart, 2000, p.9). These tasks reveal real-world situations and contexts, enabling students to face similar encounters in the target language beyond the classroom setting. Even so, Skehan (as cited in Douglas, 2000) argued that “merely making an interaction ‘authentic’ does not guarantee that the sampling of language involved will be sufficient, or [will be] the basis for wide ranging and powerful predictions of language behavior in other situations” (p.12). It must be recognized, thus, that providing authentic assessment will not automatically guarantee language competency in their future professional context. The question that emerges, then, is whether there is any certainty of actually testing representative target situation tasks. Douglas (2000) noted that authenticity in language testing is linked to situational and interactional functions:

Situational authenticity can be demonstrated by making the relationship between the test task characteristics and the features of tasks in the target

language use situation explicit. The second aspect of authenticity, interactional, (...) involves the interaction of the test taker's specific purpose language ability with the test task. The extent to which the test taker is engaged in the task, by responding to the features of target language use situation embodied in the test task characteristics, is a measure of interactional authenticity. (pp. 17-18)

Given these facts, authentic tasks seem to relate more to alternative assessments, not so to traditional testing.

Another essential factor to measure testing type effectiveness is validity, defined as "the extent to which a test measures what it is said to measure" (Coombe et al., 2007). Validity comprises content relevance or content validity, referred to assessing students with the kinds of language skills the teachers have taught in class and that the course aims at assessing (Genesee & Upshur, 1997). As Genesee and Upshur (1997) explained, the relevance of real-life outcomes cannot always be measured in a course; for instance, since the opportunity to assess students' interaction with native speakers is not always available, simulations are a way to measure students' language skills required for a real-life task. In other words, student performance may be measured "in a restricted range of situations" (Genesee & Upshur, 1997, p.65) and then, the results of this assessment may be generalized to deduce any implications of the course effectiveness for future authentic situations. Therefore, this type of validity gives valuable insight into whether an assignment can be an objective measuring instrument or not. Content relevance cannot be quantified, but only rated as "high, moderate, or low" (Genesee & Upshur, 1997, p. 66).

As for validity for testing procedures, these can be "judged by identifying the possible factors that can invalidate them" (Genesee & Upshur, 1997, p. 67). Tests scores, for example, lack validity if other factors –besides a low mastery of second language abilities– might explain the students' poor performance. Some possible causes, which Douglas (2000) pointed out, are "poor performance due to a lack of understanding of what is expected; insufficient time to carry out the task; lack of interest in the activity; or even the possibility of performing tasks in different ways that are equally valid but unforeseen by the evaluator" (p.68). Evidently, Genesee and Upshur (1997) rightly argued that assessment procedures cannot be expected to be perfectly valid due to human imperfection.

The next section of the present research study will describe the context in which the investigation was held, and inform about the instruments which were used to determine the type of testing that enhanced students' performance the most in the NBJ English course.

## **Method**

### **Subjects**

The News Broadcasters' and Journalists' English course for journalism students was offered for a time period of 13 weeks from August to November 2010

at UCR. As aforementioned, the course was team-taught by Gloriela Chacón Chavarría and Teresa Riestra Carrión; a teacher was in charge of a lesson, while the other teacher became a teacher assistant (TA) offering general technical and language support for a better development of the lesson. Every week students received two classes, for a total time length of three hours per week. There were four units in the course, which focused on speaking and written production; this is shown in the course goals and objectives included in the syllabus. The level of the students was high intermediate. During the needs analysis of the course, 27 potential students were interviewed. Nevertheless, some journalism students had schedule conflicts with other classes, and still others worked late; these overlapping responsibilities affected attendance. In the end, seven students began the course; however, the student population decreased to a minimum of four students, at times having only two present in the class. Due to her job schedule, one of the four students did not always come to class; this affected our results, for the feedback and performance of this student was not gathered in a systematic way.

### Needs Analysis

It was essential for the course designers to find out more about types of assessment in the journalism major, which is part of its “underlying methodology” (Dudley-Evans & St. John, 1998). While carrying out the needs analysis of the target population, the course designers found that journalism students are assessed in their Bachelors’ degree (BA) courses not by means of tests and quizzes but by means of projects, that is, through alternative assessment. As recommended for ESP courses, the course assessment was designed according to the needs analysis and the type of assessment used in journalism courses at UCR, in this case, projects (Dudley Evans & St. John, 1998). Thus, in the needs analysis, our target ESP students were asked if they preferred being assessed through projects, through traditional testing or both. Surprisingly, 55% of the students replied that they would like to be assessed by *both* alternative assessment and traditional testing; the rest of the students stated that they preferred to be assessed by means of projects only.

### Procedure

In order to analyze which type of assessment enhanced the students’ performance a variety of instruments were used in this research study.

Concerning assessment types, the instructors compared the grades obtained on traditional testing and alternative assessment. Traditional testing consisted of two quizzes. Alternative assessment consisted of three projects and a larger final project, as described in the syllabus. The final project consisted of an investigative report, where students had to do some research and conduct and interview to a native speaker.

To better understand the reasons for the students’ scores, the course evaluators administered two self-evaluation forms which compared the scores of their



quizzes to the ones of their project. A final self-evaluation was handed out to the students as a final reflection on the type of assessment which had enhanced their performance more. In the same way as with the course evaluation instruments, these self-evaluation forms were written in Spanish to enable the learners' to voice their opinions. An evident limitation is the fact that the grades of the fourth student are sometimes missing, for she was not able to submit some of the projects of the course. She had very little time and a hectic job schedule.

The rubrics used to assess these projects –both speech and writing performances— were also taken into account. The writing rubrics were used for each of the three projects which assessed written production of journalism subgenres –profile/feature story, press release and investigative report. Finally, these results will be contrasted with the initial preference of students toward assessment type, as was reported in the needs analysis of the NBJ English course.

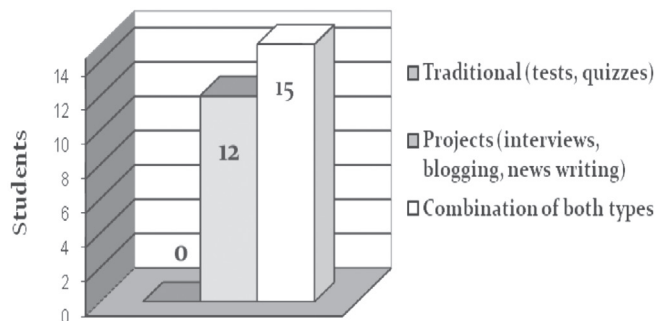
The next section of this research study discusses the main findings of this project and compares the results with the theoretical framework analyzed in the review of the literature.

## Results

The results in this study provided some insight on the type of assessment which enhanced students' performance more during the NBJ English course for journalism students.

As aforementioned, during the Needs Analysis students were asked whether they preferred being assessed through projects –alternative assessment— through traditional testing, or both. From the 27 potential students who were analyzed for the NBJ English course, 55% replied that they would like to be assessed by both alternative and traditional testing; the rest of the students stated that they preferred to be assessed only by means of projects, as can be seen in Figure 1 below.

**Figure 1**  
Preferred type of evaluation for 27 potential journalism students, UCR, April, 2010

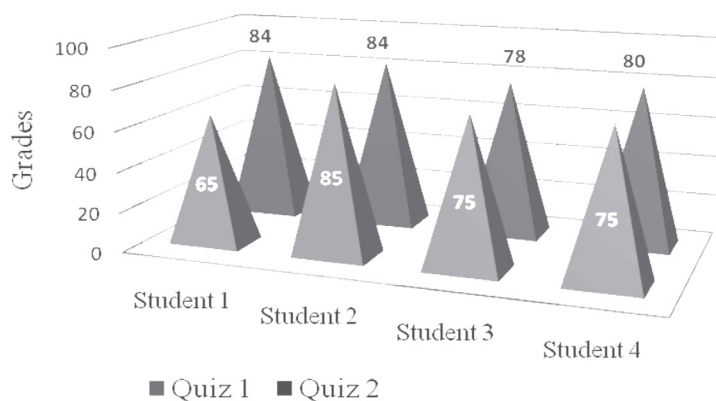


Source: Questionnaire from the Needs Analysis

The students explained the reasons for also choosing traditional testing. They mentioned that they wanted to perceive progress in clear linear outcomes; they said that traditional tests were “very effective to assess grammar points”, in other words, that tests are proof of students’ “serious [individual] preparation for the evaluation.” This preference coincides with ESP clients who “often want immediate, tangible evidence of course outcomes and improvements” (Dudley-Evans & St. John, 1998, p. 210). The instructors then decided to assess the students’ performance by using both traditional testing and alternative assessment. Nevertheless, since these students are usually assessed through projects, better performance through alternative testing was expected. This fact needed to be contrasted with the students’ actual grades.

Regarding traditional testing, the course instructors administered two quizzes, which were progressive tests. Each quiz was worth ten percent of the final grade. The results of these quizzes are shown in Figure 2 below.

**Figure 2**  
**Results of grades from traditional testing: quizzes of 4 journalism students, NBJ English Practicum course, UCR, Second Semester, 2010**



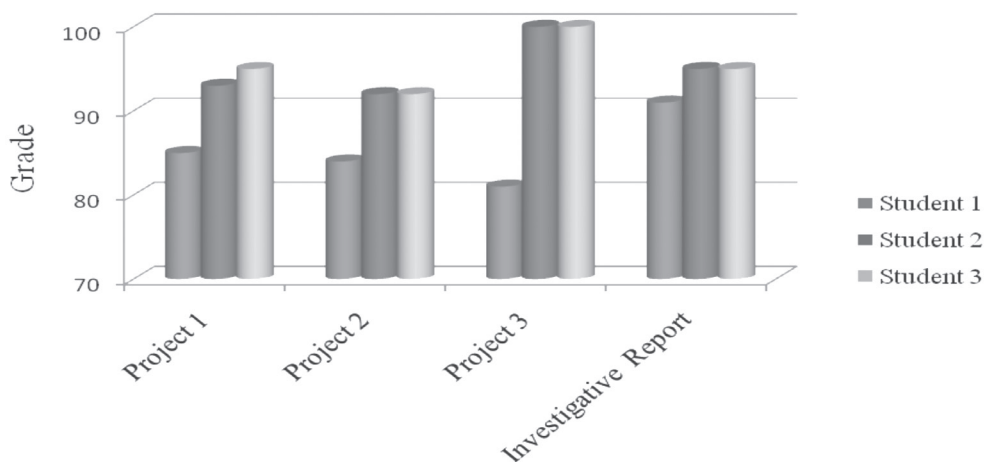
Source: Test scores

These quizzes are progressive tests because they measure class work and particular language skills such as vocabulary, grammar and writing. The grades were very similar, except in one case: Student 1 showed great improvement; her grades increased by 20 points.

As for alternative assessment, the journalism students were assessed through tasks that they would indeed perform in the future. The ESP journalism course provided students with this type of performance assessment. The students were required to carry out two individual as well as two group projects. One of the group projects involved two students in spoken interaction, which was recorded on a video. The advantage of this assignment is that it lends itself to the evaluation of accuracy in the learners’ speech production.

In addition, students have abundant time: students could record themselves the needed number of times, until they arrived at a satisfactory outcome. The remaining projects addressed the writing skill, unmistakably an ability which journalism students need to develop. Through these projects, journalism students were engaged in different kinds of news writing: profile writing, investigative news reports and press releases. They had to submit four different projects: three of them were worth 15% of the final grade –summing up to 45%– while the investigative project was worth 25%. In total, all the projects were worth 70%. The grades of these projects are shown in Figure 3 below. Projects 1 and 3 were performed in pairs, while Project 2 and the investigative report were carried out individually. Student 2 and Student 3 were one language level higher than Student 1; thus, their use of vocabulary and performance was consistently higher in the three projects. Given the fact that project 3 was done in pairs; the investigative report was the best outcome that students accomplished individually.

**Figure 3**  
**Results of grades from alternative testing: projects of 3 journalism students, NBJ English Practicum course, UCR, Second Semester, 2010**

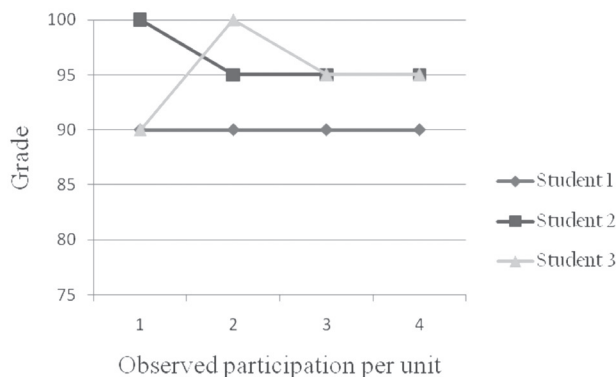


Source: Project scores

In seven out of ten self-evaluation forms, students said that they had handed in good projects. The other three times, the students replied that they had almost always handed in good projects. These behaviors are mirrored in their grades.

The remaining 10% of the final grade was allotted to on-going participation (see Figure 4). On-going assessment is another form of alternative testing. The average grade of these three students was 94%. Thus, it can be stated that students had a high degree of participation during the course.

**Figure 4**  
**Grades from alternative testing: on-going assessment of 3 journalism students, NBJ English Practicum course, UCR, Second Semester, 2010**



Source: On-going participation scores

These are observed performance samples which, as was cited in Brown (2001) “are combined to form an assessment; in-class and extra-class graded work, [are all] alternative forms of assessment (journal, observation, self-assessment, peer-assessment)” (p. 117). In-class, students made peer corrections and self-corrections. This class performance samples gathered through the on-going assessment can be compared to the perception of the students themselves. The students were asked through self-evaluation forms whether they had made an effort to pay attention and participate in class. Ten student self-evaluation forms were completed on three different occasions. The students reported eight times that they had *always* paid attention and participated in class. The remaining two times, two different students indicated that they *usually* had paid attention and participated in class. These comments confirm their excellent participation which was reflected on their grades.

When the results obtained by each student are combined, the alternative testing shows to have given the students higher marks than traditional tests (see Table 1).

**Table 1**  
**Average grades of the three journalism students for each type of assessment NBJ English Practicum course, UCR, Second Semester, 2010**

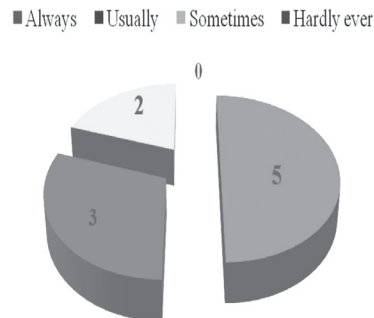
	Traditional Test Average	Alternative Test Average	On-going Assessment	Actual Course Grade
Student 1	74,5	85	90	84,15= 85
Student 2	84,5	95	96	93 = 95
Student 3	76,5	95,5	95	87.85 = 90

Source: Quizzes, Projects and On-going assessment forms

Fortunately, these grades are 70 % of the whole grade; as a result, the final grade of these students was positively influenced by alternative assessment. Since the present research study considers on-going assessment as a type of alternative assessment, its 10% added to alternative tests would sum up to 80% of the grade; this means a greater advantage for the students. There are many causes that can explain students' better performance in alternative assessment. The first is that students were more motivated to write news articles; they felt it was more useful and practical for their future jobs than taking quizzes. On the other hand, the quizzes were demanding, even though they had been practicing their content in class.

An interesting fact concerns the students' self-evaluation on the effort they made and how it is reflected on each of these grades. Half of the students mentioned that their grades *always* reflected their effort; on the other hand, on five occasions, some students replied that grades *usually* or *sometimes* reflected their effort, as seen in the following figure (see Figure 5). More than two thirds then agreed that grades mirrored their own effort.

**Figure 5**  
**Students' perceptions of their real effort being reflected**  
**on the grades of 3 journalism students, NBJ English**  
**Practicum course, UCR, Second Semester, 2010**



Source: Self-evaluation forms

Another item in the self-evaluation forms addressed the reasons for a better performance in their projects compared to the one on the quizzes. Most students commented that they had not studied enough for the quizzes or that they did not like memorization as a way of assessment. One of them mentioned that time was a factor that had affected his performance on the quiz; he was absent the day of the quiz; consequently, he had to take the quiz after the class, and he raced against time. The student who had a difficult job schedule remarked that, since she had not attended all the classes, she felt unable to successfully complete the quizzes. This argument confirms the claim that traditional testing's time constraints in a classroom setting may affect students' performance as discussed in the review of the literature. Douglas (2000) points out a lower validity of test

scores if other factors besides language abilities affect students' performance. Some of these factors include, for instance: "poor performance due to a lack of understanding of what is expected; insufficient time to carry out the task; lack of interest in the activity" (p.68), among others. Thus, the students' real abilities were just slightly reflected on the scores of their quizzes. This was better understood, when contrasted with their performance in the projects.

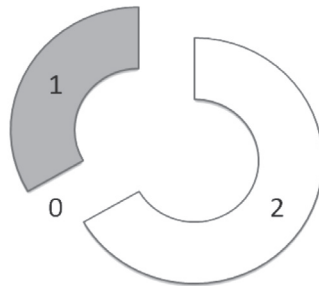
Regarding better scores on their projects rather than on their quizzes, the students mentioned that they had more time to prepare the projects well. Therefore, the product was more satisfactory. In their final self-evaluation, students were asked to identify their best project or quiz and to explain how it had enhanced their abilities. All of their answers pointed at the investigative report as their best performance. They stated that they had shown their own style, had used a variety of vocabulary and sentence structures, and even though the project had been challenging, they were very engaged in it. One student said that she had proofread her work carefully and that she thought it showed clear improvement. These findings are consistent with Brown's (2001) views on the observation of performance samples where teachers perceive "the process [of internalizing meaning] as it is performed" (p.118), as mentioned in the review of the literature.

Another question that the self-evaluations addressed was whether the feedback given by the teachers gave the students a better understanding of their strengths and weaknesses. Some of the answers from students alluded to the correction of pronunciation errors in class and written feedback received after each of their written assignments were returned. The students said that they felt they could write better and that they could speak with more confidence. Many of the course lessons were based on certain pronunciation problems which the course instructors discovered in the students' first project, which was the video broadcast. Consequently, the course content was improved by students' own projects. This is a significant result. The feedback provided by the teachers was more accurate and precise in alternative assessment than in traditional testing. In their quizzes, the students received feedback in out-of-context errors in different grammatical items; their errors were added up and their performance was reflected in a limited way on the actual scores of the quizzes. In contrast, alternative tests highlight students' strengths based on clear performance standards rather than just count the precise number of errors. All of the mistakes found in their projects were seen as a whole and classified in descriptor items as can be seen in the video project rubrics and in the writing rubrics used to assess the students' projects. In this way, the students had a clear idea of the aspects they needed to work on more.

An additional question found in the final self-evaluation form explored the students' opinions about the most appropriate type of assessment in a future implementation of the course. Based on a hypothetical situation, students were asked which type of assessment they would prefer if they had the opportunity to take the course again. As seen in Figure 6, two thirds replied that they would like to be evaluated just by means of projects.

**Figure 6**  
**Assessment type preferences of 3 journalism students on a**  
**Hypothetical enrollment of the same course, NBJ English**  
**Practicum course, UCR, Second Semester, 2010**

Only by Projects   ■ Only by Quizzes   ■ By both



Source: Self-evaluation forms

One student chose to maintain the assessment of the course as it is. This choice is slightly different from the initial preference of the 27 students in the Needs Analysis, who chose to be assessed using both types. Once again, these results seem to indicate that journalism students should be assessed in the same way they are assessed in the UCR major: through alternative assessment.

### Discussion of results

As was expected in this ESP course, alternative assessment enhanced students' performance more than traditional testing, for these journalism students were assessed through tasks that they will perform in their workplace as media professionals. Alternative assessments possess authenticity, as claimed by Hart (2000), because these tasks are "worthwhile, significant and meaningful" (p.9). In the case of this ESP course, students performed target situation tasks: writing profiles, ledes, investigative reports, editing, proofreading, interviewing, doing news casts and participating in a simulated press conference. The fact of having more time to prepare the projects gave students the opportunity "to reveal not only what students know but how well they can put their knowledge into use", as already cited by Hart (1994, p.50), in her explanation of event tasks. Furthermore, the students own unraveling of their skills constitute a positive element when assessing through projects, due to the stated "higher degree of involvement and ownership," (Dudley-Evans & St. John, 1998, p. 195). An evident example of authenticity was the task that the students had for the investigative report. They had the opportunity to interview native speakers, thanks to the contacts provided by the course instructors. They interacted using the appropriate vocabulary and intonation patterns for questions, to mention some beneficial aspects.

In terms of validity on their two quizzes, students were evaluated on the discussed grammar features and respective vocabulary of each unit. As for projects, the needed skills to accomplish them were gradually introduced in class. Students practiced broadcasting, interviewing, question formation, sentence variety, proofreading for punctuation, to mention a few. Hence, content validity was ensured.

The way each student personalizes a task expressing their own abilities and styles will lead them to use the needed learning strategies to carry out a project successfully. As said in the review of the literature, in the accomplishment of these tasks students become autonomous, and that is one of the main goals of any language course program.

### **Limitations of the study**

Evidently, the greatest limitation of this ESP course was the small number of students. As mentioned earlier in this research study, seven students actually enrolled the course on August 2010. However, the student population rapidly decreased to four students. Moreover, the small population affects the reliability of the research results themselves.

### **Conclusions**

In conclusion, the NBJ course students' performance in the different types of assessment seems to indicate that alternative assessment enhanced these students' performance the most for it reflects authenticity of tasks. The journalism students performed better through alternative assessment, a more meaningful way of showing how well they master the language in both writing and speaking tasks. On the whole, the type of assessment used in the NBJ English course positively influenced their performance, for as seen in the results, the percentage of alternative assessment was higher than the traditional testing percentage.

### **Recommendations**

For future implementation of the course, the type of assessment should rely on alternatives in assessment. Additionally, students could be evaluated for their peer feedback and self-assessments; this could be assigned 5% of their grade. Also, the course instructors vividly recommend the use of a portfolio, to place students' writing performances. This is one way to show students' own progress in process. The use of the portfolio as a whole would also be graded.



## Bibliography

- Brown, H.D. (2004). *Language assessment: principles and classroom practices*. Hertfordshire, UK: Prentice Hall.
- Coombe, C., Folse, K., & Hubley, N. (2007). *A practical guide to assessing English language learners*. Ann Arbor, MI: University of Michigan Press.
- Cushing Weigle, S. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.
- Dudley-Evans, T., & St. John, M. (1998). *Developments in specific purposes*. Cambridge: Cambridge University Press.
- Dumitresco, V. (2000). Authentic materials: selection and implementation in exercise language training. *English Teaching Forum* (38), 20-22.
- Genesee, F., & Upshur, J.A. (1997). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Graves, K. (2000). *Designing language courses*. Boston: Heinle & Heinle Publishers.
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. Menlo Park, CA: Addison-Wesley.
- O'Malley, J.M., & Valdez-Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Virginia: Addison-Wesley.
- Shohamy, E. (1996). Competence and performance in language testing. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (136-149). New York: Cambridge University Press.

