

# A Study of the Effect of a Writing Assessment System on Instruction <sup>1</sup>

Ana Muñoz, Universidad EAFIT, Colombia, [apmunoz@eafit.edu.co](mailto:apmunoz@eafit.edu.co)

Martha E. Álvarez Villa, Universidad EAFIT, Colombia,  
[ealvarez@eafit.edu.co](mailto:ealvarez@eafit.edu.co)

## Abstract

A Writing Assessment System (WAS) was validated and then implemented in the adult English program at a language center in a small Colombian private university with the intention to foster positive changes in instruction. To evaluate this impact, the teaching and assessment routines of 28 teachers were examined through portfolios gathered in 2006 and 2007. To assess the teachers' perception of the system a survey was also used and results showed that teachers greatly improved their teaching and assessment practices and that their attitude towards the system was mostly positive. However, some resistance to change was perceived regarding the teachers' attitudes towards their own language knowledge and teaching abilities. The study concludes that for meaningful change to occur, teacher educators and reformers need to understand the beliefs that teachers bring to instructional practices. Based upon this, professional development programs oriented towards reflection could be beneficial to foster desired changes.

## Introduction

In 2005 a research group of the Language Center (LC) at Universidad EAFIT, Columbia designed and validated a Writing Assessment System with the aim of improving teaching and learning writing practices in an adult English program. Intentional actions towards positive washback<sup>2</sup> require: congruity between assessment and curriculum related objectives, authenticity of tasks, detailed score reporting, teachers' understanding of the assessment criteria, and learner self-assessment (Bailey, 1996; Hughes, 2003; Messick, 1996; Shohamy, 1996). The WAS design closely followed these requirements. First, each component of the WAS—writing standards per course, rubrics, conventions, and writing tasks for mid-term and final tests—was explicitly connected to each component; second, the writing tasks were designed by considering authenticity requirements such

---

<sup>1</sup> This is a refereed article.

<sup>2</sup> Washback refers to the influence of assessment on teaching and learning (Hughes, 2003; Wall & Alderson, 1993)

as parallelism to real life situations, consistency with classroom and curriculum related objectives, and the interaction between tasks and students' background (Bachman & Palmer, 1996; Douglas, 2000; O'Malley & Valdez, 1996; Widdowson, 1979); and third, the rubrics were designed to render consistent application ( $r > 0.7$ ) (Muñoz et al., 2006).

The WAS consists of a set of writing rubrics<sup>3</sup> aligned with writing standards for each course, writing conventions to check grammar, vocabulary, punctuation, and spelling problems, and writing tasks for the tests. The system was implemented during the first academic quarter of 2006 after teachers had received training to familiarize themselves with its appropriate use. A three-module course dealing with theory and practice was offered including: 1) definition of the writing ability, 2) planning and design of writing tasks, and 3) consistent use of the rubrics and conventions. In Module 1, different approaches to the teaching of writing were reviewed and a definition of the writing ability for the LC context was presented and discussed with the teachers. Module 2 dealt with a hands-on practical understanding of writing prompts and their connection to curriculum-related objectives. In Module 3, several calibration meetings were conducted where groups of teachers worked together to score samples and to reach a shared consensus. During these meetings, teachers compared their scorings and discussed any differences of opinions they might have had. In addition to the three-module course, a training course was held to guide teachers on how to teach writing and on how to keep writing portfolios.

In 2006, a preliminary evaluation of the impact of the WAS on teaching was conducted. Results showed that teachers were not using the assessment system as required and that they needed to provide students with more detailed feedback by using the assessment tools appropriately.

In this article, I will first present a brief literature review of writing assessment, contending that meaningful assessment can motivate positive changes in the instruction and learning of writing. I will then describe the method and procedures involved in this study and present the findings and discussion. In the final section, I will offer some conclusions and implications for the classroom and for the implementation of future educational programs.

## **Review of the literature**

The primary purpose of assessment is to interpret and to make decisions about students' language ability. Based upon this, it is essential to define the ability or the construct to be measured because this determines what

---

<sup>3</sup> Scoring scales for different levels of proficiency were used to measure different aspects of writing ability: Coherence and cohesion, grammar, vocabulary, spelling, and task completion

aspects of the ability are to be measured and how they are going to be measured. The definition of the construct for the LC includes: 1) the specification of writing standards for each course; 2) the definition of the teaching approach; and 3) the definition of the aspects of language knowledge and ability in the scoring instrument.

The writing standards' specifications at the LC are based on the Common European Framework of Reference (Council of Europe, 2001) and on the particular needs of the local context. For instance, students at the elementary level are expected to be able to fill in simple forms where personal information is required, write short simple postcards, describe people, places, jobs or study experiences, write short imaginary biographies, write informal personal letters, and write stories. At intermediate levels students are required to write simple essays on topics of interest, summarise, report and give opinions, write brief reports, write personal letters and write notes asking for or conveying simple information of immediate relevance. Finally, at more advanced levels, students are expected to write clear, detailed descriptions of real or imaginary events and experiences, write a review of a film, book or play, and write an essay or report which develops an argument or presents an argument for or against a given topic.

The teaching of writing at the institution focuses on three basic aspects: 1) the process students go through when writing (prewriting, drafting, revising, and editing); 2) the accuracy, content, and organization of the writing; and 3) the particular genre the students are producing (letters, essays, biographies, reports, etc.). I believe that a focus on these three basic aspects can help students greatly improve their writing skills by considering the personal writing process, the accuracy of the language used, and the purpose of the piece of writing (Badger & White, 2000; Harwood, 2005).

Finally, the definition of language knowledge and ability was based upon aspects such as linguistic competence, discourse competence and sociolinguistic competence taken from the ACTFL (American Council for the Teaching of Foreign Languages) proficiency guidelines and the IELTS (International English Language Testing System) writing descriptors (see Internet sites in the References for more detailed information).

Once the construct was defined, it was necessary to design the assessment tools that would mirror all the components specified in the construct. This implies the design of the assessment tasks and the scoring scales. Designing tasks calls for a specification of the prompt which refers to the written instructions given to students. The prompt consists of the question or statement students will address in their writing and the conditions under which they will write (O'Malley and Valdez, 1996). The

wording of the prompt may vary in its specification. Based on a literature review, the LC considered that the prompts should:

- a. Be connected to the writing standards for the course.
- b. Include the genre or the purpose of the writing (Weigle, 2002).
- c. Include the audience, either implicitly or explicitly (Weigle, 2002).
- d. Include the process or the steps students have to follow when developing a writing piece. That is to say, the organizational plan or form of presentation which specifies how students are to develop the writing piece as well as the number of words, time allotment, sequence, or number of paragraphs (Hale et al., 1996).

In order to reduce teacher bias and increase the value of assessment, teachers have found that well-designed rubrics (or a scoring scale) can provide such a tool to promote accurate, reliable writing assessment (Stansfield & Ross, 1988; Weigle, 1994). Additionally, teachers need to be trained to reliably apply the rubrics. Sufficiently high regularity in scoring can be obtained by means of proper teacher training. Prior to the scoring stage, teachers should understand the principles behind the particular rating scales they must work with, and they should be able to interpret their descriptors consistently (Alderson & Wall, 2001).

It is widely recognized that well-designed assessments in which there are task authenticity, congruence between assessment and educational goals, detailed score reporting, teachers and students' understanding of the assessment criteria, among others, are beneficial for the learning and teaching process (Bailey, 1996; Hughes, 2003; Messick, 1996). Although different studies have been conducted on the reliability and validity of large scale writing assessments (Novak, et al., 1996; Walberg & Ethington, 1991), little has been investigated about the impact of writing assessment on teaching and learning. For instance, Stecher et al. (2004) studied the effects of a test – the Washington Assessment of Student Learning (WASL) – and a standards-based system on writing instruction in Washington State schools. The researchers found that although the process writing approach changed little before and after the test was instituted, the curriculum (writing conventions, emphasis on audience, purpose, styles and formats) and instructional methods (greatest emphasis on WASL rubrics for student feedback) did change. The study concluded that the WASL influenced instruction positively.

In another study, Lumley and Yan (2001) examined the impact of the Pennsylvania Assessment Policy on writing instruction and teaching methodology. The findings indicate that even though teachers agreed with the type of scoring and characteristics of effective writing proposed by the Pennsylvania Holistic Scoring Guide, they were reluctant to use the state rubrics, descriptors, and writing samples. The authors concluded that there may have been some deficiencies in the support material, or that

teachers used their own evaluation tools, or that they did not adopt the suggested writing approach.

The aim of this research was to evaluate the impact of writing assessment practices on the teaching of EFL writing. More specifically, following the implementation of the WAS, it was hypothesized that teacher writing instruction would improve and that teacher perception of the WAS would be positive.

## **Method**

### Participants

Twenty eight EFL teachers participated in the study. They had taught at the Language Center for at least two years. The teachers received a series of training sessions dealing with the theory and practice of teaching and assessment as well as the presentation of the WAS aims, the definition of writing ability based on the LC writing construct, the planning and design of writing tasks, and the consistent use of the rubrics and conventions among teachers (calibration sessions). In addition, training sessions were held to guide teachers on how to keep writing portfolios and on how to follow the writing approach adopted by the LC.

### Data Collection and Analysis

Teachers' improvement in writing instruction was examined by analyzing 28 writing portfolios gathered from March to October, 2007. Improvement was defined considering the LC approach to writing instruction and assessment. This consists of: 1) congruence between task and writing standards for the course, 2) appropriateness of the prompt, 3) explicitness and elaboration of techniques to generate ideas, 4) understanding of writing conventions, and 5) detailed scoring and feedback. The portfolios were distributed at the beginning of the course and steps were specified to guide teachers in the portfolio process. Teachers were expected to submit them at course end and to include the students' first drafts and final texts. Although teachers were to file students' writings, the purpose of the portfolio was to evaluate their own understanding of the writing process and scoring procedures.

The analysis of writing portfolios was conducted using a rubric designed and validated for this purpose<sup>4</sup>. The rubric measured congruence between task and writing standards, prompt design, explicitness and elaboration of techniques to generate ideas, use of writing conventions, and detailed scoring and feedback. Two researchers conducted the portfolio analysis,

---

<sup>4</sup> To determine validity, the aspects measured by the rubric were aligned to the writing construct as defined for the Language Center (Muñoz, et al. 2006). Furthermore, the descriptors for each aspect in the rubric were progressively adjusted by evaluating different portfolios used for piloting purposes.

first individually, and then in pairs to reach a consensus if discrepancies arose. The degree of suitability of the different aspects examined in the portfolios was analyzed using percentages for each category of an ordinal scale of: **excellent**, **satisfactory**, and **unsatisfactory**. To compare results from 2006 and 2007, a chi-squared test of homogeneity with 2 d.f. and 5% level of significance was used. In order to examine teacher opinion of the WAS, surveys were given to the 28 teachers and 21 surveys were returned. The surveys contained five questions enquiring about perceived changes in the activities used for teaching and learning writing, teaching methodology, improvements in learning and reasons for using the WAS. For each question, six responses (statements) were provided; teachers had to indicate their opinion about the answers given on an ordinal scale of: **disagree**, **undecided**, and **agree**. The degree of teachers' agreement on the different survey statements was also analyzed using percentages for each category of the ordinal scale of: **disagree**, **undecided**, and **agree**.

## Results and discussion

In this section seven themes will be presented based upon the results of the surveys given to the teachers. These themes are concerned with the teachers' perceptions of: improvement in writing instruction, comparison for improvement in writing instruction, reasons for the WAS implementation, teacher changes in teaching and assessment practices, types of extra teacher work required by the WAS, changes in students learning and finally teachers' willingness to implement changes in instructional practices. Each theme is represented by a figure and the results and discussion will follow.

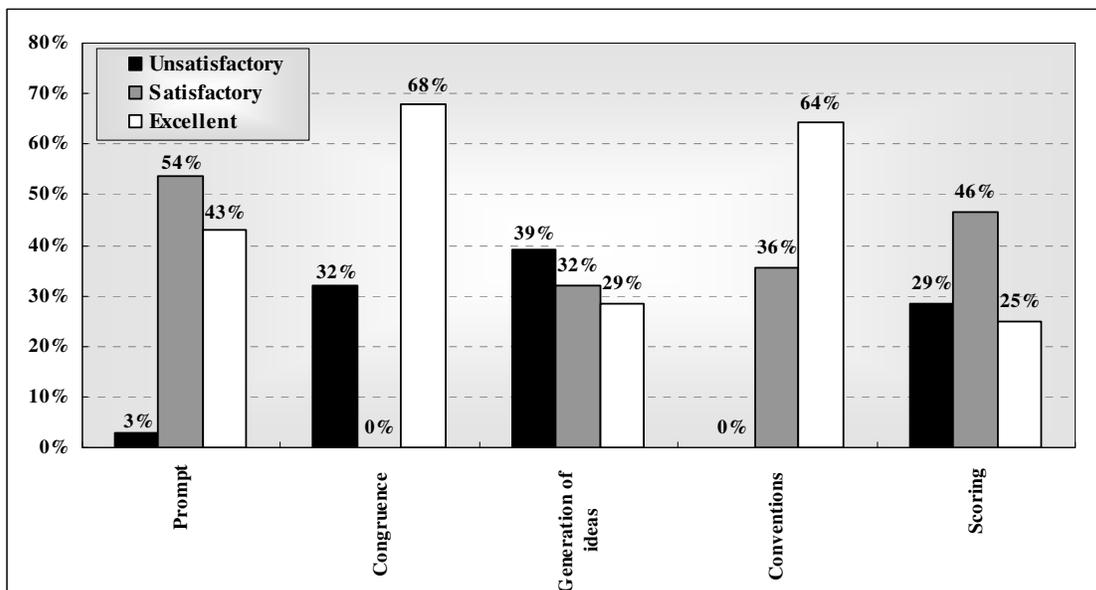


Figure 1. Teachers' improvement of writing instruction

As indicated in Figure 1, 43% of the teachers designed excellent prompts which means that the prompts were consistent with the requirements established by the LC for writing prompts, i.e., specification of genre or discourse mode, audience, and organizational plan. 54% of the teachers did not specify one of the requirements or worded the prompts somewhat awkwardly, but the prompts were considered satisfactory. Finally, 3% of the teachers did not include any of the specifications. With regard to congruence between prompts and writing standards, it was observed that while 68% of the teachers used writing tasks directly related to the writing standards, 32% used activities that had little or no relation to the standard. Even though the writing standards are clearly defined for each course, it seems that some teachers had difficulties in making this connection. This might be due to the preference of some activities given by the teacher or the students without regard to the course objectives.

The analysis further shows that 29% of the teachers were appropriately using techniques to generate ideas, such as brainstorming, listing, mind mapping, etc. These techniques were clearly presented, elaborated, and reflected in students' writings. Although 32% of the teachers clearly indicated the technique used, they did not fully elaborate on it but it was at least partly evidenced in the students' writing; therefore they were considered satisfactory. Still, 39% gave no evidence of a specific technique used. Regarding the revision process, the data revealed that 62% of the teachers made excellent use of the conventions, providing students with precise and appropriate feedback, while 36% used them only satisfactorily perhaps due to their confusion and inconsistent use of some of the correction symbols. When scoring the writings, 25% of the teachers were very specific in assigning scores for each aspect—coherence and cohesion, grammar and vocabulary, and task completion—and descriptors of the rubric and personalized comments to help students understand the score. 46% provided satisfactory scorings, meaning that they assigned scores for each aspect but did not give scores for each descriptor, yet they provided some useful comments for the students. The rest of the teachers, 29%, only assigned global grades and did not comment on the students' writings.

In general, teachers made appropriate use of the WAS, especially in relation to prompt design and the use of conventions. Providing students with well-designed prompts is obviously an important aspect of assessment because students' successful performance greatly depends on how well teachers and test developers design the tasks. Therefore task design is crucial to "allow all candidates to perform to the best of their abilities and to eliminate variations in scoring that can be attributed to the task rather than the candidates' abilities" (Weigle, 2002: 60-61). Similarly, a suitable use of the conventions may affect students' writing in

a positive way because while editing their writing, students need to exercise higher-order thinking skills, such as analysis, synthesis, and evaluation, in order to interpret the symbols and improve their writing. Although students were not formally assessed through the portfolios, the researchers saw how a great emphasis was given to these symbols in order to provide feedback. Very few of the writings contained teachers' comments related to content or style or even praise. In other words, lack of interaction between teacher and students was evident.

In the area of writing standards and tasks, the majority of the teachers utilized tasks that directly measured the standards. However, more awareness needs to be raised regarding the connection between these two aspects. When teachers and students recognize that the writing tasks directly assess the standards and that writing is assessed along clearly articulated levels of performance, teachers will probably be more motivated to change instructional practices to both teach and have students practice around these authentic assessments, and students will be more likely to buy into the value of such work (Natriello & Dornbusch, 1984). With regard to scoring, it seems essential to raise more awareness of the importance of providing detailed scorings. Score reporting may be an influential factor in performance. Several studies confirm that global skills assessments seem to be less reliable than skill specific or behaviour specific descriptors (Chapelle & Brindley, 2002; Strong-Krause, 2000). Furthermore, it is crucial that teachers not simply respond to grammar or content by means of scores but that more personalized comments should be provided so as to maintain a dialogue between the student and teacher. It is also necessary to further encourage the use of pre-writing techniques in order to spark general ideas on the topic.

Comparing the results obtained in 2006 and 2007, it is possible to say that teachers significantly improved writing instruction in most of the evaluated aspects. To determine areas of significant improvement a homogeneity test by chi-square at 5% level of significance was conducted (see Figure 2).

	Excellent		Satisfactory		Unsatisfactory		$\chi^2_{calc}$
	2006	2007	2006	2007	2006	2007	
	%	%	%	%	%	%	
Prompt	17.5	42.9	42.5	53.6	40.0	3.5	20.38
Congruence	57.5	68.0	-	-	42.5	32	1.31*
Idea generation	15.0	28.6	47.5	32.1	37.5	39.3	5.23*
Convention	22.5	64.3	47.5	35.7	30.0	-	32.59
Scoring	2.5	25.0	40.0	46.4	57.5	28.6	61.06

\* Improved, but not significantly

Figure 2. Comparison for improvement in writing instruction

As can be seen in the table, improvement was highly significant in prompt design ( $\chi^2 = 20.38$ ; p-value < 0.05), use of conventions ( $\chi^2 = 32.59$ ; p-value < 0.05), and scoring practices ( $\chi^2 = 61.06$ ; p-value < 0.05). Although there were improvements in using activities connected to the writing standards and in the techniques proposed to generate ideas, they were not highly significant ( $\chi^2 = 1.31$ ; p-value < 5.9 and  $\chi^2 = 5.23$ ; p-value > 0.05 respectively). It is interesting to note that the aspects in which there was more significance have to do with either planning or evaluating the 'product' of writing, whereas aspects where there was less significance, had to do with 'the process.'

Figures 3 to 7 below show the percentages of how teachers responded to the following questions: 1) What are the major reasons for the Adult Program to implement a writing assessment system? 2) What are the major changes that you have perceived in your teaching and assessment practices after the implementation of the WAS? 3) What kind of extra work, if any, do you think the WAS created for you in your teaching? 4) What are the major changes you perceive in students due to the implementation of the WAS? 5) What are the major changes you are willing to make in your teaching in the context of the WAS?

Statements	Disagree	Undecided	Agree
To meet the policies of the LC/University	43	14	43
To improve teachers' language proficiency	42.7	23.8	33.5
To refine assessment practices	4.8	14.2	81
To motivate students to improve writing skills	-	4.8	95.2
To encourage students to self-assess their writing	-	19	81
To encourage teachers to become more aware of their own writing teaching practices	4.8	28.2	67

Figure 3. Reasons for the implementation of the WAS

The majority of the responses (95.2%) indicate that teachers considered the WAS a motivating tool to improve students' writing skills. They also viewed the implementation of the system as a means to refine assessment practices (81%) and foster student self-assessment (81%). Based on these results it is possible to say that teachers strongly agreed with some of the WAS principles, and this agreement, in turn, represents a positive effect on teaching. Two possible circumstances may account for the percentage of undecided answers in the questions concerning teachers' language proficiency (23.8%) and teaching practices (28.2%). First, it is possible that there is a misconception regarding evaluation as exclusively oriented to or developed for students. Upon the implementation of WAS, clarity of the objectives was stressed, in order to improve teaching and learning. However, language proficiency and teaching practices continue to show certain degree of misunderstanding(s) regarding the purposes of this implementation. Second, the uncertainty in

the answers might well be connected to a possible resistance to the proposed changes in teaching and language proficiency. In other words, there seems to be a more positive attitude regarding the activities or instruments that facilitate student change than those which foster teacher change. The uncertainty in the answers did not appear when the responses were related with student improvement or change in any area.

<b>Statements</b>	<b>Disagree</b>	<b>Undecided</b>	<b>Agree</b>
Better planning of lessons	19.1	33.3	47.6
Better understanding of the connection between instruction and writing standards	-	28.6	71.4
More emphasis on writing activities	28.2	4.8	67
More emphasis on the writing process than on the final product	19	14.3	66.7
More emphasis on language accuracy	23.8	23.8	52.4
More effectiveness in teaching and assessing writing	15	-	85

Figure 4. Perceived changes in teaching and assessment practices

The most important change perceived by the teachers was connected to teachers being more effective when teaching and assessing writing (85%). They also perceived that they had a better understanding of the writing standards and instruction (71.4%). However, the 28.6% undecided responses might indicate that there is still a lack of clarity regarding the relationship between standards and instruction. The portfolio results also demonstrated that teachers need to have a better understanding of this connection. It seems that teachers continue to rely more on the pacing rather than on the writing standards which, again, reinforces the idea of a need for a time of transition due to resistance to change, especially on the part of teachers. This is not necessarily a negative aspect but rather a common and expected effect of a process of change in general (Piaget, 1972).

<b>Statements</b>	<b>Disagree</b>	<b>Undecided</b>	<b>Agree</b>
Following the course standards	76.2	14.3	9.5
Doing more lesson preparation	52.4	23.8	23.8
Providing more feedback to students' writings	28.5	9.5	62
Keeping a writing portfolio	14.3	9.5	76.2
Using the assessment tools correctly (rubrics, conventions)	57.1	4.7	38.2
Implementing more writing practices	23.8	23.8	52.4

Figure 5. Type of extra teacher work required by the WAS

Not surprisingly, teachers considered that keeping a writing portfolio added to their teaching workload (76.2%). Likewise providing more feedback was considered extra work (62%). As explained in the data collection procedures, keeping the portfolio implied that teachers had to carefully follow the writing and scoring process. It is important to note that the portfolio was used for the purpose of gathering data and it did not

constitute part of the regular use of WAS. In any case, this represents valuable information that may be considered for future decisions.

In relation to lesson planning and implementing more writing practices, there was no perception that the workload increased significantly. Undecided responses for these statements do not demonstrate an exceptional increase in the workload. If a substantial amount of extra work was required, teachers would surely not have hesitated to answer positively. On the other hand, a response reflecting uncertainty on this respect may, again, represent the teachers' fear of judgement being placed on their skills and their willingness to comply with the WAS requirements. Following the standards did not represent additional effort for most teachers (76.2%), which emphasizes the idea mentioned before that there is a progressive change regarding the understanding of standards and instruction. This progressive change, in turn, may also account for the percentage of undecided responses.

While more than half of the answers did not refer to additional work in using the assessment tools, other teachers perceived that using these tools represented extra work possibly because the WAS requires a thorough understanding of the concepts involved in assessment and a careful application of the instruments. Additionally, it is necessary for teachers to have a high language proficiency in order to correctly identify and measure the possible language problems that need correction and feedback.

Statements	Disagree	Undecided	Agree
More awareness of their writing skills	9.5	4.7	85.8
More willingness to self-assess their writing	19	42.9	38.1
More motivation to write	33.3	28.6	38.1
Improvements in grammar and vocabulary	4.8	23.8	71.4
Improvements in coherence and cohesion	9.5	28.5	62
More understanding of the prompts	9.7	28.3	62

Figure 6. Teachers perceived changes in students' learning

Most of the answers related to changes in students' learning referred to students becoming more aware of their writing skills (85.8%), which is an important positive effect of the system since awareness represents a first step towards change, as mentioned previously. Regarding students' motivation to write, teachers either disagreed (33.3%) or were undecided (28.6%) about this statement. A possible explanation might be related to a perceived lack of intrinsic motivation to learn English associated with the university's bilingualism policy, as expressed by many teachers and students in informal conversations. According to this policy, students need to demonstrate a B2 level of the Common European Framework (2001) in order to graduate from the students' undergraduate degree programs. Students can certify this proficiency through different tests such as the

TOEFL (Test of English as a Foreign Language), IELTS (International English Language Testing System), CAE (Certificate of Advanced English), CPE (Certificate of Proficiency in English), and MELICET (Michigan English Language Institute College English Test). Their proficiency can also be certified by taking and passing all the courses at the LC. This can obviously represent a source of extrinsic motivation.

Statements	Disagree	Undecided	Agree
To follow all the steps to develop writing	4.8	9.5	85.7
To apply the writing assessment instruments	4.8	9.5	85.7
To foster more students' self assessment	-	19	81
To teach towards the writing standards	9.5	19	71.5
To keep writing portfolios	38	24	38
To improve my own writing skills	14.3	23.7	62

Figure 7. Teachers' willingness to implement changes in instructional practices

In general, teachers' responses revealed a positive attitude towards the use of the WAS with the exception of keeping writing portfolios which, as previously mentioned, is not a permanent part of the WAS. The most significant percentage of undecided responses appeared in the improvement of teachers' writing skills (23.7%), which confirms the idea of a resistance to change when teachers were questioned about their language knowledge and abilities. It is also possible that teachers considered that they had little need for professional improvement. It is not uncommon for people to have certain erroneous beliefs and views about themselves.

### **Conclusions and implications**

Significant improvements were found in most of the areas observed, mainly in prompt design and the use of conventions. First, well-designed prompts may influence student learning positively because the task complexity is reduced and successful task completion is increased. Second, teachers' appropriate use of conventions may help students apply higher-order thinking through the revision and edition of their own writing texts. However, some teachers limited their feedback to the conventions without any further comments. Therefore, awareness needs to be raised on a balanced use of the conventions and on a more informative and formative type of feedback. Students who were given informative feedback that explained their strengths and weaknesses were more likely to demonstrate higher levels of intrinsic motivation towards a task (see Butler, 1988; Elawar & Corno, 1985).

Although less significant, improvements were also present in the use of the rubric(s) and in the implementation of writing tasks directly connected to the standards. Providing detailed scoring is beneficial for learning (see Chappelle & Brindley, 2002; Strong-Krause, 2000) because students can

refer to the rubrics and identify possible areas of further work. Likewise, understanding the relationship between tasks and writing standards allowed teachers to plan lessons in accordance with instructional goals and therefore direct students to the accomplishment of the standards. Multiple research studies show that students who perform better are those familiar with their learning goals (Amigues & Guinard-Andreucci, 1981; Bonniol, 1981; Jorba & Sanmarti, 1994).

Teachers also demonstrated a very positive attitude concerning the WAS. They considered the system a useful tool for raising student awareness and for improving writing skills. However, when teachers were asked about improvements in their own language knowledge and abilities, some uncertainty or possible resistance to change appeared. In other words, there was a more positive attitude regarding student change than teacher change.

In any process of change there may appear resistance or opposition. It is likely that the suggested assessment system makes new demands on the teachers' competencies and beliefs as well as the nature and goals of evaluation. The research literature suggests that beliefs and practice are inevitably related, and that teachers may have beliefs that are not compatible with the practices called for in institutional plans (Bliem & Davinroy, 1997; Borko, et al., 1997). It then follows that meaningful change in assessment practices may require changes in teachers' beliefs about such practices. As Fullan (1998: 25) suggests in his innovation theory, "change is a highly personal psychological process." This may require teacher training based upon action research which is defined as a "form of collective self-reflective enquiry undertaken by participants in social situations in order to improve the rationality and justice of their own social or educational practices, as well as their understanding of these practices and the situations in which these practices are carried out" (Kemmis & McTaggart, 1988: 5).

An important change perceived by teachers was that more emphasis was given to language accuracy. While writing texts in terms of accurate structures and lexis is important for learning, the communicative aspect of the language was somewhat ignored. It is then necessary to guide teachers' attention towards aspects such as task completion which calls for the thorough development and elaboration of ideas and the accomplishment of the specific genres and functions measured.

To summarize, the introduction of the WAS proved to be a stimulus or lever for change in some of the areas under research. Continuous efforts need to be made with in-service training and action research programs in order to maintain the system. Moreover, based on the results of this research, the program can be adjusted so that it will be improved for

future use. What is obvious with a program such as this is that communication, training, and teacher involvement are needed for the success and sustainability of new programs.

## References

- ACTFL (2001). ACTFL proficiency guidelines writing. [Internet] Available from: <http://www.actfl.org/files/public/writingguidelines.pdf> [Retrieved July 6, 2007]
- Alderson, J. C. & Wall, D. (2001). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Amigues, R. & Guinard–Andreucci, C. (1981). A propos d'une recherche sur l'évaluation formative en situation éducative. *Bulletin de Psychologie* 35 (353), 167–72.
- Bachman, L. & Palmer, A. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Badger, R. & White, G. (2000). A process genre approach to teaching writing. *ELT Journal*, 54(2), 153-160.
- Bonniol, J. J. (1981). *Déterminants et Mécanismes des Comportements D'évaluation des Preuves Scolaires*. Unpublished doctoral thesis: Université de Bordeaux.
- Bailey, K. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13, 257-279.
- Borko, H., Mayfield, V., Marion, S. F., Flexer, R. J., & Cumbo, K. (1997). *Teachers' Developing Ideas and Practices about Mathematics Performance Assessment: Successes, Stumbling Blocks, and Implications for Professional Development* (Report 523). Los Angeles: Center for the Study of Evaluation.
- Bliem, C. L. & Davinroy, K. H. (1997). *Teachers' Beliefs about Assessment and Instruction in Literacy* (Report 421). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and involvement. *Xz.* 58, 1-14.
- Chapelle, C.A. & Brindley, G. (2002). Assessment. In Schmitt, N. (Ed.), *An Introduction to Applied Linguistics*, (267-288). London: Arnold.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Douglas, D. (2000). *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Elawar, M. C. & Corno, L. (1985). A factorial experiment in teacher's written feedback on student homework: Changing teacher behaviour a little rather than a lot. *Journal of Educational Psychology*, 77(2), 165-173.
- Fullan, M. (1998). The meaning of educational change: A quarter of a century of learning. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins. (eds.), *International Handbook of Educational Change*. Dordrecht: Kluwer, 214-228.
- Hale, G., Taylor, C., Bridgeman, B. Carson, J., Kroll, B. & Kantor, R. (1996). *A Study of Writing Tasks Assigned in Academic Degree Programs*, (TOEFL Research Report No. 54). Princeton, NJ: Educational Testing Service.

- Harwood, N. (2005). The sample approach: teaching writing with Cambridge Examination classes. [Internet] Available from: <http://privatewww.essex.ac.uk/nharwoor/sampapproach.htm> [Retrieved Dec. 7, 2007]
- Hughes, A. (2003). *Testing for Language Teachers*. New York: Cambridge University Press.
- IELTS (International English Language Testing System). (n.d.). IELTS writing descriptors. [Internet] Available from: [http://www.ielts.org/lib/pdf/UOBDS\\_WritingT1.pdf](http://www.ielts.org/lib/pdf/UOBDS_WritingT1.pdf) [Retrieved July 3, 2005]
- Jorba, J. & Sanmartí, N. (1994). *Enseñar, Aprender y Evaluar: Un Proceso de Regulación Continua*. Barcelona: Ministerio de Educación y Cultura.
- Kemmis, S. & McTaggart, R. (1988). *The Action Research Planner* (3<sup>rd</sup> ed.), Victoria, Australia: Deakin University Press.
- Lumley, D. & Yan, W. (2001). The impact of state mandated, large-scale writing assessment policies in Pennsylvania. Paper presented at the annual meeting of the American Educational Research Association in Seattle, WA.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256.
- Muñoz, A., Mueller, J., Alvarez, M., & Gaviria, S. (2006). Developing a coherent system for the assessment of writing abilities: Tasks and tools. *Ikala, Revista de Lenguaje y Cultura*, 11(17), 265-307.
- Natriello, G. & Dornbusch, S. M (1984). *Teacher Evaluative Standards and Student Effort*. New York: Longman.
- Novak, J.R., Herman, J.L., & Gearhart, M. (1996). Establishing validity for performance based assessments: An illustration for collections of student writing. *The Journal of Educational Research*, 89(4), 220-233.
- O'Malley, M. & Valdez, L. (1996). *Authentic Assessment for English Language Learners*. Palo Alta, CA: Addison-Wesley Publishing Company.
- Piaget, J. (1972). *The Psychology of the Child*. New York: Basic Books.
- Stansfield, C. W. & Ross, J. (1988). A long-term research agenda for the Test of Written English. *Language Testing*, 5, 160-186.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13, 298-317.
- Stecher, B., Chun, T., and Barron, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington state. In L. Chen & Y. Watanabe, (Eds.), *Washback in Language Testing* (147-170), Hillsdale, NJ: Lawrence Erlbaum Associates.
- Strong-Krause, D. (2000). Exploring the effectiveness of self-assessment strategies in ESL placement. In G. Ekbatani & H. Pierson, (Eds.), *Learner-directed Assessment in ESL* (255-78), Mahwah, NJ: Lawrence Erlbaum.
- Wall, D. & Alderson, J. C. (1993). Examining washback: The Sri Lankan impact study. *Language Testing*, 10, 41-69.

- Walberg, H.J. & Ethington, C.A. (1991). Correlates of writing performance and interest: A U.S. national assessment study. *Journal of Educational Research*, 84(4), 198-203.
- Weigle, S. (1994). Effects of training on raters of ESL compositions. *Language Testing*, 11, 197-223.
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- Widdowson, H. (1979). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.