The effects of Curriculum-Based Measurement on EFL learners' achievements in grammar and reading

Mansoor Tavakoli
(Associate Professor, University of Isfahan, Iran)
mr.tavakoli14@gmail.com

Samira Atefi Boroujeni
(MA Graduate, University of Isfahan, Iran)
samiraatefi@gmail.com

(Received: 21.06.2012, Accepted: 28.8.2012)

Abstract

The purpose of this study was to examine the effect of using Curriculum-Based Measurement (CBM) on the learners' achievement in L2 grammar and reading in an EFL context and to further investigate whether the students' classroom performances would predict their final exam results. To conduct this study, two classes, each containing 30 female students, were selected among the existing 1st grade high school classes in a high school. Prior to the treatment, all participants were given both reading and grammar tests so as to make sure of any initial differences among them. CBM was randomly implemented in one of these classes (as an experimental group) once a week over a 12-week period to monitor their progress in reading and grammar. At the end of the treatment, all participants took the achievement test. The results of two independent-samples t-tests indicated that CBM significantly improved students' L2 reading and grammar achievement during the term. Also, the results of two linear correlations revealed that CBM outcomes during a semester significantly predicted final exam results.

Keywords: Curriculum-Based Measurement; formative assessment; summative assessment; reading achievement; grammar achievement.

Introduction

Traditionally, assessment is defined as an information-gathering activity. We assess in order to gain insights into learners' level of knowledge or ability and the information gained through assessment procedures would be welcomed, and viewed as an integral component of good teaching. However, a number of questions are in order: Is this uni-dimensional view of assessment deemed sufficient for tapping precise information regarding the student's language ability? Or do we need to apply some alternative procedures in order to

provide multiple sources of information for showing a complete picture of students' progress and ability? This latter question was also the concern of researchers such as Gipps (1994) and McNamara and Deane (1995). Although Such questions have roughly been investigated, it is hoped to shed further light on them in this study. More specifically, this study was to probe into the effect of curriculum-based measurement (henceforth CBM) on Iranian EFL learners' development of L2 grammar and reading.

Curriculum-Based Measurement (CBM), as an objective system of ongoing measurement, is used by classroom teachers to fulfill two purposes, i.e., to assess the students' outcome behaviors and to increase educational decision-making. It is objective because it uses explicit rules and procedures, and it is ongoing because it frequently occurs after a while in the classroom. Curriculum-based measurement procedures were developed to index the effects of instruction on student performance within the curriculum (Christ, 2006). CBM was primarily developed for the purpose of monitoring students' performance in curriculum (Deon, Fuchs, Marston, & Shin, 2001; Fuchs & Fuchs, 1993). This formative function of curriculum-based assessment is thus under-explored especially in the EFL context, and the present study is an attempt in this direction.

Theoretical background

Formative assessment in L2

The last decade has witnessed a significant shift in thinking about the role of assessment in language learning programs (Brindley, 2007). At the forefront of this change has been the increased experimentation with learner-centered 'alternative' assessment methods. From among different possible alternatives has emerged formative assessment, which, as its central premise, sees the goal of assessment as an index to learning processes (Leung, 2004; Hagstorm, 2006; Ke, 2006). In many second and foreign language instruction contexts, assessment practices have increasingly moved away from objective mastery testing of instructional syllabus content to on-going assessment of the effort and contribution learners make to the process of learning(Ross, 2005). Teachers should build in many opportunities to assess how the students are learning and then use this information to make beneficial changes in instruction.

According to Bachman (1990), formative assessment is intended to provide feedback for the ongoing teaching by providing important information regarding learners' strengths and weaknesses that can then be used for subsequent instructional decisions. Teachers are now more aware of the various roles that they can adopt to aid their pupils' learning in a more proactive way than in the past, and so are more focused on pupils' learning as opposed to their own teaching. That is, the focus is more on the changes taking place in pupils' minds as opposed to the effectiveness of the teacher's performance (Harris, 2007). According to Murphy (2006), learning processes can be improved if formative assessment procedures are applied appropriately.

Formative or "for-learning" perspective is quite different from the summative and "of-learning" perspective in terms of theoretical and educational orientation. By definition: summative assessment is more quantitative in nature than formative assessment because it is formally used to assign grades or marks so as to make judgments regarding students' achievement at the end of a particular term

or an educational program. But formative assessment is process-based, and is used for assessing students' learning in the classroom usually for the purpose of keeping records of their progress overtime (Harlan & James, 1997; Ke, 2006; Harris, 2007).

Formative Assessment has a considerable body of research validating its effectiveness. As recent contributions to the literature on second language assessment would suggest, conventional summative testing of language learning outcomes is gradually integrating formative modes of assessing language learning as an on-going process.

However, in spite of all the research that has been done on this issue, there is a need to expand empirically -and theoretically-informed approaches to the investigation of how formative assessment is accomplished in the classroom (Leung & Mohan, 2004). Besides, of key interest is whether formative assessment manifests itself in observable changes in how learner achievement evolves over time and how putative changes in achievement spawned by innovations in assessment practices influence changes in language proficiency. Therefore, empirical research is required on the impact of formative assessment on actual learning success (Ross, 2005). It is now time to turn our attention to one of these formative classroom systems, i.e., curriculum-based measurement, which is under scrutiny in this study.

Curriculum-Based Measurement

There are several types of formative assessment measures, including authentic assessment, portfolio assessment, and performance-based assessment. One type of formative assessment is Curriculum-Based Measurement (CBM). Popham (1993) refers to it as "measurement-driven instruction". This is also called "curricular-driven assessment" (Poehner, 2007), which is described as playing a mediation role between instruction and assessment.

CBM, for the first time, was developed by Deno (1985), who defined it as a frequent measurement of students' curriculum so as to examine the effect of instructional program on the effectiveness of teaching methods and the improvement of learners' success. CBM is a reliable and valid system of measuring students' over time in the classroom. This way, teachers can use the obtained information from the ongoing assessment to monitor learners' progress in due course and resolve "when and how" they can fine-tune instructional objectives to enhance teaching effectiveness (Fuchs & Fuchs, 1993).

Poehner (2007) believes that in this approach assessment procedures are not developed a priori and then imposed upon institutions and classroom teachers but instead emerge from a grounded analysis of instructional interactions and pedagogical practices as observed in the classroom. This approach enables classroom teachers to assume a more active role in determining assessment practices. An added advantage of curricular-driven assessment is that it lends itself well to evaluations of program effectiveness. In other words, because the assessment is derived from curricular objectives, students' performances can be taken as indicators of how well those objectives are being met (Poehner, 2007).

As observed in the literature, The critical feature of CBM is its documented "technical adequacy" (Deno, 1985). This way technically sound measures are significant parts of any assessment system utilized for decision making function regarding

students' test or class performance. The use of CBM procedures for assessing ongoing student progress and for making instructional decisions has been investigated for validation (Stecker, Fuchs, & Fuchs, 2005). Using collective information from multiple assessment procedures results in the reduction of measurement error and permits the teacher to make judgments about whether the student shows that he is on the right track toward achieving the long-term goal and to make decisions correctly regarding the effectiveness of any instructional programs (Stecker et al. 2005). With the help of such indicators of performance, teachers can measure the relevant standing point of an individual at a particular time or can indicate the student's progress over time (Deno, 1985).

Three features distinguish CBM from most forms of classroom assessment (Fuchs & Deno, 1991 cited in Stecker et al., 2005). They are:

- First, CBM is standardized. So the behaviors to be measured and the procedures for measuring those behaviors are specified.
- Second, the CBM testing methods and the difficulty of the tests remain constant, with equivalent weekly assessments spanning a full school year.
- Third, each week's test content reflects the performance desired at the end of the year, and therefore samples the many dimensions of the year's curriculum.

Although there is a robust research literature on CBM in psychology and general education (Stecker et al., 2005), the approach is relatively unknown in applied linguistics. Indeed, with very few exceptions, L2 performance has not been examined from this perspective. "Beginning in the mid-1970s through the early 1990s, research on CBM focused on students with disabilities. It examined whether use of CBM-aided instructional decisions produced differential achievement among students" (Stecker et al., 2005; p. 799). Many studies (such as Stecker & Fuchs, 2000; Fewster & Mcmillan, 2000; Fuchs, Fuchs, & Compton, 2004; Maxwell & Delaney, 2004) reported the efficacy of CBM in improving the achievement of students with learning disabilities in academic skills.

Studies such as Fuchs and Fuchs (1993) and Stecker and Fuch (2000) investigated the effect of using CBM on reading, writing and spelling achievement of the students in primary schools. Results indicated that teachers were able to implement CBM with relatively large numbers of students with fidelity and that their overall satisfaction with CBM procedures was high and CBM could significantly improve early literacy skill achievement of primary students.

An alternative approach to traditional tests is the collection of ongoing data through multiple, brief assessments that allow for consideration of the student's response instructional approach based on both level of performance and growth over time (Francis, 2005). Curriculum-based measurement (CBM) has demonstrated to be a potential method for assessing both level and growth of student performance in skill achievement.

The study done by De Ramirez and Shapiro (2006) was among the first studies that used CBM in L2 language context. Using CBM procedures they examined the performances of eighty-three Learners. The investigation was guided by the following questions: Did students in the English general education

curriculum have significantly higher levels of reading English than did Spanish-speaking ELLs in the bilingual education curriculum? The findings obtained from their experimental study revealed that Spanish-speaking learners of English read English passages more fluently than general education students. Moreover, results regarding the comparison of general education students reading in English and Spanish-speaking ELLs reading in Spanish showed that general education students outperformed Spanish-speaking ELLs in term of fluency. From their findings it can be concluded that CBM proved to be workable concerning reading fluency in EFL contexts.

As it has been said, despite the extensive body of research literature on CBM in general education and psychology, it is relatively unknown in applied linguistics. Therefore, this study was an initial attempt to represent an in-depth treatment of CBM and applications of its principles to L2 context. In other words, the study attempted to examine the instructional role of CBM both as a technique which helps EFL students learn better and as a predictor of their end-of-the term performances.

The present study

The major purpose of this study was to investigate how the results of a CBM procedure used during the educational semester would predict the performance of students at end-of -the-semester summative evaluation. It was further an attempt to investigate the effect of a CBM procedure on the grammar and reading achievement of Iranian high school learners, illustrating how this type of assessment could be accomplished in the classroom context. To this end, this study was conducted to demonstrate the formative value of CBM in assisting students to foster their English grammar and reading ability.

Considering the aforementioned problems and purposes, the following null hypotheses were set forth to be investigated in this study:

1) There is no significant difference between reading achievement of students who receive CBM with those who do not receive any especial kinds of measurement during the course.
2) There is no significant difference between English grammar achievement of students who receive CBM with those who do not receive any especial kinds of measurement during the course
3) There is no relationship between the results obtained based on CBM regarding reading during a semester and the student's end-of-the-semester performances.
4) There is no relationship between the results obtained based on CBM regarding grammar during a semester and the student's end-of-the-semester performances.

Method

Participants

Two classes, each containing 30 female students, were selected from among the three existing 1st grade classes of a high school in a rural area. Regarding educational background, they were nearly the same. Indeed, due to the difficulty of conducting research in a classroom context in our high schools, it was difficult to randomly select and divide the participants into groups as it is in a true experimental design, so an intact-group design was taken to conduct the study. All the participants had three years of experience in English language learning in

grades one, two, and three of secondary school.

Before starting the treatment, a pretest was used to assure the equality of the participants in terms of reading and grammar proficiency. The results of two independent-samples t-tests indicated that there was no significant difference between the learners' reading and grammar performances on pretest and two groups were equivalent in terms of reading and grammar achievement before starting the treatment (Tables 1 and 2; Tables appear after 'references').

Table 1: Independent-Samples t-test of Reading Pretest

Table 2: Independent-Samples t-test of Reading Pretest

Instrumentation

As to the purpose of the present study, a number of instruments were prepared and used which will be described in order.

The achievement test

An achievement test was used as a pretest in this study so as to control the initial differences among the groups in terms of reading and grammar achievement. This achievement test was developed by the researcher. It includes two reading and grammar subtests, 10 items per each subtest. The items of the grammar part are completely related to the English language structures of grade one book (modals such as "could"," had to"," must", and "should", expletive "It", comparison with "as+ adjective+ as structure" and comparative and superlative forms of adjectives), which was taught and assessed through CBM during the term.

The text of the reading comprehension subtest is also at the same length of the last reading text of the grade one book (150 words) as it is considered to be the students' reading goal in the semester. Reading comprehension subtest is an equivalent form of the other reading comprehension tests administered during the semester.

In order to assure the reliability and validity of the test, a pilot study was conducted. In this phase of study, the test was administered to 20 grade one students in the same school. The reliability measures obtained using Cronbach's alpha coefficient was .79 for the reading subtest and .84 for the grammar subtest. Prior to piloting, attempts were also made to make sure of the content validity of the test. That is, the test was sent out for two university professors, who were competent in English so as to render their value judgments on its content and format. The results of their evaluation of the test confirmed the close correspondence between the content of the test and the content of the materials to be tested. So, the test was deemed appropriate as to its content validity.

Curriculum-Based measurement (CBM) tests

This study used CBM as one of the methods of formative assessment which is characterized with the features of standardized testing. CBM uses the typical paper-and-pencil tests in a formative way to monitor students' gradual progress. An explanation of the tests is provided as follows:

Reading tests

In this study, twelve teacher-made ten-item multiple-choice equivalent reading comprehension tests were used in order to measure student's reading achievement during the treatment. These tests were informal tests designed for the purpose of implementation of CBM. All passages used in the reading comprehension tests were about 150 words long, comprising a complete story, taken from different sources and for which equivalence was determined using the Flesch readability estimate. These passages were at the same level of difficulty with the reading text of the achievement test. The readability indices of all the selected texts fell between 70-79 which means the texts were fairly easy. The topic of the passages was relevant to topics used in the1st grade book. Each test included 10 items and students had 10 minutes to complete the test.

Grammar tests

Twelve teacher-made grammar tests were used in order to measure student's grammar improvement during the treatment. All the tests included 10 multiple-choice items. Every week, the students' achievement of one of the structures of 1st grade book that is targeted in the study was evaluated through one of the tests. Each test included 10 items and students had 10 minutes to complete the test.

Procedure

The procedure employed in this study was as follows: the first stage was the administration of pretest (as noted above) to ensure the equality of participants in terms of reading and grammar proficiency prior to the treatment and further to ensure that none of the selected structures have been previously known. After the participants took the pretest, they started the educational semester, which lasted 12 weeks consisting of 2 sessions per week, one session 90 and the other 45 minutes.

Having administered the pretest, the instructor started to employ CBM with the experimental group in order to record pupils' progress in reading and grammar during the course from the first week. Every week, students took a test, including two grammar and reading sub-tests, which lasted 20 minutes. Students' progress was measured by a quiz including both reading and grammar subtests every week, at least four times each month, twelve times a semester. Every week, the gradual improvement of each student over time on reading and grammar was shown on two different graphs. These two graphs also represent the overall improvement of the class on reading and grammar. They are a kind of Linear Regression Graphs and are comprehensive indicator of the students' reading and grammar achievement (see Figure 1 & 2 below). In these graphs, a steep line shows that the students' reading and grammar achievement is improving and a flat line identify inadequate students.

In the courses in which the students' performances are being assessed by the methods of formative assessment progressively, it is incumbent upon the teacher to react to the students' weaknesses to compensate for their inadequacy during the semester (Leung & Mohan, 2004). In CBM, this can be done by taking different actions such as changing instruction, lowering the goals of learning, and providing interventions. To fulfill such an objective, this study used weekly sessions of interventions. It means that every week, the mean score and the standard deviation of the class were calculated and the students whose scores were 1 standard deviation below the mean were identified as weak students. This

empirical evidence potentially informs decisions about which students require a level of intervention that exceeds what is ordinarily provided within the education. Weak Students entered a group tutoring as a kind of intervention.

It should be noted that during the semester, the control group took just a midterm exam which was administered in the midweek of the semester. Formative assessment as done in the CBM group was not implemented there. Moreover, the effect of teaching on the participants' achievement was controlled by using the same teacher and the same teaching method for both control and experimental groups.

At the end of the term, once again the test which was used as pretest was administered as a posttest to find the probable changes in the reading and grammar achievement of the individual learners.

Tutoring sessions

Good formative assessment will support good judgments by teachers about student progress and levels of attainment and it will provide feedback that can be used to help learning. In the present study, tutoring sessions were held to help the students improve their weaknesses in the problematic areas. These 40-minute sessions are based on the principles of the usual instruction and they have a clear scope and sequence of lessons that had followed a weekly test and provide cumulative review and practice. Weak students were tutored once a week, two days after the weekly tests. In order to further clarify what went on during the tutoring, a sample session is provided here in detail. The session was held on the first week. The grammar point of the test was the modal "could" as a past tense of "can' and 4 students were participated in this tutoring session. This session was recorded and it was transcribed later by the researcher.

At first, teacher was going to present a complete and comprehensive explanation of the form and meaning of the modal 'could". To do so, she wrote these examples on the board:

- My father couldn't swim last Friday.
- Jack couldn't ride the bicycle three weeks ago.
- Mary couldn't speak English last year. But she can speak English now
- She spoke with a very low voice, but I could understand what she said.

-Teacher: look at these sentences. What does it mean? Which tense does it refer to?
- [silence]
- Teacher: ok, what do you think about the second sentence? What time does it happen?
- Z (student): present
-Teacher: why? How about three weeks ago? Don't you think…
-F (student): past, past I think. Ago refers to the past time.
-Teacher: right. Let's look at the third sentence. What is your idea about the third sentence M (student)?
- [at first, M translated the sentence into the Persian language] I think "now" refers to the present time but last year….
- [F interrupted] past time again.
- Teacher: so, you think "could" is used for the past tense F, yes?
-F: yes.
-teacher: and what does it mean?
- F: "be able to"
-Teacher: good, exactly. S (student), would you give us an example of what you couldn't do in the past?

-S: [thinking] yes, I couldn't played the guitar.
-Teacher: what's wrong with the sentence students?
- [silence]
-Teacher: "couldn't play" or "couldn't play"?
F: I think "couldn't play"?
-Teacher: why?
- [silence]
-Teacher: ok, because, as you know, we use bare infinitive after Modals, yes.
- Students: yes.
Then, teacher gave the students some completion exercises. They were supposed to complete the sentences using "could", "couldn't", or "can". An example is provided here:
EX: my grandfather was a very clever man. He ……speak five languages.

Having done the exercises, the students were asked to discuss the answers. At the end of the session, the teacher assigned some homework exercises for the students.

Weekly analysis of subjects' performance

In Tables 3 and 4 below, weekly analysis of the participants' performance on reading and grammar exams is shown. In these Tables, class statistics are provided which potentially inform decisions about which students require a level of intervention that exceeds what is ordinarily provided within general education. As it is mentioned earlier, these students should enter a one-session group tutoring. The mean CBM score for the class is shown, along with the standard deviation on that mean and a discrepancy CBM score (average score minus 1 standard deviation) for signaling an inadequate performance level relative to classmates.

Table 3: Weekly analysis of participants' performances on the reading exams

Table 4: Weekly analysis of participants' performances on the grammar exams

As it can be seen in the Tables above, the number of the students who are supposed to take part in tutoring is decreasing toward the end of the semester.

The results of CBM tests are also shown on Linear Regression Graphs. They indicate all students' reading and grammar improvement one by one during the educational semester. Graphs 1 and 2 below are two examples of students' performance on reading and grammar tests during the educational semester.

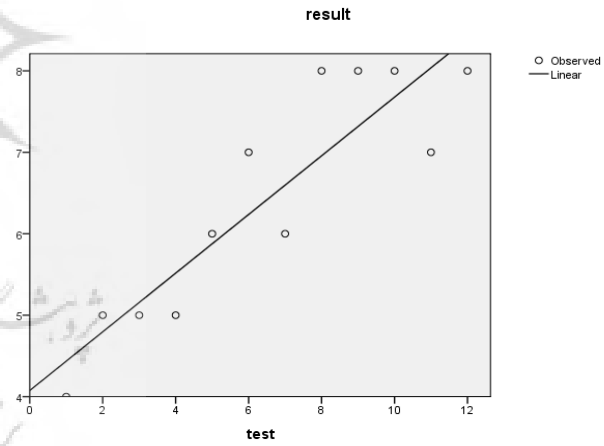

Figure 1: Example of participants' CBM reading graphs

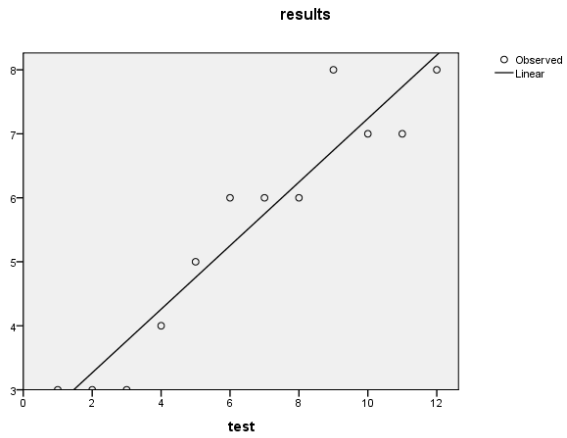

Figure 2: Example of participants' CBM grammar graphs

Results

The effect of CBM on students' achievement in reading and grammar

The first and the second null hypotheses were intended to investigate the CBM effect on the achievement of reading and some English structures of Iranian high school students. In order to investigate these two hypotheses, after the treatment, the posttest was administered to both groups to compare the subjects' performances. Table 5 shows descriptive statistics of the posttest in both groups.

Table 5: Descriptive statistics of the posttest

Two independent- samples t-tests were run to compare control and experimental groups' performances. The results showed that there is a significant difference between the performances of the control and experimental groups on the reading and grammar posttest (Table 6 and 7). Considering descriptive statistics and mean scores, it can be inferred that the experimental group performed better than the control group on the posttest. Therefore, it can be concluded that CBM can significantly improve the learners' achievement and the first and the second null hypotheses were rejected.

Table 6: Independent-sample t-test of the reading posttest

Table 7: Independent-samples t-test of the grammar posttest

The relationship between students' classroom performances and their final-exam results

The third and fourth hypotheses of the study were that there is no relationship between the results obtained based on CBM regarding reading and grammar during a semester and the students' end-of-the-semester performances.

To investigate these hypotheses, two linear correlations were run to explore how well the mean scores of students' performances during the semester can predict their final exam result. (See Tables 8 and 9).

Table 8: Linear regression analysis of reading CBM and the final exams results

Table 9: Linear regression analysis of grammar CBM and the final exams results

As can be inferred from the above Tables, there is a strong positive relationship between the students' performances during the semester and at the end of the term both in reading[r=.825, n=30] and grammar[r=.856, n=30]. Furthermore, the obtained significant levels, for both reading and grammar results, are lower than

significant level of .05 and as a result, reading and grammar test results during the semester can significantly predict final exam results. Therefore, it can be concluded that the performances of students who were assessed by CBM during the semester can be a significant predictor of their final exam results and the third and fourth null hypotheses were rejected.

Discussion

The present study was an effort to apply CBM in foreign language learning context. As it was pointed out in the Background section, much research has been done on CBM in special education, general education and psychology area, but in second language context, it is relatively unknown. Therefore, although all the components of the treatment and data analysis were completely based on the CBM research and theory in special and general education, this study was among the first studies which brought the CBM in to the second language assessment classroom domain.

As a consequence of the analysis of the test results during the term, and comparing the results of the pretest and posttest in the experimental group, it was concluded that CBM can be an effective method for improving the grammar and reading achievement of the learners. The effectiveness of CBM is enhanced if the learners' performances are monitored step-by-step using graphs and if the students' inadequacies are compensated using interventions. CBM caused a gradual growth in the students' performances. This result can be in line with the conclusion which has been made in an overview of CBM research by Stecker et al. (2005). They concluded that teachers can expect significant growth with CBM progress monitoring if they simultaneously implement modifications or interventions when warranted by student data; however, frequent progress monitoring alone did not appear to boost student achievement.

After the treatment, two independent-samples t-tests were run to compare the performance of the experimental and the control group. The results indicated a significant difference between the two groups on the posttest. CBM apparently caused experimental group to improve in comparison with the control group. This finding is in line with many studies which used CBM in general education (e.g., Fuchs & Fuchs, 1993; Stecker & Fuchs, 2000; Fuchs, Fuchs, & Compton, 2004; and Maxwell & Delaney, 2004). De Ramírez and Shapiro (2006) also confirmed the effectiveness of CBM for bilingual children and for learning second language in bilingual schools.

The findings in this regard also lend support to Murphy's (2006) argument that the learning process can be improved if formative assessment is conducted properly. In sum, this study demonstrated the formative potential of CBM to help students foster their English grammar and reading ability in the classroom. Interestingly, this formative function of CBM as an assessment procedure underscores the agreed upon statement that, such formative assessment data can help teachers identify areas of strength or weakness of the students and help them make informed decisions for future teaching and learning process (Weir, 2001; Ellis, 2003).

The third and fourth hypotheses of the study dealt with the investigation of the relation between the formative assessment during the semester and summative assessment at the

end of the semester. The results of the correlation between the test scores during the term and on the final exam revealed that the results of CBM, as one of the methods of formative assessment, can predict summative results. To the best of our knowledge, there are not any studies that have investigated this issue in second language learning domain to date. Yet there is a study that investigated the relation between formative and summative assessment in undergraduates in oral surgery (Anziani, Durham, and Moore, 2008). In that study, no correlation was found between the overall grades for the formative and summative assessment. The obtained results of that study seemingly contradicted the present study. Maybe, this contradiction is due to many factors such as using different methods of formative assessment, different educational settings, very different participants, and the like. In conclusion, because of the controversy observed in the data, this issue needs to be investigated more in different settings in the future.

The findings obtained also underscore the washback effects of CBM in the sense that this way of formative assessment would drive teaching and hence learning. It is also referred to as 'measurement-driven instruction 'by Popham (1993). The measurement-driven instruction is achievable by encouraging the match between the content and forms of the tests and the content and forms of the curriculum. This is referred to as 'curriculum alignment' by Sheppard (1993). In this study, the implementation of curriculum-based measurement during the experiment could roughly help the instructor to match the content and forms of the assessment procedures with those of the curriculum in the high school. Also, the results of the experiment reported above confirm that the match between assessment procedures and the content of the instruction practiced in this study is beneficial to the subjects and hence furthers their learning.

To this end, The findings obtained from the implementation of CBM as a formative assessment procedure can be significantly explained in the light of the conclusion made by Ke (2006), that "Such formative testing allows our teachers to tailor their teaching energies toward continuing instruction and toward providing timely feedback for developmental purposes" (p. 216). It can thus be concluded that more empirical evidence is needed for future research to cast light on the relationship between the components of the curriculum such as the course objectives, program goals, and the washback that practitioners and teachers obtain.

Conclusion

As noted above, two findings were obtained as a result of data analysis in this study. First, the implementation of CBM in EFL context was shown to be useful, that is, the subjects in the experimental group were more successful in improving their L2 grammar and reading as compared with those in the control group. Second, the assessment of the subjects' class performances positively predicted their final scores on the achievement test. Thus, based on the findings obtained in the present study, the following concluding remarks are worth mentioning as to the application of CBM in EFL classrooms:

- Monitoring through using CBM in the classroom allows for the systematic collection of comparative data to determine the significance or effect of instruction and intervention

on individual learners or groups of learners.

- CBM helps second language teachers to have a step-by-step account of their students' progress during a semester. It can aid teachers to judge students' ability, growth, and efforts during the term, and help them before their final exam.
- CBM typically uses materials from the student's curriculum and is administered in a standardized format so that a given student's performance is comparable from one assessment to another. The 'motion picture' that develops as the results of several assessments are plotted on a graph, reflects the student's progress.
- CBM results, as one of the methods of formative assessment, can yield consequences that can significantly predict the course outcome.

Therefore, formative assessment can be a significant predictor of summative assessment.

Moreover, a number of implications can be drawn from the results of this experimental study that may possibly be useful for both practitioners and teachers in EFL contexts. First and foremost, CBM can help teachers to find out how students are progressing in basic skills such as reading, grammar, and spelling since in CBM, each learner has a chance to be assessed and graded several times during the term. Second, the results of the study can be also very helpful for stake-holders, such as students, teachers, administrators, parents and education board; in this way, CBM can help them communicate with each other more constructively. Teachers also can use the CBM graph in conferences with their colleagues and administrators, as it gives them specific information about the students' progress and the success of the instructional methods being used. Finally, CBM, as an overlooked assessment system but efficient method of formative assessment can possibly open new horizons in the domain of second language testing and assessment. Language testing researchers can investigate different aspects of this reliable and pragmatic means of measurement in assessing second language skills.

However, due to the nature of experimental research in classroom context, a number of caveats may still limit the findings obtained in this study. First, since data were collected from a small size population, care must be exercised in generalizing the findings of the study to larger populations. With more participants, the results of the study would be more reliable and the obtained data would be more generalizable. Finally, If random assignment of the participants to different groups were done in this study, the obtained findings would be, for sure, more dependable and generalizable than they are now.

References

Anziani, H., Durham, J., & Moore, U. (2008). The relationship between formative and summative assessment of undergraduates in oral surgery.

European Journal of Dental Education, 12(4), 233-238.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Brindly, G. (2007). Editorial. *Language Assessment Quarterly, 4*(1), 1-5.

Christ, T. J. (2006). Short-term estimates of growth using curriculum-based measurement of oral reading fluency: estimating standard error of the slope to construct confidence intervals. *School Psychology Review, 35*,128-133.

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219–232.

Deno, S.L., Fuchs, L.S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review*, 30, 507–524

De Ramírez, R. D. & Shapiro, E. S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review*, *35*(3), 356-369.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford, England: Oxford University Press.

Fewster, S. & Macmillan, P. D. (2002). School-based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial and Special Education, 23*(3), 149-156.

Francis, N. (2005). Bilingual children's writing: Self-correction and revision of written narratives in Spanish and Nahuatl. *Linguistics and Education, 16*, 74-92.

Fuchs, L.S., & Deno, S.L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488–501.

Fuchs, L. S. & Fuchs, D. (1993). Formative evaluation of academic progress: how much growth can we expect? *School Psychology Review, 22*(1), 27-57.

Fuchs, L.S., Fuchs, D., & Compton, D.L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71*, 7–21.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational measurement*. London: Falmer.

Hagstorm, F. (2006). Formative learning and assessment. *Communication Disorders Quarterly, 28*(1), 24-36.

Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education, 4*, 365-379.

Harris, L. (2007). Employing formative assessment in the classroom. *Improving Schools, 10*, 249-260.

Ke, C. (2006). A model of formative task-based language assessment for Chinese as a foreign language. *Language Assessment Quarterly, 3*(2), 207-227.

Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice and change. *Language Assessment Quarterly, 1*, 19—41.

Leung, C. & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: assessment as discourse and assessment of

Discourse, *Journal of Language Testing*, 21, 335-359.

Maxwell, S.E., & Delaney, H.D. (2004) *Designing experiments and analyzing data: a model comparison perspective*, Second Edition, Lawrence Erlbaum Associates.

McNamara, T., & Deane, D. (1995). Self-assessment activities: Toward autonomy in language learning. *TESOL Journal*, *5*, 17-21.

Murphy, R. (2006). Evaluating new priorities for assessment in higher education. In C. Bryan & K. Clegg (Eds.), *Innovative assessment in higher education* (pp.37-47). New York: Routledge.

Poehner, M. (2007). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Pennsylvania: Springer Science.

Popham, W. J. (1993). Measurement-driven instruction as a 'quick-fix' reform strategy. *Measurement and evaluation in counseling and development*, *26*, 31–34.

Ross, S. J. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics*, *26*(3), 317-342.

Shepard, L. A. (1993). The place of testing reform in educational reform: A reply to Cizek. *Educational Researcher*, *22*(4), 10–14.

Stecker, P. M., Fuchs, L. S. & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: review of research. *Psychology in the Schools, 42*(8), 795-819.

Stecker, P.M., & Fuchs, L.S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research and Practice, 15*, 128–134.

Weir, C. (2001). The formative and summative uses of language test data: Present concerns and future directions. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, et al. (Eds.), *Studies in language testing* 11: *Experimenting with uncertainty—Essays in honor of Alan Davies* (pp. 117-125). Cambridge, England: Cambridge University Press.

Tables

(1)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
RPRETEST	Equal variances assumed	.422	.518	.215	58	.831	.033	.155	-.277	.344
	Equal variances not assumed			.215	57.460	.831	.033	.155	-.277	.344

(2)

		Levene's Test for Equality of Variances		***t-test for Equality of Means***						
		F	Sig.	t	do	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
GPRETEST	Equal variances assumed	.031	.860	-.478	58	.634	-.100	.209	-.519	.319
	Equal variances not assumed			-.478	56.451	.634	-.100	.209	-.519	.319

(3)

	Mean Score	Standard Deviation	Discrepancy Criterion	Number of Inadequate Students
Week 1	2.80	1.400	1.40	6
Week 2	3.60	1.354	2.24	3
Week 3	4.23	1.633	2.62	2
Week 4	4.53	1.925	2.61	2
Week 5	4.97	1.771	3.2	3
Week 6	5.00	2.068	2.94	1
Week 7	5.63	2.092	3.54	2
Week 8	5.87	2.193	3.68	2
Week 9	6.07	2.116	3.96	2
Week 10	6.57	2.144	4.43	2
Week 11	6.57	2.223	4.35	2
Week 12	6.90	2.234	4.67	2

(4)

	Mean Score	Standard Deviation	Discrepancy Criterion	Number of Inadequate Students
Week 1	3.07	1.760	1.31	6
Week 2	4.03	1.712	2.32	5
Week 3	4.33	1.605	2.73	4
Week 4	4.47	1.717	2.76	4
Week 5	4.70	1.878	2.83	3
Week 6	5.30	2.087	3.22	2
Week 7	5.63	2.076	3.56	4
Week 8	5.97	2.205	3.77	3
Week 9	6.67	2.057	4.62	3
Week 10	6.63	2.173	4.46	3
Week 11	6.87	2.129	4.75	3
Week 12	6.90	2.234	4.67	3

(5)

	N	Minimum	Maximum	Mean	Std. Deviation
EXPG-R	30	2	10	7.30	2.136
CG-R	30	1	10	5.10	2.249
EXP-G	30	2	10	7.53	2.177
CG-G	30	0	10	4.07	2.477

(6)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
GPRETEST	Equal variances assumed	.031	.860	-.478	58	.634	-.100	.209	-.519	.319
	Equal variances not assumed			-.478	56.451	.634	-.100	.209	-.519	.319

(7)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
GPOSTTSET	Equal variances assumed	.385	.537	5.758	58	.000	3.467	.602	2.262	4.672
	Equal variances not assumed			5.758	57.063	.000	3.467	.602	2.261	4.672

(8)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	1.949	.729		2.673	.012			
	Rmeans	1.023	.133	.825	7.713	.000	.825	.825	.825

(9)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	1.791	.688		2.603	.015			
	GMeans	1.066	.122	.856	8.757	.000	.856	.856	.856