

The Difference between Formative Assessment through Multiple Choice Questions and True False Questions in Iranian Intermediate EFL Learners' Grammar Learning

Mansoor Fahim

Allameh Tabataba'i University, Iran
Email: Dr.mfahim@yahoo.com

Farnoosh Fahim

Allameh Tabataba'i University, Iran
Email: f_fahim@yahoo.com

Abstract—Assessment is a prominent topic including everything from large scale tests to district benchmark or interim tests to everyday classroom tests. In order to discover what seems to be an overuse of testing, educators should frame their views of testing as assessment and that assessment is getting information rather than making learners stressed. The more we know about the students, the clearer the picture we have about achievement or where gaps may occur. This study attempts to find the difference between using multiple choice type of quizzes and true/false type of quizzes as means of formative assessment in teaching intermediate adult students. What we exactly tried to do was checking that whether learners can improve their knowledge of grammar and have better results in their achievement test based on different exam types they take during the term and in the class. The findings showed that there was no significant difference between the learners exposed to multiple-choice exams and those who were given true/false tests. The type of exam which was used for final exam was even a multiple choice one but this did not lead to a significantly better performance for the ones who had the same exam type during the instructional sessions.

Index Terms—formative assessment, summative assessment, interim tests, achievement test

I. INTRODUCTION

Most of the successful schools engage students in all aspects of their learning. There are many strategies to accomplish this. One of these strategies could student-led conferences. As a teacher or administrator, how is it possible to ensure that the information shared in a student-led conference provides a suitable picture of the students' strengths and weaknesses? One possible way to do so is to balance both summative and formative classroom assessment practices and information gathering about student learning.

Formative assessment is a way of checking the learners regularly. This way, students will have better results during the course and both the teachers and the students can improve or even correct themselves according to the results they get out of the exams. There are various ways of doing formative assessment, such as: summaries, presentations, debriefing, quizzes, journal entries, observations,...etc.

Quizzes assess students for some actual information, related concepts and discrete skill. There is usually a single best answer and not more for each question. Some quiz examples are: multiple choice, true/false, short answer, paper and pencil.

The terms "formative" and "summative" do not seem to be difficult, however the definitions have become confusing over the past few years. This is true for formative assessment particularly. In a balanced assessment system, both summative and formative assessments are inseparable parts of information gathering. Depending too much on one or the other and the reality of student achievement in your classroom results in ambiguity.

Summative Assessments are given periodically, mostly as achievement tests, to determine at a particular point in time what students know and what they do not know. Many associate summative assessments only with standardized tests, yet they are also used at and are an important part of district and classrooms and schools programs. Summative assessment at the district and classroom level is a kind of measure which is generally used as a means of the grading process.

The key is to think of summative assessment as a means to judge the students' learning at a particular time, the learning that is related to content standards. Although the information gained from this type of assessment could be significant, it can only help to evaluate certain aspects of the learning process. Since they are not very regular in time

and occur after instruction every few weeks, months, or once a year, summative assessments are tools to help evaluate the efficiency of programs, school improvement goals, effectiveness of curriculum, or student placement in some specific programs. Summative assessments happen too far down the learning path to provide the people in charge with some needed information at the classroom level and to make instructional adjustments and changes during the learning process. It takes formative assessment to accomplish this goal.

Formative Assessment is part of the instructional process. When incorporated into classroom practice, it provides the needed information to adjust learning and teaching while they are happening. In this sense, formative assessment can inform both students and teachers about student understanding at a point when these kinds of timely adjustments can be made. These adjustments help to ensure students achieve those pre-determined standards-based learning goals within a set time frame. Although formative assessment strategies appear in a variety of forms, there are some clearly different ways to distinguish them from summative assessment.

One distinction is to think of formative assessment as “revision” or better to say “practice.” We should not hold students accountable in “grade book fashion” for skills or concepts they have just been introduced to or what they are learning. We must let them practice. Formative assessment helps teachers decide about or even predetermine next steps during the learning process as the instruction approaches the summative assessment (which could be of any format) of student learning.

Once before, I read somewhere that when you are going to have a driving test in order to get your driving license, it’s not a good idea to pass every step by getting a particular grade on that. If it is so, then the one who is practicing would face some problems that may result in ‘stress’ which is one of the most bothering parts of every test.

Another sharp distinction made here is that formative assessment involves student. If students are not involved or engaged in the assessment process, formative assessment is not implemented effectively. Students need to be involved both as self-assessors or self evaluators and as resources to other students. There are various kinds of strategies that teachers can make use of to engage students. In fact, research shows that the involvement in and ownership of their own work can easily increase students’ motivation to learn something and this great. This does not mean that there is no need of teacher involvement. To the contrary, teachers are very significant and necessary in identifying learning goals; designing assessment tasks and setting clear criteria for success that provide evidence of student learning.

One of the key components of involving students in self- assessment is providing them with descriptive feedback as they are learning. In fact, research shows descriptive feedback is one of the most significant instructional strategies that very easily move students forward in the process of their learning. Descriptive feedback provides students with an understanding of what they are doing efficiently, links to classroom learning, and gives some specific input on how to reach the next step in the learning progression. In other words, descriptive feedback is not a score, a level, or “good job !” A prominent body of research indicates that such insufficient feedback does not lead to improved student learning.

There are many classroom instructional strategies that are part of the successful teaching. When teachers use vocal instructional practice for the purpose of gathering information on student learning, they are using this gathered information in a formative way. This way, formative assessment is pedagogy and clearly cannot be far from instruction. It is what most of the good teachers always try to do.

The distinction lies in what the teachers actually do with the information they gather. It’s not just the matter of collecting information/data on student learning; it’s what teachers have to do with the information they gather.

As we have seen, assessment in English is a challenging work. The issues that are at the center of our attention when talking of summative assessment are very similar to those that we have when we think about formative assessment. What makes English hard to assess also makes progression complex to define. Yet here we are not going to define any concept and what we merely try to do is showing some differences between applying different kinds of strategies.

This study attempts to find whether different test formats that we use to formatively assess our students have any significant influence on their marks in final exam, so the following research question is the main one for this research:

Does the format of the test through which the students are formatively assessed during the class have any significant influence on the result of the scores they receive in the final exam?

II. REVIEW OF LITERATURE

There is considerable research evidence that show effective feedback leads to better learning. Black and Wiliam (1998) have done over 250 studies of feedback carried out since 1988, including nearly all educational sectors. The studies focused on real teaching occasions, and the selection included teacher-made assessments and peer and self assessments. While the bulk of Black and Wiliam’s data came from the school sector, their review and that of others (e.g. Hattie, 1987; Crooks, 1988), provides enough convincing evidence of the value of feedback in promoting learning. Moreover, there is a large body of complementary research studies demonstrating the influence of self and peer feedback on students learning (e.g. Boud, 1995; Boud et al., 1999). Nonetheless, while the work of Black and others has had significant effects on teaching practices in schools (Black et al., 2003) it has so far had much less influence on higher education.

One of the most influential articles underpinning Black and Wiliam review, and works of other researchers (e.g. Yorke, 2003), is that of Sadler (1989). Sadler found three necessary conditions for students to benefit from feedback in academic tasks. He mentioned that the student must know:

1. what good performance is (i.e. the student must possess a concept of the goal or standard being aimed for);
2. how current performance relates to good performance (for this, the student must be able to compare current and good performance);
3. how to act to close the gap between current and good performance.

From this analysis Sadler made an important observation: for students to be able to compare actual performance with the standard one (as suggested by 2), and take action to close the gap (3), then they must already possess some of the same evaluative skills as their teacher (Sadler, 1989). For some writers, this observation has led to the conclusion that, as well as improving the quality of feedback messages, teachers should also put much more effort on strengthening the skills of self-assessment in their students (Boud, 2000; Yorke, 2003). Sadler's argument, that students are already generating their own feedback, also could help account for the common finding that students still make significant progress in their learning even when the external or extrinsic feedback they receive is quite poor.

In spite of the appeal of self-regulation as a construct, it is significant to recognize some basic assumptions underlying its use. While it is assumed that students had better self regulate internal states and behaviors as well as some important aspects of the environment, it does not mean that the students can always have full control on what they are doing. The tasks set by teachers, marking regimes and other course requirements are not under students' control, even though students still have latitude to self-regulate within such constraints. Students often learn in implicit or unintentional ways without explicit regulation as well.

There is much empirical evidence, mainly published in the USA, showing that learners who seem to be more self-regulated are more effective learners: they are more persistent, resourceful, confident and higher achievers (Pintrich, 1995; Zimmerman & Schunk, 2001). In addition, when students are more self-regulated, they think that they have much more control over their own learning and this really helpful for them. (Zimmerman & Schunk, 2004). Also, researches show that any student, even those 'at risk', can easily learn to become more self-regulating (Pintrich & Zusho, 2002). The development of self-regulation in students can be facilitated by structuring learning environments in ways that make learning processes explicit, through meta-cognitive training, self monitoring and by providing opportunities to practice self-regulation (Schunk & Zimmerman, 1994; Pintrich, 1995). The contribution of this article is to identify how formative assessment can affect the students learning and see whether different ways and techniques of formative assessment have different influence on their achievement test.

If we assume that the purpose of a test is to make sure whether or to what extent the learner knows the language, obviously, the basic need to prepare valid tests of language proficiency will be a theoretical question of what it means to know a language. Corder (1975) believed that our ability to do a good job of measuring the learner's knowledge of the language depends upon the learner's knowledge of the language depends on the adequacy of our theory about the language, our understanding of what is meant by knowledge of language.

While the positive role of formative assessment has now apparently been widely accepted in mainstream education circles there has been, according to Black (2000), an absence of an appropriate and coherent program of research to support both the practical and theoretical development of the formative assessment process. For instance, a review of research into formative techniques found that key work in this area showed just little overlap—'it seems that most researchers are not studying much of the literature that could inform their work' (Black, 2000, p. 409). In fact, this difficulty is made of different notions about what the term 'formative assessment' actually means.

One possible result of this changing position of formative and summative techniques is that, on the one hand, we can have a system of high stakes summative national testing that is very unique among European countries; and on the other hand, a 'crisis account' of low standards and students being failed which is not pleasant for teachers or administrators (Gorard, 2001). This might lead a person to consider that the current testing system has proven unsuccessful in some respects, and that, ironically, the assessment method that could have produced the continuous improvement in student performance apparently recommended or better to say demanded by some governments might more properly be formative, rather than summative.

Those who believe in assessment reforms, today focus on the need for a closer substantive connection between assessment and meaningful instruction. They are somehow overreacting against documented distortions in recent decades where teachers in the contexts of high-stakes accountability testing have shaped instructional activities again to conform to both the format and content of external standardized tests, thereby reducing the complexity and demands of the curriculum recommended by governments and at the same time lowering the credibility of test scores gained from the exams. In describing present-day practice, for example, Graue (1993) suggests that both the assessment and instruction are "conceived as curiously separate concepts," a kind of separation that Graue relates to technical measurement concerns.

III. METHOD

A. Participants

The participants in this study were 28 Iranian adult students who were studying English at the intermediate level at one of the English institutes called Kalam. These students were all males and were chosen from among 34 based on the scores they gained in a homogeneity test given to them. Those 28 students were then divided into two groups both exposed to 10 same instructional sessions with different test formats during the sessions in order to find its influence on

the final exam. The students in group one were given a multiple choice test every two sessions to assess their learning and the students in group two were given a true/false test with the same intention. Finally, they all were given a test of multiple choice to see whether the format of the tests in the classroom can make a significant difference between two groups or not.

B. Instrumentation

As it was mentioned before, the participants of this study were 28 male students who had been previously participated in a placement test held by the institute and since they were all put in the same class of instruction, they were regarded to be homogeneous, so there was no exam of homogeneity for them. The tests used here were 5 multiple choice and 5 true/false teacher made quizzes which the students were supposed to take during the instructional sessions as means of formative assessment. A standardized multiple choice test was used as the posttest of this study for both groups and the scores were compared using an independent t-test.

C. Procedure

This study was done based on a comparison between two groups: group 1 and group 2. All the 28 students were exposed to 10 same instructional sessions. The first group which was consisted of 14 students received the instruction and after two sessions they were given a quiz of multiple choice related to the lesson they were taught in the 2 previous sessions. This trend repeated for the rest of the course, meaning that the students received 10 sessions of instruction and therefore 5 quizzes.

The second group had exactly the same situation but instead of multiple choice types of questions they were given true/false ones. The participants weren't aware of this experiment and they were present in all the instructional sessions. The quizzes were all teacher-made and had content validity; they consisted of only 10 related questions to the content of the taught lessons and were used as means of revision and finding the problematic areas but not as something to make them scared or stressed.

IV. RESULTS

Finally, after providing the students with 10 sessions of instruction and 5 related quizzes, the final test was given to the students in order to see whether their performance differ according to the types of tests they have taken based on the taught lessons or not.

After having the students participated in the same final exam, it was revealed that the type of quizzes through which the students were formatively assessed had no significant influence on their level of learning due to the scores they received in the exam. By comparing their final scores through an independent t-test the following result were gained.

TABLE 1.
GROUP STATISTICS

	Testformat	N	Mean	Std. Deviation	Std. Error Mean
testscore	mc	14	17.7321	1.15386	.30838
	Tf	14	16.3036	1.51967	.40615

As it is seen in table 1, the means of both group1 and group 2 are not very much different from each other, but we know that observing the means merely is not sufficient to evaluate the groups' performances. So, the means of the groups were also compared through an independent t-test. The table below (table 2) provided us with more information of participants' performances.

TABLE 2.
INDEPENDENT SAMPLES TEST

		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
Testscore	Equal variances assumed	1.074	.310	2.801	26	.009	1.42857	.50996
	Equal variances not assumed			2.801	24.250	.010	1.42857	.50996

As the table shows, the difference between these two groups is not significant, so the test formats through which the students were formatively assessed could not be very influential on the scores they gained in their final exam.

Actually, the participants with which we worked were not randomly chosen, and they were only a group of students who had registered to study English at an institute. Besides, they were all males, so the results of this study shouldn't be easily generalized to a very much bigger population which has different conditions.

V. CONCLUSION

Assessment is an important part of every educational system. Nowadays, students need to be assessed in a way that they are completely free of stress. Teachers or other people in charge, have to provide the learners with some periodical tests in order to find their progression, but more significant than that is the way they make them ready to be successful in such exams. The trainers need to give them some tests or quizzes as means of formative assessment from time to time and let them know about their weaknesses and improve them. If not, although it'll be assumed that the desired learning has taken place, the learnt lessons will be soon forgotten. In addition to that, formative assessment can be very helpful and the revision will lead to better marks at the time of the final exam.

As mentioned before, there was more emphasis put on the issue of formative assessment and there has been some important distinctions made between summative and formative assessments over the past few decades.

REFERENCES

- [1] Black, P. & Wiliam, D. (1998). Assessment and classroom learning, *Assessment in Education*, 5(1), 7–74.
- [2] Black, P. (2000). Research and the development of educational assessment, *Oxford Review of Education*, 26(3&4), 407–419.
- [3] Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into practice*. Berkshire, England: Open University Press.
- [4] Boud, D. (1995). *Enhancing learning through self-assessment* (London, Kogan Page).
- [5] Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society, *Studies in Continuing Education*, 22(2), 151–167.
- [6] Boud, D., Cohen, R. & Sampson, J. (1999). Peer learning and assessment, *Assessment and Evaluation in Higher Education*, 24(4), 413–426.
- [7] Corder, S. Pit. (1975). *Introducing Applied Linguistics*. Penguin Education, Hazell Watson and Vincyltd.
- [8] Crooks, T. J. (1988). The impact of classroom evaluation practices on students, *Review of Educational Research*, 58, 438–481.
- [9] Gorard, S. (2001) International comparisons of school effectiveness: a second component of the 'crisis account'? *Comparative Education*, 37(3), 279–296.
- [10] Graue, M. E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, 1, 293–309.
- [11] Hattie, J. A. (1987). Identifying the salient facets of a model of student learning: a synthesis and meta-analysis, *International Journal of Educational Research*, 11, 187–212.
- [12] Pintrich, P. R. (1995). *Understanding self-regulated learning* (San Francisco, CA, Jossey-Bass).
- [13] Pintrich, P. R. & Zusho, A. (2002). Student motivation and self-regulated learning in the college classroom, in: J. C. Smart & W.G. Tierney (Eds) *Higher Education: handbook of theory and research* (vol. XVII) (New York, Agathon Press).
- [14] Sadler, D. R. (1989). Formative assessment and the design of instructional systems, *Instructional Science*, 18, 119–144.
- [15] Schunk, D. H. & Zimmerman, B. J. (1994). *Self-regulation of learning and performance: issues and educational applications* (Mahwah, NJ, Lawrence Erlbaum Associates).
- [16] Yorke, M (2003). Formative assessment in higher education: moves towards theory and the enhancement of pedagogic practice, *Higher Education*, 45(4), 477–501.
- [17] Zimmerman, B. J. & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: theoretical perspectives* (Mahwah, NJ, Lawrence Erlbaum Associates).
- [18] Zimmerman, B. J. & Schunk, D. H. (2004). Self-regulating intellectual processes and outcomes: a social cognitive perspective, in D. Y. Dai & R. J. Sternberg (Eds) *Motivation, emotion and cognition* (Mahwah, NJ, Lawrence Erlbaum Associates).



Mansoor Fahim was born in Iran in 1946. He received a Ph.D. in TEFL (Teaching English as a Foreign Language) from Islamic Azad University in Tehran, Iran in 1994, an M.A. in General Linguistics from Tehran University in Tehran, Iran in 1978, and a B.A. in English Translation from Allameh Tabataba'i University in Tehran, Iran in 1975.

As for his professional background, he was the chairman of the EFL department at Allameh Tabataba'i University from 2003 to 2007 and a member of the faculty of English Language and Literature at Allameh Tabataba'i University in Tehran, Iran from 1979 to 2008 when he was retired as an associate professor of TEFL. He has also taught English at a welter of universities and language schools. At present, he runs Psycholinguistics, Applied Linguistics, First and Second Language Acquisition, and Discourse Analysis courses at M.A. and Ph.D. levels at a number of universities in Iran, including Allameh Tabataba'i and Islamic Azad Universities. Moreover, he has several published articles and books mostly in the field of TEFL including: Fahim, M. & Nezakatgoo, B. (2006). *GRE general words for graduate and post-graduate students*. Tehran: Rahnama Publications. Ghobadi, A. & Fahim, M. (2009). The effect of explicit teaching of English 'thanking formulas' on Iranian EFL intermediate level students at English language institutes', *System*: 526–537. Fahim, M., Bagherkazemi, M. & Alemi, M. (2010). The relationship between test takers' multiple intelligences and their performance on the reading sections of TOEFL and IELTS. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 1 (3) : 27–41.

Dr. Mansoor Fahim is currently a member of the editorial board of the Iranian journal of Applied Linguistic Studies, Sistan & Baloochestan University, Iran; Journal of Language Studies, Shahrekord University, Iran; and Journal of English Language Studies, Islamic Azad University, Tehran, Iran.



Farnoosh Fahim was born in Tehran, Iran in 1987. She is a M.A student in TEFL(Teaching English as a Foreign Language) at Allameh Tabataba'i University. She also received her B.A in English literature from the same university in Tehran, Iran in 2009.

As for her professional background, she has been teaching English for about 6 years, since 2005 until present, in some English institutes in Tehran and during these years she has taught many textbooks including those designed for business people and IELTS candidates. She has also held several phonetics workshops and panel discussion sessions in English across the country for Iranian adult advanced EFL students.

Ms. Farnoosh Fahim is currently teaching business English and IELTS in some English schools and working on her dissertation at the same time.