Involving language learners in assessment: A new paradigm

Andrew Finch

Kyungpook National University

Abstract

The field of language assessment has been slow to catch up with research findings in other areas of language learning, largely because of the impracticability of applying concepts such as autonomy, collaboration and student-centered learning in the high-stakes context. Because of this, standardized tests around the world continue to use a summative, product-oriented, competitive format that has been found to be lacking in pedagogic terms. Given the importance of such tests to the lives of the language learners, it is inevitable that teaching in this situation becomes test-driven, motivation becomes extrinsic, and teachers instruct their students in 'test-taking' strategies. Learning how to pass the test becomes more important than gaining mastery of the target language, and the development of higher-order thinking skills is sacrificed to memory-based acquisition of discrete grammatical constructions. The final 'result' is beneficial to nobody except university administrators, who need to manage the intake of students each year (gate keeping).

In view of this situation, this paper examines whether it is practical for high stakes testing to take on a different form, more closely suited to the findings of language-learning research. In particular, it investigates how current theories regarding student-centered learning might be applied to internal and external assessment in schools. (204 words)

1. Introduction

The topic of involving students in assessment is not a new one for TEFL (Teaching English as a Foreign Language) professionals. Academics such as Oscarsson (1997), Blanch (1988), Blue (1988) and many others have investigated the topic extensively, so that 'self-assessment' and 'peer-assessment' are terms familiar to most language teachers in the 21st century. However, in the current testing frenzy that is sweeping education the world over (Eason, 2005; Gardner, 2000) with dubious results (Popham, 2005), the need for the field of assessment to catch up with important findings in ELT (English Language Teaching) research has become more urgent. This paper therefore investigates how modern educational concepts regarding student autonomy, collaboration, meaning, authenticity, emotional growth, and development of higher-order thinking skills might be incorporated into testing, on the assumption that assessment and teaching should supplement and augment each other as complementary aspects of learning, both at local and at standardized levels of assessment. This topic includes Classroom-based Assessment (CBA), in which portfolios, presentations, self/peer-assessment, projects and learner journals can be effectively employed. It

also extends to performance assessment and test design. Finally, a logical application of current educational thinking leads to the design of review tests and even end-of-semester tests by the students themselves.

  It might seem that the ongoing and personal nature of such innovations can only be appropriate at the local, classroom level, with teachers and students cooperating in meaningful, authentic assessment (Kohonen, 1999). However, portfolios and presentations on learning projects are gradually becoming accepted as indicators of higher-order learning at a national level in other countries. Thus, part of the forward-thinking that must take place now in the testing arena in Korea is to realize that High-stakes tests are not the only way to evaluate schools and students. As Nichols and Berliner (2005) point out:

> Finland, the highest achieving country in the world in reading, mathematics and science, have [sic] no standardized tests that resemble ours whatsoever, though they use teacher made tests in their classroom and school accountability system. Their system uses high standards for allowing teachers into the profession, awards high pay and bestows high status to those that enter teaching, provides rigorous and extensive professional development for the teachers, and depends on trusting relationships to improve academic achievement. Clearly there are highly successful models of how to build a national school system that we should study before assuming that our corrupting high-stakes accountability system is the only one that will work. (Nichols & Berliner, 2005, pp. 165-166)

2. The situation

  English Language Teaching (ELT) has experienced many changes and paradigm shifts during its brief lifetime (Graddol. 2007; Legutke & Thomas, 1991, pp. 4-5; Richards & Rodgers, 1986 p. vii; Underhill, 1989, p. 250). The propositional-process paradigm shift documented by Breen (1987), for example, was one of a number which led to the student being placed at the center of learning, managing the learning process. In another tectonic movement, work in the early part of the 20th century identified group work and collaboration as more effective than individual competition in terms of effectiveness of learning (Vygotsky, 1978; Kohn, 1992).

  At the root of many changes has been the problem that students do not learn what teachers teach (Allwright, 1984) and cannot apply the things they have learned in school to real life situations. This has led to various shifts in educational thinking:

- from a utilitarian view of language education to one which aims to promote the overall development of future citizens (Sano *et al.,* 1984, p. 170);
- from teacher-centered classrooms to learning environments which place the learner at the center of the learning process (Tudor, 1996);
- from transmission of knowledge towards transformation of knowledge and to integration

of knowledge with existing personal constructs and meanings (Kohonen, 1999, p. 280);

- from education from the neck up (Rogers, 1975, pp. 40-41) to recognition of affect (anxiety, confidence, motivation, attitudes to learning) and emotion as important factors controlling the learning process (Stern, 1983).

Language assessment has been slow to catch up with these research findings, largely because of the impracticability of applying concepts such as autonomy, collaboration and student-centered learning in the high-stakes context. Because of this, high-stakes tests around the world continue to use a summative, product-oriented, competitive format that has been found to be lacking in terms of pedagogic justification. Given the importance of such tests to the future lives of the language learners, it is inevitable that teaching becomes test-driven, motivation becomes extrinsic, and teachers instruct their students in 'test-taking' strategies. Students learn how to pass the test rather than how to use the target language, and the development of higher-order thinking skills is sacrificed to memory-based acquisition of discrete linguistic knowledge. However, changes which occurred in the 20th century in social science, psychology, philosophy, and political science, indicate that in modern society, learning and understanding metaskills (problem-solving, critical thinking, etc.) are more important than knowledge (Finch, 2006).

The test-driven classroom is responsible for four forms of alienation (Stevick, 1976), which have produced the failure of modern language teaching:

1. alienation of the learners from the materials;
2. alienation of the learners from themselves;
3. alienation of the learners from the class; and
4. alienation of the learners from the teacher. (Stevick, 1976, p. 225)

Other disadvantages of high stakes testing include the excessive time needed for testing, and preparing for the test, along with the decreasing range of subject matter areas that are assessed. It is a sad fact that subjects valued by many citizens and educators (e.g., creative writing, art, social studies) are often dropped from school curricula, with students learning more and more about less and less. However, intensive test preparation and narrowing of the curriculum both serve to distort the things being measured. Students who study test-preparation strategies 12 hours a day are not providing information on their linguistic abilities when they take the final test. In this case, the high-stakes indicator of learning has become corrupted, and is virtually meaningless.

This leads us to the issue of the morale of teachers and administrators. Whatever they think about the test-driven classroom, many high-school teachers have no choice but to 'buy into it,' since to do otherwise would anger parents, students and politicians. The result in the USA is that "Good administrators and good teachers are being driven out of the profession" (Nichols & Berliner, 2005, p. 169). When we consider the fact that teaching is a vocation which traditionally attracts people who are extraordinarily dedicated to helping children – a workforce that other professions can only dream of - "Reform minded governments would do well to consider what is lost by squandering

such a resource … and what the costs would be of finding an equally effective replacement" (Leithwood, et al. 2002, p. 115).

The final 'result' is beneficial to nobody except university administrators who need a way of controlling the intake of new students each year (gate keeping). As Darling-Hammond points out, high stakes testing has "failed wherever it has been tried" (Darling-Hammond, 1997, p. 238), and such tests do not measure the real goals of education.

> Standardized tests can't measure initiative, creativity, imagination, conceptual thinking, curiosity, effort, irony, judgment, commitment, nuance, good will, ethical reflection, or a host of other valuable dispositions and attributes. What they can measure and count are isolated skills, specific facts and functions, the least interesting and least significant aspects of learning. (Ayers, 1993, p. 116)

A radical rethinking of assessment policies in Korea is therefore necessary, since the present model would be unsatisfactory even if it worked! Imagine a society full of A⁺ students as defined by traditional education. Who would drive the trains, clean the streets, grow the food, and deliver the newspapers? Such roles are integral to society, yet their artisans are seen (and perceive themselves) as unsuccessful products of the educational system. Young people who possess practical skills are forced to attend secondary institutions that tell them they are failures. They then move on to Technical Colleges, learning skills that are the lifeblood of the community, but which are seen by everyone (including themselves) as 'second best' options.

Educational theorists are currently addressing the problems associated with an under-performing education system by revisiting ideas that Rogers, Dewey, Bruner, Frière and Vygotsky were expressing before Applied Linguistics was born. Indeed, a holistic view of education, which can be traced back to Aristotle, Plato, and early oriental philosophers, represents a return to basic principles, rather than simply another fashionable trend. In Korea, the ideal of *Hongik-Ingan* (contributing to the overall benefit of humankind) has always been at the heart of education. Such an emphasis must surely be given utmost urgency in these times of natural and man-made disasters; times in which society, in its deification of monetary gain, has neglected moral education; times in which an ethical awareness must extend to every aspect of life.

3. Questions for assessors

Language education is typically poor in producing learners who can be termed successful, even within the narrow criterion of linguistic proficiency. When we consider the additional need to facilitate affective, cognitive, social and cultural learning, then the question "How can language be taught and assessed effectively?" must be exchanged for "How can the language classroom and language tests become instruments of positive attitude change?" In other words, "How can language classrooms and assessment mirror changes in social development, and produce future citizens

equipped to take on the challenges of a century in which the only constant factor will be change?"

Legutke & Thomas (1991, pp. 7-10) pose a number of questions which are relevant at this point. These questions have been adapted to the purpose of this paper, but retain their original meaning. It is interesting that they do not mention assessment in their original form, being concerned only with the learning environment. By adding a focus on assessment, however, they become more meaningful to teachers and assessors, since they acknowledge and emphasize that assessment is an integral part of teaching and learning – that the ultimate goal of assessment is to provide information to teachers and learners so that learning can become more effective.

Question 1: Is it possible and feasible to turn learners' classrooms [assessment] into whole-person events, where body and soul, intellect and feeling, head, hand and heart converge in action?

Question 2: Can second-language (L2) learning [assessment] be a satisfying activity in itself? What adventures and challenges are possible under the very conditions of L2 learning [assessment]?

*Question 3: W*hat needs to be done to regain some of this creative potential in the L2 classroom [assessment]? Do we have to consider individual and cultural differences?

Question 4: What needs to be done to create situations and scenarios where communication in the target language [assessment] could be more meaningful? What are the roles of teacher, learners, topic and input in such scenarios? Could even inter- and intra-student discourse be carried out in the target language?

Question 5: What needs to be done to develop in learners such a capacity for critique? How can they become co-managers of their learning and participate in their own teaching [assessment]? How do we create the learning space so that learners can take initiatives to pursue their own learning [assessment] for their own benefit, and to discover their own learning styles? (Legutke & Thomas, 1991, pp. 7-10)

These questions highlight both the problem and a means of addressing it. It is no longer defensible for educators to use discrete-item testing of dubious linguistic constructs. Instead, the need to understand performance itself and the processing (and affective) factors which influence it, require reflective forms of assessment in instructionally relevant classroom activities (communicative performance assessment, language portfolios and self-assessment), which focus on integration of instruction and assessment. In this two-way process, the essentially interactive nature of learning is extended to the process of assessment (Williams & Burden, 1997, p. 42), examining what learners can do with their language, through real-life language use tasks (cf. Weir, 1998, p. 9).

The four alienations identified by Stevick can be reduced if classroom events and assessment include:

- learning in the here and now;
- student-offered material as the basis for learning and practicing language structures;

- interpersonal sharing; and
- self-awareness and realization. (Galyean, 1977)

4. Process-based, naturalistic assessment

Formative, process-based evaluations are an attempt to integrate teaching with assessment, and focus on what is actually going on in classrooms, rather than what is assumed to be going on: "One can then determine not only whether a program really works, or works better, but if so, why, and if not, why not" (Long, 1984, p. 422). This need to assess what happens in the classroom from the perspective of the participants has given rise to forms of illuminative and naturalistic (non-scientific) evaluation (Williams & Burden, 1994, p. 23), in which the evaluator is involved in the day-to-day working of the course, and data is yielded that assists and guides decision-making. The role of the evaluator/teacher is to produce an "interpretation of a highly complex system" (Parlett, 1981, p. 421), understanding the questions raised by participants rather than setting his/her own questions for investigation, and taking into account the background, culture, politics, and aims of different participants, as well as the hidden curriculum and opinions of the teacher. (Shin, 2003).

Naturalistic inquiry is shaped by the belief that there can be no meaningful separation of facts from values, and that "phenomena can be understood only within the context in which they are studied" (Guba and Lincoln, 1989, p. 45). According to this approach, the language course is seen as a process that is continuously changing.

A variety of information-gathering techniques are used in naturalistic assessment: i) interviews; ii) questionnaires; iii) observation; iv) diaries; v) student records; and vi) portfolios. Naturalistic enquiry can thus provide important information to the most important people in the learning process – the students and the teacher. From the point of view of the students, there can be attention to product as well as process, in that they can have a learning journal (dairy) and one or more portfolios as evidence of the growth that has occurred during the language course. They might also have videos of projects designed and performed by them. These will all assist in the formation of positive attitudes to learning, and will therefore improve the quality of learning itself (success breeds success).

Naturalistic assessment happens all the time. The teacher sets up a non-threatening environment, with portfolios, journals, self-assessment, etc., and then observes the results and the process. As time passes, trends appear, and it becomes evident that certain aspects of learning need extra attention. At that time, the teacher can adjust his/her teaching accordingly, and repair the learning process at first hand, without delay, and on an individual, group or class basis. The tools for doing this are the assessment instruments:

- Learning journals can uncover anxieties and emotional problems that are interfering with learning;
- portfolios can show that organizational skills have been acquired;

- observations can bring interaction issues to light;
- open-ended interviews can reveal concerns previously unimagined by the teacher; and
- well−crafted performance assessments give students the chance to show what they know and can do and to provide teachers with the tools to assess these abilities. (Furger, 2002)

5. Negotiation

5.1 Learning conferences

The traditional view of assessment at every level is external and 'teacher-centered', in that the teacher or the external examiner designs and administers the tests. However, if we apply principles of autonomy and student-centeredness to assessment, we can see that students need to be involved in assessment at every point. This leads to the concept of "learning conversations" (Harri-Augstein & Thomas, 1991), in which students discuss the learning that has occurred up to that point, and make further learning goals based upon their conclusions. Such conversations can occur as a regular part of the coursework, or can be used as content-based oral performance assessments, with the teacher observing the language used by the students. Test-anxiety is lowered by allowing students to attend in self-selected groups and to determine the format of the discussion. As they talk about the semester and their performance in terms of learning goals, they refer to their textbooks, portfolios, speaking journals, and any other materials they choose to bring along. In part 1 of the evaluation session, students talk about their initial learning goals and whether these have changed over the semester. In part 2, they then discuss the projects they devised and performed during the semester, and the teacher is allowed to join in the conversation to offer advice on issues that arise. Students are thus encouraged to see this as a useful (and painless) learning tool, rather than a proficiency test.

5.2 Negotiating grades

Another way to involve students in their own assessment is to ask them to discuss their grades. Grade negotiation is an interesting method of promoting realistic attitudes to learning, and functions in a similar format to the learning conversations of the previous section (the final learning conversation in a semester can easily contain grade negotiation). Figure 1 shows information that can be given to students prior to the assessment session and Figure 2 shows a self-assessment sheet.

FIGURE 1: Grade negotiation: Information sheet (Finch & Shin, 2005, p. 501)

Grade Negotiation

Self-assessment and peer-assessment are used during the course, so that you can learn how to assess yourself realistically and honestly. This is an important life-skill. Now, please think about your performance this semester:

- Have you improved?
- How is your portfolio?
- How is your Learning journal?
- How was your presentation?
- How was your participation?
- How was your cognitive growth?
- How was your affective growth?
- How was your social growth?
- If you could give yourself the final grade, what would it be?

Growth is an important criterion. It is normal to have problems such as overwork, lack of time, other commitments, lack of confidence, not feeling good, etc. How did you deal with these problems? Did you keep doing your best? Did you learn to trust yourself and to trust other people? Did you cooperate with your presentation group members? Can other people count on you? What do your group members say about your teamwork?

Your proficiency level at the beginning of the course is important only because it is a starting point. Do not compare your level with other people. Think about how you have developed from your starting point. What have you learned (cognitively, affectively, and socially)? What have you done this semester?

FIGURE 2: My grade: Self/peer-assessment (Finch & Shin, 2005, p. 502)

Self/Peer-Assessment: My Grade

1. Pairwork:
- Look at the Grade Negotiation Table on this sheet.
- Talk about the topics in pairs.
- Agree on a grade for yourself and your partner.

2. Groupwork
- Talk about the topics in groups.
- Tell the others in the group why you think you deserve your grade.
- Agree on grades with your group members.
- Write your grade in the Class Journal.

9, 10 = Wonderful (I did my best all the time and I overcame my difficulties.)

7, 8 = Good (I did my best most of the time, and I had some valuable learning experiences.)

5, 6 = OK (I worked hard, though some problems were too much for me.)

3, 4 = Below OK (I did not come up to my usual standards. I will be more diligent next time I have the opportunity.)

0, 1, 2 = Poor (I could have done much better. I need to organize myself and my studies)

Grade Negotiation Table	0 to 10
1. My participation in class activities during the semester was …..	
2. My portfolio (preparation, organization) was …	
3. My Learner Journal was …	
4. My presentation lesson (preparation, collaboration) was …	
5. My cognitive development (linguistic and thinking skills) was …	
6. My affective development (confidence, motivation) was …	
7. My social development (cooperation, responsibility) was …	
8. My self-assessments (honest, regular) were …	
9. The peer-assessments done by me (honest, regular) were …	
10. My entries in the Class Journal (honest, regular) were …	
TOTAL/100	

In my opinion, the grade I deserve for this semester is (Circle the grade):

A+	A^O	A-	B+	B^O	B-	C+	C^O	C-	D	F
97-100	94-96	90-93	87-89	84-86	80-83	77-79	74-76	70-73	60-69	0-59

.

As we can see in Figure 1, students are asked to consider every aspect of learning that occurred during the semester. They are also asked to think in terms of absolute assessment. Students are now taking on the teacher's function of making a report of growth over the semester, based on the achievements and learning events. This has added meaning, since it is the people concerned who are producing the reports, explaining themselves to their peers, and agreeing on a suitable grade. This negotiation session has therefore taken the issue of student responsibility to its logical conclusion, by involving the learners in the grade-decision process. It is up to the teacher, as final arbiter, whether to accept these suggested grades, but if students have successfully acquired assessment skills and social skills during the semester, then the results of such a session should be well thought-out, well argued and honest. If the negotiation session is carried out as a learning conversation, with the teacher present, then there will be opportunity to discuss the grades with the students, and to ask them to support their opinions, though this one-on-one approach also has potential to be threatening.

The issue here is one of trust. If there are problems with students not taking the session seriously, then further training in assessment skills and collaboration skills will be needed before they can be involved in another negotiation session. If there has been sufficient preparation, however (notification on the assessment schedule for the semester, information about the session on the course notice board, etc.), students are usually eager to take the opportunity to discuss this final aspect of assessment seriously and honestly. We (teachers) must remember that it is our duty to allow positive outcomes in class, in addition to being prepared for negative ones. If students are ready for a negotiation session, it can represent the culmination of a semester of self/peer-assessment, and can finish off the assessment process by allowing students to reflect deeply on their

achievements and future goals.

6. Collaborative test-making

The purpose of this paper is to show that formative process-based assessment can be used effectively and efficiently in EFL classes, and that it can improve the learning in the class through heightened awareness of the learning needs on the part of the teacher and the students. Because it is process-based, student-centered, and formative, such assessment i) is ongoing; ii) is integrated with instruction; iii) is meaningful to the students; and iv) feeds information back into the course, thus modifying the teaching and learning that occurs. It follows from these characteristics that assessment is happening all the time. The teacher observes classroom events and takes appropriate actions; the students perform self/peer-assessments and make new learning goals based upon their reflections.

In this situation, portfolios, learner journals, interviews, observations and self-assessments will take up most of the students' and teachers' time during the semester. However, there are times when more formal assessment is required. Some examples are placement tests, proficiency tests, and end-of-semester tests. These are usually required by the educational institution, and teachers have no option but to prepare and administer the tests. It is at this stage that teachers face the dilemma of whether to give a CSAT[1]-style test in order to prepare students for the university entrance test, or whether to test performance and growth over the semester.

If we continue using the principles of autonomy and 'learning by doing' that have illuminated the paper thus far, we can see that it makes sense, even at this stage, to involve students in their own test-making. Whereas test-preparation is a short-term learning strategy, the ability to write tests calls upon all the higher-order thinking skills required by the 7th National Curriculum, and empowers students with assessment skills for their entire lives. Students who experience test-making are practicing a particular type of self-assessment, and thus developing cognitive and linguistic skills. By making a review test based on a chapter of the school textbook (for example), students apply critical thinking skills to the content of the chapter. When deciding what should be tested, and how it should be tested, they are understanding the learning content and working with it. Just as the best way to learn something is to teach it to someone else (preferably in their Zones of Proximal Development), so it can be said that comprehension, problem-solving and critical thinking are enhanced by writing a test (Shin, 2002). It can therefore be very interesting to ask students at any level and any age to write the end-of-term test, or simply the regular review tests for each chapter of the school textbook (Finch, 2005).

This approach gives great freedom to teachers and students. It not only assesses learning, but also critical analysis, in that students have to examine the textbook and decide upon appropriate test topics and questions. If we extend the format to include open-book questions, the opportunity for assessing understanding and higher-order thinking skills are increased. An open-book test

---

[1] CSAT = College Scholastic Aptitude Test

minimizes the attention to memory that dominates most formal tests. By allowing students to access the required information in their textbooks, the test questions can focus on higher-order thinking skills, and can ask students to perform tasks with the information, rather than just repeating it. Information is accessible in many media in the 21st century, and to reward students only for their memory skills is to ignore an important fact of life – the need to know <u>how</u> to find and use information.

In addition, a group-test format can be employed. Although most tests are taken individually and competitively, we must remember that collaboration is a more effective learning tool than competition, and that teamwork is highly valued by future employers. It is reasonable, therefore to 'practice what we preach' by testing students in groups. There will be cases in which some students rely on other group members excessively, but there will be more instances in which students who have grasped the value of working together will produce excellent results. Some groups will work through the questions together, producing a consensus result, while others will allocate different questions to different group members, who will report back to the group when they have finished. In this sense, the test is not only evaluating course content, but is also promoting collaboration and social skills and offering a further opportunity for development. There are always ways of obtaining individual scores, but very few means of finding out how effective students are at working together.

Additional advantages are that the test need have no time limit, and there can be one test-booklet for each group. This format encourages real group-reporting skills, and simulates real-life situations more accurately than normal tests. Students are free (and have the responsibility) to turn in the best effort they can manage together. Lastly, students can be encouraged to discuss test-taking strategies and to make a plan for the test. In other words, a real-life situation (the planning meeting) is being simulated. Videos of students making and taking group-tests can be viewed at: www.finchpark.com/videos/alt_assess/.

7. Multiple-choice test-items

If teachers feel unsure in terms of asking their students to write regular review tests, or the end-of-semester test, they will have to make them by themselves. In this case, multiple-choice items are usually used, since they are easy to administer, easy to mark, and suitable for "number-crunching" (using quantifiable results "objective" of tests to select students for a limited number of school/scholarship/job positions). When using multiple-choice tests, however, we must constantly ask ourselves whether we are formatively evaluating language performance, or whether we are testing the students' ability to take multiple-choice tests.

On the positive side, multiple-choice tests can be used to measure simple linguistic aspects (grammar, vocabulary, structure), reading comprehension, and even (when used appropriately) some higher-order thinking skills. There are also, however, a number of dangers and pitfalls, of which we need to be aware:

1. grammatical inaccuracies;
2. incorrect answers to multiple-choice questions and reading passages;
3. more than one appropriate response to multiple-choice questions;
4. poorly-constructed passages (not conforming to the rules of paragraph construction);
5. illogical passages (poor expositions of unfounded, opinion-based arguments);
6. inappropriate use of the target language; and
7. inappropriate content (racist, sexist, discriminatory) (Finch, 2004, p. 310)

In addition to these problems, we also have the issues of test-item validity and reliability. If we make the answer options too close in meaning, then students can claim that the test is invalid. However, if we make the options too unreasonable, then the correct answer will be chosen by everybody, and certain options will become redundant. For this reason, multiple choice tests can only function well when piloted responses to the test-items undergo thorough analysis.

When making test-items, therefore, we need to decide what we wish to assess, and then make simple, clear, unambiguous (piloted) test-items. Rather than evaluating low-order knowledge, we should try to develop the skills necessary for reading in real life. This can be done by making test-items based on: i) recalling facts; ii) understanding facts; iii) recognizing words in context; iv) distinguishing fact from opinion; v) keeping events in order; vi) making correct inferences; and vii) understanding main ideas (Spargo, 1998).

Figure 3 offers an example of how the traditional multiple-choice format might be changed to avoid the problems outlined above and to encourage higher-order thinking skills. This test item requires students to recognize words in context (inferring), and refers to a reading passage.

FIGURE 3: Recognizing words in context (Adapted from Spargo, 1998)

| **Question 1: Instructions** |
| Find the word <u>vital</u> in the passage. One definition below means the same or almost the same thing. It is a <u>synonym</u> for <u>vital</u>. One definition has the opposite or nearly opposite meaning. It is an <u>antonym</u>. The other has a completely different meaning. Circle the correct options: S (synonym), A (antonym), or D (different). |
| ① unimportant   S   A   D
 ② lively   S   A   D
 ③ crucial   S   A   D |

More cognition is going on now, as students comprehend the instructions and then answer this question. The use of the meta-language terms <u>synonym</u> and <u>antonym</u> is no real problem. They represent useful new vocabulary for the students, and they help to raise awareness of these concepts. The main point is that students are choosing the correct definition for <u>vital</u>, dependent upon context. If there is a problem, then it is that the synonym <u>crucial</u> is more difficult to understand than <u>vital</u>. When testing vocabulary, it is important (even vital!) that the options are at the appropriate level

and are easier to understand than the word they are being compared with.

As we can see, teachers can make such an item with ease, and need not worry about inappropriate answers. Instead of choosing one 'correct' answer, students have to find responses for each option, and these responses must be the result of cognitive analysis of the reading passage. There are no redundant options, and the answers cannot simply be deduced from the test-item itself. This approach can be applied to i) recalling facts; ii) understanding facts; iii) recognizing words in context; iv) distinguishing fact from opinion; v) keeping events in order; vi) making correct inferences; and vii) understanding main ideas (Spargo, 1998).

8. Conclusion

In view of the high-stakes, test-driven nature of secondary education in Korea, it can be difficult and even depressing for teachers who wish to take a student-centered, communicative, collaborative approach, in order to promote meaningful learning and holistic, intrapersonal/interpersonal growth. Any divergence from the textbook, the test-determined syllabus, or the teacher-fronted grammar-translation teaching method, can incur immediate protests from students, parents and principals, all of whom are ruled by the all-powerful testing culture. However, it has been shown that that "working together to achieve a common goal produces higher achievement and greater productivity than does working alone [and] is so well confirmed by so much research that it stands as one of the strongest principles of social and organizational psychology" (Johnson & Johnson, 1991, p. 40). Therefore, if we can encourage students to study together and to see each other as collaborators rather than as competitors, it is quite likely that they would all have improved scores on standardized tests.

Asking high school students to make review tests based on the textbook has the potential to promote critical thinking, collaboration, comprehension and problem solving as part of the test-driven scenario, rather than in opposition to it (Finch, 2005; McClean, 1995). However, we need to consider the implications carefully. If it can be shown that any educational innovation is an efficient way of improving their scores on the university entrance test, then it would not be long before every teacher would be required by the powers-that-be to carry out this practice. Test scores in general would rise, and the cut-off point for acceptance to the best universities would also rise. In other words, nothing would change; the same propositional (educationally unsound) tests would continue to determine the lives of every individual taking the test countrywide, and the English language lesson would continue be dominated by short-term, extrinsic motivation.

When considering alternatives to this situation, educators must remember the basic, holistic purpose of education, before asking the question of what leads to excellence. The 'winners' in the present situation are those with the best memories, and this is of little use for a society that needs creative, inspired team-work: "it is quite clear that it [competition] has no adaptive value whatever in the modern world" (Montagu, 1973, p. 72.). If the goal of education is to produce "A person who seeks to develop his/her own individuality on the basis of well-rounded and wholesome

development" (Korea Institute of Curriculum and Evaluation, 2007), then teachers need to construct their learning environments in a way that will produce such a result. This will involve replacing the present test-driven approach with public accountability, continuous assessment, portfolios, language projects, and encouragement of intrinsic learning. In view of the general consensus that this is highly desirable but impractical, teachers need to ask themselves whose interest they are serving in the present situation:

> Are we creative agents of social change or are we dispensers of band-Aids to the injured and facilitators of adjustment to 'the way things are'? (Elleson, 1983, p. 198)

The administrative pressure for teachers and professors to 'come-up with the numbers' and to rank language achievement 'objectively' has resulted in summative, memory-based testing, with its avoidance of personal, social and affective development. However, such testing, which concentrates on the "target-like appearance of forms" (Larsen-Freeman, 1997, p. 155) ignores the fact that "we have no mechanism for deciding which of the phenomena described or reported to be carried out by the learner are in fact those that lead to language acquisition" (Seliger, 1984, p. 37), as well as the fact that the learner's interlanguage is not a steady commodity and often deteriorates prior to internalizing new content. Even if we could ignore such considerations, we would still have to recognize that the competitive testing culture has been shown to be unsuccessful according to its own criteria. Thus, many people who excel in high-stakes tests such as the Test of English for International Communication (TOEIC) are unable to communicate effectively and fluently in the target language. Knowledge-based testing fails to assess whether the learning content has been internalized and can be applied (or is applicable) in non-rehearsed situations.

It is vitally important, therefore, that educators seriously address the understanding crisis and reinvestigate the purpose of education and assessment, the role of the teacher, and the expectations of society. In an era in which unlimited information is publicly available on the Internet, and in which rapid advances in technology make the conventional teacher's learning irrelevant, we must look at the assumptions that underlie our educational practices and ask why it is that they continue to use pedagogically unsound high-stakes testing.

This paper has shown that research findings concerning the desirability of a student-centered, autonomous, collaborative approach to learning can be applied to assessment within the classroom, with resulting promotion of cognitive, affective, social and cultural growth. In terms of standardized testing, it is necessary to identify national priorities. If cost is more important than producing informed, competent, creative citizens, then multiple-choice, high-stakes testing will continue. As the saying goes, 'You get what you pay for.' If educational assessment is judged by its cost-effectiveness, then the results will match that criteria. However, if assessment is seen as vital for the growth of the nation, its citizens, and the economy, then standardized assessment must 'move with the times' and must be made accountable to current trends in society. The current individualized, competitive test format is irrelevant in terms of preparation for life, and memory-based trivia learned for such tests are duly forgotten by test-takers.

We now have the means to take a different path. Countries such as Finland and Denmark have adopted a process-based approach to assessment, and consequently have transformed their economies by injecting creative, competent, skilled citizens into their societies. If Korea is to survive as a global power in the coming century, it must examine these options and radically transform its assessment policies.

In conclusion, Gardner proposes a single criterion for effective education: "*An education that yields greater understanding in students*[2] (Gardner, 1993, p. 145). He also (1999) asked those who support high-stakes testing:

> … to think of what it means to be educated in a discipline where one learns to think like a scientist, a mathematician, an artist or an historian. He asked them to think of what it meant in the disciplines to pursue the meaning of truth and its equally important opposite, what is false or what is indeterminate. He asked them to think of what it means to understand beauty, and its equally important opposites, ugliness or kitsch. He challenged them to think of what it means to deal with what is good and what is evil in this world. After we engage in Gardner's exercise, we must ask ourselves if the assessment tools used for high-stakes testing are designed to measure these things, or are they likely to miss them completely? (Nichols & Berliner, 2006, pp. 169-170)

References

신동일. (2002). 영어시험 사용자 관점에 관한 새로운 고찰: 반실증주의적 패러다임논의로부터.
　　응용언어학. 18(1), 293-312.

신동일. (2003). *한국의 영어평가학1: 시험개발편*. 한국문화사.

Allwright, R. L. (1984). Why don't learners learn what teachers teach?: The interaction hypothesis. In D. M. Singleton, & D. Little (Eds.). *Language learning in formal and informal contexts*. (pp. 3-18). Dublin: Irish Association for Applied Linguistics.

Ayers, W. (1993). *To teach: The journey of a teacher*. New York: Teachers College Press.

Blanche, P. (1988). Self-assessment of foreign language skills: implications for teachers and researchers. *RELC Journal, 19*/1, 75-93.

Blue, G. M. (1988). Self-assessment: the limits of learner independence. In A. Brookes & P. Grundy (Eds.). *Individualization and Autonomy in Language Learning. ELT Documents, 131*. (pp. 100-118). London: Modern English Publications and the British Council,

Breen, M.P. (1987). Contemporary paradigms in syllabus design, part I. *Language Teaching, 20*/2, 81-91.

Darling-Hommond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass.

Eason, G. (2005). Testing link to school disruption. *BBC News*. Retrieved June 3 2007, from: http://news.bbc.co.uk/1/hi/education/4384399.stm.

---

[2] Italics in original text.

Elleson, V. J. (1983). Competition: A cultural imperative. *The personal and guidance Journal,* December 1983, 195-198.

Finch, A. E. (2004). An analysis of EBS preparation materials for the CSAT, English section. *English Teaching. 59/3,* 305 – 329

Finch, A. E. (2005). Students as test-makers. *Modern English Education. 6*(2), 3-25.

Finch, A. e. (2006). The Postmodern Language Teacher. *Studies in British and American language and literature, 78,* 221-248.

Finch, A. E. & Shin, D-I, (2005). *Integrating teaching and assessment in the EFL classroom: A practical guide for teachers in Korea.* Seoul: Sahoipyungnon Publishing Co., Inc.

Furger, R. (2002). Assessment for understanding. *Edutopia.* Retrieved July 2 2007 from http://www.edutopia.org/node/937.

Galyean, B. (1977) Language teaching as a means to self knowledge, self acceptance, and communication with others. In *CATESOL Occasional Papers*, *No. 3.* Winter, 1976-77.

Gardner, H. (1993). *The unschooled mind: How children think and how schools should teach.* London: Fontana Press.

Gardner, H. (1999) *The disciplined mind: Beyond facts and standardized tests, the K-12 education that every child deserves*. New York: Simon and Schuster.

Gardner, H. (2000). The testing obsession: In the quest to improve public schools, we've made test performance more important than education. *Los Angeles Times.* Retrieved June 3 2007, from: http://csmp.ucop.edu/cmp/comet/2001/02_12_2001.html.

Graddol, D. (2007). *English next: The future of English*. UK: The British Council

Guba, E. G. & Lincoln, Y. S. (1989). *Fourth generation evaluation.* California: Newbury Park.

Harri-Augstein S., & Thomas, L. (1991). *Learning conversations: The self-organised learning way to personal and organisational growth*. London: Routledge.

Johnson, D. W. & Johnson, R. T. (1991). *Learning together and alone: Cooperative, competitive, and individualistic learning.* Englewood Cliffs, N.J.: Prentice-Hall.

Kohn, A. (1992) *No contest: The case against competition.* New York: Houghton Mifflin Company.

Kohonen, V. (1999). Authentic assessment in affective foreign language education. In J. Arnold (Ed.). *Affect in language learning.* (pp. 279-294). Cambridge: CUP.

Korea Institute of Curriculum and Evaluation, (2007). National Curriculum: The direction of curriculum design. Retrieved July 2 2007 from: http://www.kice.re.kr/kice/eng/info/info_2.jsp

Legutke, M & Thomas. H. (1991) *Process and experience in the language classroom.* Harlow: Longman

Leithwood, K., Steinbach, R., Jantzi, D. (2002). School leadership and teachers' motivation to implement accountability policies. *Educational Administration Quarterly, 38*, 1, pp. 94-119.

Long, M. H. (1984). Process and product in ESL program evaluation. *TESOL Quarterly*, *18*/3, 409-425.

Montagu, A. (1973). *Darwin, competition and cooperation.* Westport, Conn.: Geenwood Press.

Nichols, S. L. & Berliner, D. C. (2005). The inevitable corruption of indicators and educators through high-stakes testing. Education Policy Studies Laboratory. Retrieved July 2 2007 from http://edpolicylab.org/.

Oscarsson, M. (1997). Self-assessment of foreign and second language proficiency. In *Encyclopedia of Language and Education. Vol 7: Language testing and assessment*. (pp. 175-187). Dordrecht: Kluwer Academic,

Parlett, M. (1981). Illuminative evaluation. In P. Reason, & J. Rowan. *Human enquiry*. (pp. 219-226). Chichester: Wiley.

Popham, W. J. (2005). F for assessment. *Edutopia, April 2005*. Retrieved July 2 2007 from: http://www.edutopia.org/f-for-assessment.

Richards, J. C., & Rodgers, T. S. (1986). *Approaches and methods in language teaching*. Cambridge: CUP.

Rogers, C. R. (1975). *Freedom to learn*. Colombus, OH: Merrill Publishing Co.

Sano, M., Takahashi, M. & Yoneyama, A. (1984). Communicative language teaching and local needs. *English Language Teaching Journal*, *38*/3, 170-177.

Seliger, H. (1984). Processing universals in second language acquisition. In F. Eckman, L. Bell, & D. Nelson (Eds.). *Universals of second language acquisition*. Rowley, MA.: Newbury House.

Spargo, E. (1998). *Timed readings plus: Book 10*. Lincolnwood, Illinois: NTC/Contemporary Publishing Group, Inc.

Stern, H.H. (1983). *Fundamental concepts of language teaching,* Oxford: Oxford University Press.

Stevick, E. W. (1976). *Memory, meaning and method*. Rowley, M.A.: Newbury House.

Tudor, I. (1996). *Learner-centredness as language education*. CUP: Cambridge.

Underhill, N. (1989). Process in humanistic education. *English Language Teaching Journal,* 43/4, 250-260.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, Mass.: MIT Press.

Weir, C. J. (1998). *Communicative language testing*. Exeter: University of Exeter Press.

Williams, M. & Burden, R.L. (1997). *Psychology for language teachers: A social constructivist approach*. Cambridge: CUP.