

Group testing – Student-Designed Tests

Dr. Andrew Finch

Kyungpook National University

Abstract

In this study, which took place in a university in South Korea, two classes of students (Sophomores and Juniors) designed review tests for each unit of their course textbooks. It was hypothesized that involving students in every stage of the evaluation process would encourage greater comprehension of the concepts in the textbook, in addition to helping students to set goals and reflect upon achievement. Results confirmed these expectations, with students typically gaining a great deal from the group decision-making format, allocation of responsibility, and promotion of intra-personal and inter-personal responsibility. This presentation described the process of writing, implementing and marking a student-designed test, and showed a video of the group test-making process, taken during the semester. This video can be viewed online at www.finchpark.com/videos/alt_assess.

Introduction

Recent years have seen an increasing amount of attention to learner autonomy, cooperative learning, management of affect and promotion of learner responsibility in second-language learning literature (Arnold, 1999; Benson & Voller, 1997; Cotterall, 2000; Gremmo, 1995, p. 151). Contemporary language-testing theories and practices have also acknowledged the need to broaden the scope of language assessment to include cognitive, affective and social learning. Examples of this trend are authentic assessment (Kohonen, 1999) and classroom-based assessment (self/peer-assessment, portfolios, learner-journals, language conversations, etc.) (Richards, 1999). Such advances and the compelling pedagogical justifications behind them have, however, passed largely unnoticed by the proponents of high-stakes testing. This is especially evident in Korea, where the summative, multiple-choice, university entrance test continues to turn every 3rd grade high school English lesson into a test-preparation class, and where test-related student suicides are an annual event (Soh, 2003). The fact that this situation is not confined to Korea (Herbert [1983, p. 30] cites "competition

pressures" as prominent factors in the alarming rise in suicide rates among young people) does not make it any more acceptable, but points to a malaise prevalent in education systems around the world.

In view of this situation, along with the wealth of literature supporting self/peer-assessment as a valid educational goal (Achara, 1980; Blanche, 1988; Blue, 1988; Cram, 1997, etc.), the purpose of this study was to examine whether English teachers might develop enhanced confidence, motivation and attitudes to learning in their students, by involving them in the entire process of test-design. It was also hypothesized that the cooperative learning context in which this test-design would occur would promote intra-personal responsibility, as well as inter-personal responsibility, along with long term autonomy and awareness of learning strategies. If this were shown to be the case, then test-preparation might become an indirect, non-threatening, means of developing meaningful life skills in students.

Any test that is used in any English class must have a definite purpose, since tests are intrinsically threatening, and can cause undesirable levels of anxiety and worry, which can harm the learning process. Traditional, low-order-thinking (memory-based), multiple-choice, 'trivia'-based tests have the inevitable outcome that students who are relatively unsuccessful in recalling rote-learned lexis items, who have learning styles incompatible with high-stakes tests, and who have worked out some time ago that they were not going to get a good score, simply give up trying to learn: "Those who believe they will lose may see little point in trying hard" (Kohn, 1992, p. 56). High levels of stress and internalization of failure also result from this evaluation system, since "When emphasis is placed not upon the demonstration of competence, but upon winning, then the competitor eventually 'comes to believe that he is defective and deserves to fail'" (Walker, 1980, p. 250). This process "of coming to equate losing with being a loser ... is a gradual process, but it is taking place at any given instant in our playing fields and classrooms, our offices, and even our homes" (Kohn, 1992, p. 109).

When we consider, however, that the prime educational goal of assessment is to inform the learners of their achievements, and to help them in further (realistic) goal-setting, then it can be seen that testing need not be so stressful, and can occur informally, in many forms. In fact, language performance can be satisfactorily evaluated in a non-threatening, student-centered manner.

There are times when more formal assessment is required, of course. 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. This presentation described a study carried out in a university in

Korea, in which students were asked to make the final test for their courses “English Teaching Methodology” and “Classroom-Based Assessment.” Students discussed how to make a test, how to choose test formats, and how to use the test to assess growth in a non-threatening manner. For the purposes of showing the process and the product, the test based on *How To Teach English* (Harmer, 1998) is used here, and rather than considering the various issues in isolation, matters relating to the What? Why? and How? of test-design are addressed as they appear in the examples.

Test-design

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 course textbook, 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 content and working with it. Just as the best way to learn something is to teach it to someone else (within their ZPD), so it can be said that comprehension, problem-solving and critical thinking are enhanced by writing a test. 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. When doing this, it is important to give them a framework within which to work. This might be a template of the school test, or some examples of test types.

There is another side to this issue, which is that high-stakes tests are intrinsically competitive, and that team-work and the ability to work with others is never evaluated by them, despite the fact that it is an important factor for employers when hiring new staff. 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 (even in the testing situation), it is quite likely that they will achieve improved individual test scores in the more traditional testing environments.

The test

As mentioned above, it is important to give students a framework within which to work, however free the format might be. Such a framework involves organisation and

accurate, comprehensive test-design signposting. The first point to note in this test (Figure 1) is that it is well organized (face validity), and that all the demographic information, including space for the final marks (out of 100), appears on the cover page.

FIGURE 1: *How To Teach English* test, cover page

My name:	Date:	
Group members:		
<p>Open-book, group test of the course textbook: How to Teach English</p>		

- There is only one test-book for each group.
- Complete the test-book cooperatively.
- Everyone in the group will receive the same mark for this test.
- Pay attention to the instructions for each question.
- If there are no instructions for a question, decide (in your group) what to do.

When you do the grading, write the score for your section here:					
Part 1	/20	Part 2	/20	CH 2:	/4
CH 5	/24	CH 6	/12	CH 10	/10
CH 13	/10	Total:		/100	

Secondly, we can see that it is a group test. While most required tests are individually-based, we must remember that collaboration is a more effective learning tool than competition (see above). It makes sense, therefore to “practice what we

preach” by testing students in groups. Students taking this test have thus experienced all the factors involved in collaborative test-design (brainstorming of goals and procedures, allocation of responsibility, accountability for assigned tasks, etc.) and are now experiencing collaborative test-taking. 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 (collaboration and social skills), but is promoting those 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. It should also go without saying that this test would not be the only measure by which student performance would be judged. (In this case, the designing, administering, taking and marking of the test was worth 20% of the grades for the semester.) A video of this group-test can be viewed at: www.finchpark.com/videos/alt_assess/index.htm.

Thirdly, Figure 1 is an open-book test, and thus 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 – that we need to know how to find and use information.

Fourthly, the test has no time limit, and there is only 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 (in terms of the cover page, Figure 1), students are encouraged to discuss test-taking strategies and to make a plan for this test. In other words, a real-life situation is being simulated – the planning meeting.

Although open-book tests in general promote higher-order thinking skills, Part 1 of the teacher-designed test (Figure 2) tests factual knowledge. This might seem a contradiction, but it is also a learning opportunity. Groups of students who have not acquired these basic terms while reading the book will need to find them in order to complete the page, whereas those who know these terms will be able to fill them in and thereby save a great deal of time.

FIGURE 2: *How To Teach English* test: Part 1

• **Part 1: Simple information**

1. (Chapter 1)	What is TTT?	(1 point)
2. (Chapter 1)	What is STT?	(1 point)
3. (Chapter 2)	What is instrumental motivation?	(1 point)
4. (Chapter 2)	What is integrative motivation?	(1 point)
5. (Chapter 2)	Fill in the names for student levels: beginners 	(1 point)
6. (Chapter 3)	What is proximity?	(1 point)
7. (Chapter 3)	What is appropriacy?	(1 point)
8. (Chapter 4)	What does ESA mean?	(1 point)
9. (Chapter 4)	What does PPP mean?	(1 point)
10. (Chapter 4)	What does CLT mean?	(1 point)
11. (Chapter 4)	What is a discovery activity?	(1 point)
12. (Chapter 5)	What is a Complement?	(1 point)
13. (Chapter 5)	What is a Particular preposition?	(1 point)
14. (Chapter 5)	What is a Determiner?	(1 point)
15. (Chapter 5)	What is a Paralinguistic feature?	(1 point)

Figure 3 shows an activity from one of the student-designed tests. This is one of the activities designed by a group working on Chapter 2 of *How To Teach English*. As we

can see, this is a matching exercise, though it is not a simple one-on-one exercise, since there are fewer boxes in the left column. There can be a problem in marking matching exercises, since one mistake causes at least one other choice to be wrong. Test-designers therefore need to be aware of this, and should be careful in the allocation of marks.

FIGURE 3: Student-designed test of *How To Teach English: Chapter 2*

31. Some learners seem to be more engaged with a progress of learning and take advantage of what's going on more than others. They possess some common characteristics. Match the items on the right to the items on the left (using lines). (4 points)

A willingness to listen	They prepared to take risks, to try things out and see how it works. Of course, not all successful language learners are extroverts, but the urge to use the language is important one.
A willingness to think about how to learn	They bring or invent their own study skills when they study on their own or come to class, for example, the best way to read a text.
A willingness to ask questions	They are keen to get feedback from the teacher and act upon what they are told. (Teachers can offer constructive criticism rather than castigating them for being wrong.)
A willingness to experiment	They listen to What's going on - not just in the sense of paying affection, but also in terms of really listening with eagerness and intelligence.
	They ask teachers difficult and/or irrelevant questions judging whether it is appropriate to do so or it isn't. (Teachers should create an atmosphere which shows them that their questions are welcome.)

It is also worth noting that there are some grammar mistakes in the top right-hand box: “They [are] prepared to take risks, to try things out and see how it [they] works [work].” This is mentioned here, not as an exercise in error correction, but in order to illustrate an important point, i.e. that test questions (especially in high-stakes test) must be: i) grammatically correct; and ii) semantically unambiguous. If the grammar is open to question, then test-takers can justifiably complain about the question and ask that it be cancelled. If the question words support more than one interpretation (semantically ambiguous), then test-takers can demand that their version of the answer be accepted. Writing test-items thus requires a great deal of rigor. It is easy to write “What-am-I-thinking?” questions, only to find that students have understood the question in a different way. Test questions must be re-read, piloted on sample test-takers, revised, and re-revised, until there is no possibility of misinterpretation.

Figure 4 shows a feedback page. This is very helpful for test-takers and test-designers. For test-takers, it helps them to organize their thoughts about their learning needs, based on the ease or difficulty they experienced while completing the test. If we remember that assessment is primarily intended to give information to students, then it follows that even this test must help students to formulate new learning goals. For the test-designer, this feedback sheet provides important information, not only regarding the achievements and abilities of the students, but also their perceptions of the test and the relative validity/usefulness of the test-items

FIGURE 4: *How To Teach English* test: Feedback

1. How did you feel about this “teamwork” test?	
2. Did the test tell you about your learning process?	
3. Did the test help your language skills?	
4. Did the test help your creative-thinking skills?	
5. Did the test help your critical thinking skills?	
6. Did the test raise your social consciousness?	

7. Did the test help you in any way?	
8. What was the purpose of the test?	
9. Did the test encourage collaboration?	
10. How could this test be improved?	

Results

There is little space here to perform detailed analysis of the results of this study. Instead, it must suffice to say that it has been proceeding for two years, that four groups of students have so far experienced this test-design format, and the results, as measured by the questionnaire in Figure 4, along with classroom-observation of the test-making and test-making process, have been positive. An advantage of the tertiary learning environment is that once the pilot course has been taken, the information about course procedure disseminates very quickly, so that subsequent students enter the course already conversant with the requirements and expectations. So it was that the second year's student-designed tests proceeded quite smoothly, with none of the testing-culture-shock of the first year. Students were quite happy with the fact that they were expected to design an open-book, higher-order-thinking test for their peers, and that they were going to mark the results of that test.

It is interesting that the four tests taken to date have shown little difference in terms of grade results. Each test was allocated 20% of the final grade, and of this, 10% was given to the test-design process, as observed and evaluated by the teacher. Only 10% was given to the actual results of the test, but it has been noticeable that most students have scored in the 90s (out of 100). One could put this high scoring rate down to peer-pressure and "friendly" marking, but the writer has observed that students do not want to score badly on tests written by their peers, and that they study the textbook intensively in order to avoid this. It has also been noticeable that the test-design process (also 10%) has been characterized by effective collaboration, group decision-making, and considerable out-of-class, self-directed 'homework.' Finally, students have commented on the questionnaire (Figure 4) that the test-design process raised their awareness of assessment methods and purposes.

Conclusion

In conclusion, this open-book, group test uses a number of test formats and asks a variety of questions of the test-takers. There are no language performance tasks, no listening activities, and no explicit speaking activities (the business-meeting format of this test, resulting from group discussion of goals and allocation of tasks is in itself a speaking/listening activity, and teachers can use this opportunity to observe students using the target language), but the test has done its job of getting students to review the material in the textbook. It has also carried out its role in a non-threatening manner, using collaboration rather than competition, and simulating a real-life, report-writing situation.

References

- Achara, W. (1980). Self-assessment in English skills by undergraduate and graduate students in Thai universities. In J. Read, (Ed.). *Directions in Language Testing*. (pp. 240-260) Singapore: Singapore University Press.
- Blanche, P. (1988). Self-assessment of foreign language skills: implications for teachers and researchers. *RELC Journal*, 19/1, 75-93.
- Arnold, J. (Ed.). (1999). *Affect in language learning*. Cambridge: Cambridge University Press.
- Benson, P. & Voller, P. (1997). *Autonomy and independence in language learning*. Harlow: Longman.
- 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*. London: Modern English Publications and the British Council, 100-118.
- Cotterall, S. (2000). Promoting learner autonomy through the curriculum: principles for designing language courses. *English Language Teaching Journal*, 54(2), 109-117.
- Cram, B. (1997). Training Learners for Self-Assessment. *TESOL in Context* 2/2, 30-33.
- Gremmo, M-J. (1995). Autonomy, self-direction and self-access in language teaching and learning: The history of an idea. *System* 23(2), 151-164.
- Harmer, J. (1998). *How to Teach English*. Harlow: Longman
- Herbert, H. (1983). *Suicide in America*. New York: Norton.
- 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, N.Y.: 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: Cambridge University Press.
- Richards, J. C. (1999). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.
- Soh, J. Y. (2003) College exam drives students to suicide. *The Korea Times*. Retrieved May 9, 2004, from <http://times.hankooki.com>
- Walker, S. H. (1980). *Winning: The psychology of competition*. New York: Norton.

Andrew Finch, Ph.D.

Kyungpook National University, Teachers' College, Department of English Education
1370 Sankyuk-dong, Buk-gu

Daegu City, 702-701, Republic of Korea.

School: (053) 950-5832

School Fax: (053) 950-6804

aef@mail.knu.ac.kr; aef@finchpark.com

Andrew Finch, whose Ph.D. (Manchester University 2000), described the setting up and evaluation of a task-based language program in Korea, is currently assistant professor of English Education at Kyungpook National University. Andrew's research interests centre around language-learning as education, and he has co-authored a number of task-based conversation and assessment books, which can be accessed at <http://www.finchpark.com/books>