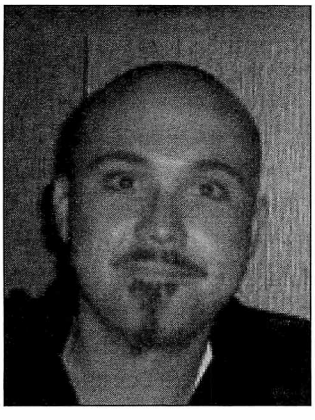
Part 2: Inquiry into language learners

Taking the teacher out of the test: Exploring student autonomy in EFL classroom testing



Marlen Elliot Harrison

Indiana University of Pennsylvania

scenteur7@yahoo.com

About Marlen

Marlen Harrison is currently pursuing his doctoral degree in English Composition and TESOL after four years of teaching in Japan. Co-coordinator of JALT's Learner Development Special Interest Group, Marlen's research interests include email exchange projects, project-based learning, and alternative approaches to testing and evaluation. He currently resides in Indiana, Pennsylvania.

マーレン・ハリソンは、日本で4年間教師を務めたのち、現在は英作文とTESOLの博士号取得を目指しています。JALTの学習者ディベロップメントグループのコーディネーターである彼が関心を持つ研究対象として、メール交換プロジェクト、プロジェクトに基づく学習、試験や評価への代替的アプローチなどがあります。現在、ペンシルバニア州インディアナに在住しています。

ABSTRACT

Is it possible to "take the teacher out of the test?" Wanting to give students the chance to create, administer, and evaluate their own tests, I have been examining creative methods for assessment as a means to help students synthesize their knowledge and decrease their test anxiety. I will explore ways in which I transformed classroom testing into opportunities for student-centered, collaborative learning.

「試験から教師を除く（組み入れない）」ことは可能なのだろうか？学生に自身の試験を作成、実施、評価してもらいたいという希望から、私は、学生に知識を統合させ、試験に対する不安を軽減させる手段となる独創的な評価法を検討している。本稿では、教室での試験を学生中心の共同学習的な機会へと転換させる方法を検討する。

In 1995, Tim Murphey wrote a paper titled "Tests: Learning through negotiated interaction" in which he explored collaborative testing as an approach to "putting students more interactively in the center of creating and administering...tests" (p. 12). Murphey's ideas have been an absolute delight to review, validating and further exploring many of my own testing practices as a language teacher. Having given a lot of thought to both why and how I test my students, I have concluded that my beliefs about testing have had two major influences, the first being my own experiences as a student, and the second being the current paradigms of learner development and autonomy. What follows then is a review of my attempts to remove myself from the testing process as much as possible, hoping to maximize opportunities for student autonomy, collaboration, and creativity.

LEARNING FROM EXPERIENCE

I have always suspected that exams were excellent tools for synthesis, if not necessarily evaluation. Some of my most exciting "wow" memories as a student are those last few moments of a specific test or exam where I suddenly saw the bigger picture, realizing just how much I had learned, and feeling a sense of accomplishment because of it. As a high school student, writing papers for my literature class, I often recognized that a successful essay was one in which I was able to "put it all together," working pieces of a puzzle into an ordered, intricate design. I distinctly remember remarking to a teacher that I felt almost selfish in completing his written exam. He had the task of reading and evaluating my ideas, while I had the pleasure of putting the ideas together for him. I believe I compared it to a show where I was the actor and he the critic whose review was moot because the performance itself had resonated so deeply for me.

I also have fond memories of a college study group where we made tests for ourselves and then shared them with each other as a way of helping us prepare for our psychology and biology exams. The process of reflecting on what we had learned, and formulating anticipated test questions really helped synthesize the information and was as useful in aiding understanding as taking the practice tests we had created. The ensuing discussions about our tests, and the realization that we all had differing points of view as to what was important, also provided further opportunities for learning.

Years later as a university teacher in Japan, while discussing graded readers with my reading classes (2nd year English Literature and Language majors), I decided that I wanted to motivate students to read on a schedule and concluded that grades might be the best motivator. After a few weeks of thinking up quizzes, I suddenly had an idea. Instead of supplying students with questions that I thought were important, why not allow them to create their own questions, reflecting on ideas that were important to them? Additionally, I had hoped that the test creation process would actually be a good review of the material, akin to a study session. After students

had worked together in small groups, using questions from both the classroom and their graded reader activities to create tests, they exchanged papers, answered the questions, and then returned the tests. And it didn't stop there! I then asked students to grade the tests and turned this process into a conversation task using a previously taught dialogue. Whew! I was amazed at how much activity could be derived from one simple task.

A BRIEF WORD ABOUT TESTING

Deciding just what exactly a test should measure and how to best go about doing so is quite a tricky task. The current evolution of assessment in language learning has been well documented in the past few years (Finch, 2004; Gorsuch, 2000; Shaaban, 2001; Smith, 2001) with the recognition that traditional pencil-and-paper summative methods of evaluation in the classroom do not necessarily reflect the experience of students whose studies are based on communicative activities. Moreover, instructors have begun to recognize the need to honor "uptake," a term coined by Allwright (1984, p.11) as "whatever it is that learners get from language lessons," as equally important as what the instructor thinks the student should learn. When I refer to testing and evaluation with regards to the activities prescribed within this paper, I am not referring to placement or aptitude testing, but to classroom evaluation and assessment.

DEVELOPING AN APPROACH TO STUDENT-CENTERED TESTING

I had not spent hours planning how I could increase student autonomy in the testing process, but rather, it just "came to me" while reminiscing about my own learning processes. I thought about the amount of work I was doing as the teacher and realized that my students had strong enough analytical and language skills to be doing the work themselves. In the past, I had asked students to evaluate each other's essays, proofread each other's writing, or work collaboratively to grade tests, but I had never before actually tried having them create their own tests. My mind was suddenly filled with questions: How could I be sure that the test items were valid? How could I be sure that the grades were representative of student knowledge? How could I be sure that this was an effective testing process?

As the students continued the test creation process, I noticed that some groups were quiet – deciding that each individual should create his or her own test questions and then combine these upon completion. Why were they approaching the task in this way? Simple: They lacked sufficient conversation skills that would allow them to communicate with each other and they lacked sufficient experience with test creation to give them the confidence to collaborate. This left little room for student interaction and had me wondering if asking them to complete the collaborative process strictly in English would serve to either hinder their reflection process or aid their communicative abilities by allowing focused and real attempts at constructing meaning.

I then made a change and allowed students the opportunity to communicate in Japanese during the creation of a subsequent test and noticed that although certain groups continued to have each member work independently, the amount of "checking in," or "negotiation" as Murphy might phrase it, increased. Students were more involved in asking each others' opinions about their items and checking for meaning. I had to make a decision: Would I focus the test creation process on conversation practice, or would it be better to allow students the opportunity to process content?

Additionally, I noticed that many of the students' questions mimicked those they had previously completed in their books' activity sections, while other questions were directly

copied from those previously reviewed in class. Only a small group of questions were original and attempted to employ yet other formats, but the one thing that kept troubling me was that all of the formats were written; sadly, the oral communication component was missing.

REFINING THE PROCESS

The following semester, I created a template (see MAYA! website for Appendix 7A) of question formats for my students and worked with them throughout one class period to help them better understand how to design their tests. This added support from the instructor made a marked difference as students both mimicked and manipulated the template resulting in the creation of rather individualized tests (see MAYA! website for Appendix 7A). Students began to add images and new question formats, and as evidenced by the atmosphere in the classroom, and in my own emotional appraisal of the experience, they truly seemed to enjoy the entire process. Having them work in groups, they could then forward the test to me for proofreading, grammar correction, and suggestions.

Earlier, I posed the following questions:

1) How could I be sure that the test items were valid?
2) How could I be sure that the grades were representative of student knowledge?
3) How could I be sure that this was an effective testing process?

I gave a lot of thought to the first research question and concluded that the validity of the test items and the actual test itself just might be secondary to the experience gained from preparing the tests. Having the students work together in groups chosen by the instructor allowed for a variety of student input based on individual experience in the classroom. Finally, students themselves overwhelmingly praised the activity and volunteered that they had felt much more relaxed, as one student noted on an end-of-semester questionnaire, "It's a test and doesn't feel like a test. I'm exciting to see their [classmates'] tests, and if it's same as my test. I'm not really thinking this is a test."

One by-product of this new student-centered testing approach that I had not expected was the decrease in performance anxiety. But why should students feel more relaxed? Another student offered, "I practiced a lot for making the test. I think I know it [the material on the test] now so I didn't really study like I always study. I can relax and have fun." Some students became quite competitive with their creations, aiming to top each other in presentation and content. I was pleased to see the range of ideas expressed in the test questions and the earnest pleasure students took in evaluating their classmates' responses.

The evaluation process of test responses was undertaken by the groups of students who created the respective instruments. Again, I was faced with allowing Japanese conversation to ease negotiation of meaning, vs. strict English use as a means of practice. To my surprise, although I allowed Japanese, a number of students incorporated pre-taught English classroom dialogs into their evaluation process, illustrating the effectiveness of providing students with the opportunity for authentic communication.

BACK TO THE DRAWING BOARD: COMMUNICATIVE TESTING?

As my experience with student-centered testing continued, I realized that what I had really wanted was to make a three-part testing process: a) reinforcement of learned material via the test-creation process, b) application and recall via taking the test, and c) examination of pragmatics via peer assessment and negotiation of grades. To summarize Kitao and Kitao (1995):

> Communicative language tests are those which make an effort to test language in a way that reflects the way that language is used in real communication. It is, of course, not always possible to make language tests communicative, but it may often be possible to give them communicative elements. This can have beneficial backwash effects. If students are encouraged to study for more communicative tasks, this can only have a positive effect on their language learning.

With this in mind, I decided that I needed to re-think the way in which I asked students to undertake evaluation. I realized that the oral aspect of communication was being neglected in the actual tests and wondered how I could change that.

After some brainstorming, I thought about classroom activities that required oral communication, a balanced amount of communication within groups, and opportunities for realistic conversation, Moreover, I wanted the activity to be easy to evaluate and fun! My solution? A game!

I showed students examples of two board games, one a photocopied template from the internet, the other a game created by one of my previous students. Again, working in groups chosen by the instructor (to ensure even distributions of language skills), students were asked to create a game that would meet the above criteria. They were given class time to brainstorm a proposal which then had to be presented to and accepted by the instructor. They were asked to complete the creation of the game as out-of-class work with clearly written English instructions were required.

A variety of ideas were presented, everything from quiz-style games, to role-play, and even a trivia game. Assessment was completed based on the number of questions a student answered, and if the student did not know the answer, they were instructed to respond in English accordingly. I experimented with two different formats for assessment: In one class, students were asked to assess their own performance and confidentially, their classmates' performances based on the following Likert scale criteria:

1) Participation – "I spoke English only," and "I tried my best with a positive attitude."
2) Content – "I was able to answer the questions using English," and "I could respond in a timely manner."
3) Performance – "My spoken English (accent) was easy to understand."

The second format was a simple pass/fail format where students had to earn a minimum number of points – points were earned for both answering questions and participating in the conversations of others using English. Each group also had a team point keeper. With both formats, essentially any student who participated could pass the exam. Murphey (1995) reflects on his own testing process and writes, "The main purpose of these [types of] tests is not evaluation but rather stimulation of effective language learning processes that can later be used by learners to help them learn whatever they want" (p.13).

WRAPPING IT UP

Murphey (2003) notes that "...during testing (or gathering data) students are still constructing knowledge, using the tests and evaluations as learning events" (p.4). To that effect, by having had my students play a major role in their testing, I hoped to accomplish the following:

1) Allow students to more actively participate in all aspects of the testing process.
2) Use the typical "study process" for tests as an interactive task that can be undertaken during class time.
3) Offer opportunities for collaborative learning.

4) Support communicative competence via conversation tasks related to the tests.
5) De-emphasize what I as the teacher feel is important, and ask the students to examine what they found important.
6) Increase student participation and autonomy in the classroom and raise motivation to "study."

Baron (2005) writes, "If students are to take more responsibility for their own learning, they should be involved in the assessment process, not only in seeing the results but also in designing some of it" (p.1). Students may play a larger role in the assessment process and have greater opportunity to practice effective communication in English by creating their own test questions, grading the tests and discussing the tests. In this paper I have highlighted a variety of testing methods that accentuate student autonomy and collaboration, while minimizing teacher participation using actual examples from my courses and feedback from students, but truly, the possibilities for creative testing are endless.

CRITICAL READER RESPONSE 1
MIKE NIX

"It's not about finding the answers; it's about asking better questions."

This maxim for intellectual inquiry, which I recall from my own days as a university student, seems to nicely articulate the purpose of Marlen's thoughtful and thought-provoking explorations of student-created tests. So what are some issues his paper highlighted for me?

Firstly, Marlen's paper made me think about how previous learning experiences can be both a resource for, and a limit on, the development of learner and teacher autonomy. Making practice tests in his study group at university provides a model for Marlen's initial break as a teacher with conventional teacher-devised tests. And "fond memories" of the value of those practice tests seem to give affective support for his experiments when worries about the validity and effectiveness of student-created tests emerged. But when his students first made their own tests, they were limited by the kinds of question types encountered during his course and, it seems likely, by the non-communicative nature of the conventional English tests they experience so often in Japan.

Describing how he helped students to be more creative, communicative and collaborative in their test-making, Marlen's paper asks what it takes for learners and teachers to both build on, and break with, their learning experiences to develop their autonomy. Or, put another way, having observed that it is actually not enough to simply 'take himself out of the test,' he explores what a teacher needs to do to scaffold students' growing control over their own learning and use of English. Marlen's account here suggests that teacher direction – templates of alternative test formats and more explicit explanation – is initially useful for the development of autonomy, with students' own noticing of various alternatives in the work of their peers becoming gradually more important. His explorations also suggest that for a teacher's autonomy to grow beyond the limits of their own formative learning experiences, a consciously exploratory and reflexive practice of scaffolding autonomy, and the repeated re-thinking or theorizing of previous learning experiences and current teaching practices, are essential.

The shift to test-making and taking as a communicative, collaborative learning process also raises the question of what learners should be evaluating: the development of language competence? control over learning processes? participation and use of English? Marlen wants students to reflect in test-making on ideas "important to them" but doesn't really address how the learners' own goals, as well as their "uptake" from classes, can be integrated into the testing/evaluation stage of ongoing, autonomy-inviting cycles of learning.

And, if the aim is to make evaluation an integral, formative part of the language learning/using process, not an end-point assessment, why stick with tests, even student-created ones? Why not use, say, learner journals, or portfolios? Perhaps the testing 'wow' factor has, in the end, become a constraint on Marlen's own autonomy to choose appropriate forms of evaluation. Or, on second thought, is it actually more autonomy-inviting for students to explore, de-construct, and subvert the very conventions of test-based evaluation that constrain so much of the rest of their language learning experience?

CRITICAL READER RESPONSE 2

ETSUKO SHIMO

In one of my university lecture courses, the professor asked us students to make questions. The professor told us if our questions were chosen for the final exam, we would get bonus points. In the test-making process, we reviewed, reused, and recycled previously learned knowledge. In Marlen's chapter, the idea resonated with me that evaluation activities should be additional learning opportunities for students. Reading through the chapter, I came to think that "collaborative assessment," where student-centeredness and student initiatives are encouraged, seems to be the key to such opportunities.

O'Malley and Valdez Pierce (1996) explained how students and teacher engage in "collaborative assessment" by means of portfolios in language classrooms: "[S]tudents get individual feedback on how to set and achieve goals, and teacher gets individual feedback on how to make instructional activities more meaningful and useful to students" (p.43), while they evaluate the student work together. Assessment by student-made tests can also be described as collaborative assessment in the sense that students and teacher give feedback to each other's "course work," or what they each do in or for the course.

In that case, what is the "feedback on how to make instructional activities more meaningful and useful to students" that the teacher receives from student-made tests? Students make test questions based on what they think is important and what they have learned. I concur in evaluating highly students' *can-do's*. On the other hand, students have to learn points of use in language learning even if they have not noticed them by themselves. If students fail to include such points in their test questions, the teacher will be reminded that he/she might want to redesign the course activities to emphasize these points by using more explicit instructions.

This chapter seems to indicate that Marlen did go through these thinking and acting processes. He tried to integrate oral communication skills, which were lacking in earlier student-made test questions. At the same time, I became more interested in what reading activities were used, when Marlen referred to his reading classes in his chapter. If his students made questions in the reading class, did the questions include items to test reading skills and those to measure reading comprehension level? For example, did they provide reading comprehension questions which required skimming or scanning skills? Did they make questions asking for reactions to what they had read to promote deeper understanding of the materials? When the teacher looks at student-made questions, he/she can think over the meaning of course objectives. If students did not include questions that measure or reflect achievement of the course objectives, then the teacher should think of what should have been more emphasized and how, and sometimes also whether the course objectives were appropriate or not.

Thus, collaborative assessment can promote not only student reflection but also teacher reflection. Reflection is essential for developing autonomy, because through reflection, students can set their learning goals which are more effective and appropriate for their own level. Once they set their goals, their actions will be more focused. Students then continuously engage in the cycle of "goal-setting → action → reflection → goal-resetting → action → reflection → ..." and develop their autonomy. This cycle works in a similar fashion for teachers, with the only difference being "goals" means "teaching goals" (or "learning goals about teaching!") Through such reflection, the teacher can expand their autonomy—their responsibility and control over their teaching processes. Student autonomy and teacher autonomy become synergistically enhanced when cycles of student and teacher reflection are connected through "student-teacher collaboration."