<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Issues in the evaluation of instruction and learning in distance education in Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Ysusaku Otsuka</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>1994</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/1263">http://hdl.handle.net/10220/1263</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td></td>
</tr>
</tbody>
</table>
Issues In The Evaluation Of Instruction
And Learning In Distance Education In Japan

By

Ysusaku Otsuka
Issues in the Evaluation of Instruction 
and Learning in Distance Education in Japan *

Ysusaku Otsuka 

National Institute of Multimedia Education 
Chiba, Japan

Introduction

It has been eleven years since we began distance higher education with an application of a broadcasting system in Japan. We have encountered problems peculiar to the long distance education, which are not generally regarded as issues in the traditional higher education system requiring face-to-face interaction between teachers and students. Distance education through a broadcasting system is characterized by such features as one-way delivery of information from instructors to learners, openness to a variety of audience, and self-initiated and individualized learning. New technologies of the media have enabled us to develop these dimensions of distance education making a delivery of information quite efficient. These technologies, however, have posed a critical problem in terms of the demand for two-way communication between instructors and students, which cannot be readily enhanced. In essence, this is a problem of feedback between instructors and students, an issue embedded in evaluations of teaching and student coursework. In the past exploratory instruments have been developed to evaluate instruction and learning with particular attention to promoting students' motivation.

The purpose of this paper is to explore problems of evaluation in distance education. First, I will briefly discuss the programs developed at the University of the Air; second, I will

* This paper is presented in the session on "The New Communication Environment and Media Education" at AMIC Conference on "Communications, Convergence and Development (The Changing Communications Scene in the Asia-Pacific: Implications for Development)" in Bangkok, Thailand, on June 25, 1994. I would like to express my deep appreciation to the following for their great assistance: Prof. Nobuo K. Shimahara, Rutgers University and Prof. Delwyn L. Harnisch, University of Illinois.
focus on problems of feedback between instructors and students. In addition, I will discuss issues of testing related to distance education.

A Trial of Distance Higher Education in Japan: The University of the Air

Modelled after the Open University in England, the University of the Air (UA), the first distance higher education institution in Japan that offers all courses through radio and television, was established in 1983 and began enrolling students in 1985. The National Institute of Multimedia Education (NIME), which was instrumental in establishing UA, has been producing nearly all UA programs at NIME's studios.

UA has only liberal arts programs with six areas of concentration and broadcast lectures in more than 300 courses, including foreign languages and some interdisciplinary studies. Anyone who wants to study can access any UA program at home or office. UA has exclusive TV channels and radio frequencies through which all lectures are broadcasted in the greater Tokyo metropolitan area, the so-called Kantoh area. UA is planning to expand the broadcasting area throughout the nation by using a satellite system. Probably UA is the most privileged university with exclusive channels in the world.

UA prints and delivers to its students the textbooks for courses that its instructors have written. Students study by following mainly textbooks and broadcast lectures. In order to earn credits they are expected to submit a mid-term report to their instructors and must pass final examinations. Students can earn two credits per a regular course, which involves fifteen 45-minute-broadcast lectures.

In addition, UA has several local study centers, where face-to-face instructions and final exams are offered and where students can also have access to the library and review taped lectures in the audio-video room. Moreover, a few classroom instructions, most of which have a subsidiary role vis-a-vis broadcast lectures, are given providing traditional face-to-face interaction between instructors and students. In this setting students must complete at least 20 classes, equivalent to 20 credits to graduate.

In seeking admission to UA as a regular student, it is necessary only to be a high school graduate. If applicants do not meet this minimum requirement for admission, they can still be admitted as special students who may be later admitted as degree students upon completion of 16 credits. It is evident that UA is open to nearly everybody over the age of 18.

Once they have been admitted, UA students must register for courses they are interested in studying, and they are expected to complete mid-term course assignments and send them to
UA instructors by mail. If they fail to complete the mid-term assignments, they are not qualified to take final exams. For the final examinations they must go to study centers. To meet the graduation requirements, students have the option of writing a thesis after they have completed 96 credits. They can graduate upon completion of a minimum of 124 credits in addition to other graduation requirements.

Diversity is a salient characteristic of UA students. For example, unlike traditional students, students' age varies from eighteen to over one hundred. The mean age of the students, however, is about 40. They have a variety of full-time work, for example, as teachers, public officers, businessmen, and housewives. It suffices to point out that students enrolled in the UA programs are diverse with respect to their backgrounds and age.

The Issues of Evaluation in the Distance Education

From the preceding discussion it is apparent that UA is a new and unique university in Japan. There are considerable numbers of unsolved and challenging problems that need to be attacked if its system is to be improved. Among them are technical and psychological issues related to the evaluation of learning and instruction, which will receive attention in this section.

Technical Problems

Here I will discuss two technical issues. One is how to evaluate courses or instruction, and the other is how to evaluate students' coursework.

(a) Course Evaluation

Regarding the evaluation of courses, UA instructors have only limited information about the effectiveness of their lectures and the quality of students' work. Most instructors can get information about students' coursework only from their mid-term assignments and final exams. It goes without saying that UA instructors need to develop a system of feedback to understand better characteristics of their students, the quality of their course work, and their perceptions of teaching. Developing such a system, however, is a very challenging task in Japan. Unlike most institutions of higher education in the United States, we have not yet established the practice of evaluating lectures. Most instructors at Japanese traditional universities appear to have fear that they may not receive fair evaluation from students. For example, lazy students who complain about trivial things may rate teaching unfairly low.
These instructors are afraid that such students may manipulate course evaluation to receive credits. In the past we have conducted a survey in which students evaluated their UA instructors to construct an evaluation instrument. Although we provided the instructors with the data we collected from students, they made little use of them to improve their courses. It seems that in addition to their fear, UA instructors find it difficult to readily improve teaching based on such data, given the fact that UA courses are restructured every four years.

We developed a strategy for evaluation of courses. Fig. 1 shows an outline of the UA course evaluation scheme. All UA lectures were taped. We used those taped lectures as materials for instruction in other contexts. Those lectures were given by such veteran and nationally noted professors that the instructors who were teaching at the other universities could refer to UA lectures as a model for their lectures or use a part of the UA lectures as teaching materials. We conducted a survey in which the taped lectures and textbooks were distributed to about a thousand faculty members at other universities, along with a questionnaire requesting their evaluation of the course materials. They evaluated the materials after they had used them in their classes, enabling us to receive course evaluations from a professional point of view. We are also considering constructing an evaluation instrument used by students as schematically presented below.

![Diagram of the UA-lecture Evaluation Scheme]

**Fig. 1  Outline of the UA-lecture Evaluation Scheme**
Based on this scheme we have developed an evaluation instrument, but we have encountered problems in making it a universally useful instrument.

First, our instrument involved only one questionnaire that includes more than 100 question items for all courses. We have found that it is difficult for evaluators of particular courses to respond to some items, because the teaching methods of those courses are either unique or situationally specific. To solve this problem, we are considering developing a more adaptable course evaluation instrument where instructors can select relevant items for evaluating their courses. In this proposed instrument there are a universal category of items that all instructors use and a category from which they may choose only relevant items, including items requesting the evaluators' written comments.

Second, we are reminded that we must take into consideration the features specific to UA's distance education in developing a valid and functional evaluation instrument. In addition to teaching, we must consider the process in which the teaching is the final product. UA courses are organized not only by the instructors but also by the producers for broadcasting programs who offer technical support. Teaching in this context is a joint product. We should consider relevant items specific to distance education, especially broadcast programs. Most of the items in our exploratory questionnaire are comparable to the items included in course evaluation instruments used in some American universities. This questionnaire, however, does not include items related to program or course production. For example, the use of computer graphics or animation, and the duration of bust shot of the talking instructor are highly important aspects of teaching in broadcasting courses. We are exploring such features specific to UA courses calling attention in the evaluation of teaching.

Third, it is difficult to provide appropriate and effective feedback to the instructors and producers with respect to the results of course evaluations partly because students' responses to the evaluation questionnaire are highly diverse. A statistical index such as mean score is insufficient in terms of substance and not so practical for improving courses. On the other hand, a detailed analysis of all items would make the feedback of results extremely complex. We are exploring a model of analysis that will enable us to provide substantive feedback to the instructors and producers by breaking down the evaluation data provided by students.

(b) Testing: Item-banking and an Application of IRT

It is an another big challenge to produce appropriate tests in such a huge system like UA. There are a few courses which enroll more than a thousand students. As a result, instructors need to construct a multiple-choice test for the efficiency of scoring. But most UA instructors
are not acquainted with developing such an objective test, which might be too heavy a burden on them.

In addition, it is difficult to maintain the same criteria for constructing tests every term. Fig. 2 shows an example of changes of mean scores over 9 terms for a UA course. The horizontal axis indicates terms. The vertical axis on the left side indicates success rates displayed by the solid line. The dotted line is the mean scores indicated by the vertical axis on the right side. In the first four terms, changes in mean scores and success rates are very drastic. A test maker for this course who was not the instructor, revealed that in the first term he could generate relatively general test items but that in the second term he paid specific attention to the content of the course in constructing them. In other words, the test items he constructed for the first two terms varied because he assumed that he must not use the same items for two consecutive terms. As a result, many students failed the test in the second term.

After the 5th term, an item-bank system was created for the course. Consequently, students' mean scores and success rate became relatively stable. By using the item-bank system, test makers can use test criteria more readily in constructing content-specific tests.

Item Response Theory (IRT), which is a relatively new theory for analysis of test scores, is useful to equate scores between terms. We can obtain an expected (statistical) distribution of test scores based on IRT, which enables us to anticipate the mean score, success rate, and information function.

![Fig.2 Changes of mean scores over 9 terms for a UA course](image-url)
Fig. 3 shows an example of expected test scores in the course in the fifth term in which the items used previously were used again. There is no perfect correspondence between the expected distribution and actual distribution of scores, as indicated by white bars and black bars, respectively. Fig. 3 suggests that the examinees' actual text scores are lower level than the theoretically expected scores, although it should be noted that variation between the two is relatively small. IRT is especially useful in predicting test scores for UA students whose academic quality is relatively stable over time. Students who cannot pass the exam at first can retake another exam in the next term. In UA large numbers of new students are not enrolled each year as in traditional universities. The level of students who take the exam in the second term may be slightly lower. We can infer a small variation resulting from this situation.

Since 60 is a cut-off point in a course at UA, it is desirable that the peak of information curve displaying the precision level of measurement, is around 60-point. The information curve of Fig. 3 also suggests that we can get a relatively exact prediction of the percentage of students who will earn credits by passing the tests.

Such a test-bank system has not been used widely in Japan, because it is cumbersome for instructors to construct many items at first. Moreover, it seems to me that they do not prefer a multiple-choice item format. Japanese professors are inclined to believe that in higher education it is better to evaluate students' coursework based on essay-type assignments to encourage students to develop conceptual skills than by objective-type tests, which are
believed to measure only students' memorization.

Moreover, there are some problems in the current test–bank system, such as its inability to include special characters or pictures and difficulties to treat long sentences. Thus we are trying to develop a new test–bank system to make it more flexible and more widely applicable.

At present UA students must go to specified study centers to take examinations, while they can study essentially anywhere and anytime. This policy seems discrepant. As an alternative we are considering computerized testing as a practical and feasible method. Probably it will take a considerable time to develop such a system, but I think it is vital for educational technologists to develop it.

**Psychological Issues**

Even if the technical problems are solved, we cannot evaluate teaching and learning sufficiently without determining what it is that we are trying to evaluate. In the UA programs it is very laborious to determine the intrinsic goals of distance education for individual students. Nevertheless, we can identify relatively common characteristics of UA students, which we do not often see at other universities. One prominent, shared characteristic of UA students is that they display a high level of intrinsic motivation for learning as part of life–long learning.

Table 1 stems from the results of our survey of UA students and students at other universities conducted in 1987 (Otsuka, 1989; Otsuka & Iwanaga, 1989). We constructed scales based on several items in the questionnaire and standardized the means at 50 with standard deviations at 10. This table presents results of the survey on motivation for learning and on the evaluation of universities as a whole. In comparison with traditional university students, UA students are highly motivated toward learning. The latter are not enrolled in UA programs to seek a degree or to gain promotion in their job. This appears to be a universal attribute of many US students.

Most distance education programs in other nations are designed to orient coursework to an attainment of degrees and new jobs or promotion to better positions. In contrast, UA's central philosophy is to promote life–long learning without specific focus on attaining extrinsic rewards. One may wonder how UA students maintain a high level of intrinsic motivation. I am sure that getting new skills and knowledge must be an important purpose of education. Nevertheless, I would like to emphasize that enhancing such intrinsic motivation should be another vital purpose of education, especially in the case of distance education in Japan.
Table 1. Motivation for learning and evaluation of university

<table>
<thead>
<tr>
<th></th>
<th>UA regular students</th>
<th>other traditional university students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>scale</td>
<td>mean</td>
</tr>
<tr>
<td>Motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>general culture</td>
<td>52.9</td>
<td>9.3</td>
</tr>
<tr>
<td>degree or job</td>
<td>48.6</td>
<td>9.2</td>
</tr>
<tr>
<td>friendship</td>
<td>47.1</td>
<td>8.5</td>
</tr>
<tr>
<td>learning itself</td>
<td>53.8</td>
<td>7.4</td>
</tr>
<tr>
<td>accessibility</td>
<td>53.9</td>
<td>9.5</td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>satisfaction with education</td>
<td>54.6</td>
<td>8.5</td>
</tr>
<tr>
<td>appropriateness of education</td>
<td>52.3</td>
<td>9.5</td>
</tr>
<tr>
<td>sense of fulfillment of campus life</td>
<td>46.4</td>
<td>8.0</td>
</tr>
<tr>
<td>difficulties of learning</td>
<td>52.9</td>
<td>9.7</td>
</tr>
</tbody>
</table>

The majority of UA students have full-time jobs and experience constraints in finding time for study much more than other university students as readily seen in Table 1. In sharp contrast, other university students enjoy leisurely campus life and associations with their peers, as indicated by the fulfillment of campus life scale. Yet UA students' rates of satisfaction and appreciation of education are much higher than the other university students'. It is significant to note that although there are a number of major problems in the UA education system to be resolved, UA students are satisfied with the fact that they have the opportunity to study. However, apart from their satisfaction with the programs in which they are enrolled, we should take the profiles of UA students into careful consideration to improve their programs.

It is vital to further investigate the level of UA students' motivation toward studies to
enhance it. Although we need more research to identify factors contributing to their high intrinsic motivation, I suggest a few extrinsic factors underscoring their education for discussion. For example, an actively interactive learning, which is promoted by multimedia materials, might be an important influence on sustaining students' high motivation. And as motivation theory suggests, their sense of control over what they learn can be another relevant source of motivation. A sense of control stems from the controllability of knowledge that they acquire and the pace with which they learn it. So the locus of control in learning seems to play an important role in enhancing students' motivation toward studies. An exploration of factors contributing to the development of students' intrinsic motivation should include developing an evaluation system to improve teaching.

Summary

In this paper I have discussed salient characteristics of the UA programs that enhances students' goals for life-long learning rather than immediate-job related aspirations, a primary concern of recurrent education. In order to promote the programs, it is critical to develop an effective system of feedback between instructors and students. I have reported our surveys and current project focused on course evaluation and student motivation. In the future teaching methods, including course evaluation and testing, will be computerized and personalized. In a computerized environment it is vital to maintain a balance between the computerization of teaching and human psychology. It is people's motivation to learn that takes precedence over electronic technologies. That is a governing principle in a computerized higher education.

Reference


RESUME

Yusaku OTSUKA, 5/25, 1994

Name: Mr. Yusaku OTSUKA  (age: 41; birth: July 29, 1952)

Educational Background

M.A. in Educational Psychology, the University of Tokyo, March 1980.
B.A. in Educational Psychology, the University of Tokyo, March 1978.
B.A. in Physics, the University of Tokyo, March 1976.

Occupation

1986-Present: Associate Professor, The National Institute fo Multimedia Education
1982-1986: Research Associate, The National Center for University Entrance Examination

Fields of Research

Educational Evaluation; Cognitive Psychology; Educational Psychology; Data Analysis

Publication