

**ASSESSING THE LANGUAGE COMPONENT OF
THE MANOA GENERAL EDUCATION REQUIREMENTS**

Report on the College of Languages, Linguistics, and Literature (LLL) Assessment Committee's
2002-04 assessment activities ¹

Kimi Kondo-Brown
East-Asian Languages and Literatures
University of Hawaii at Manoa
1980 East-West Road
Honolulu, HI 96822
kondo@hawaii.edu
(808) 956-2037

April 2004

¹ This project was funded by the 2002-03 University of Hawaii Assessment Fund. The 2002-03 LLL assessment committee members were (alphabetical order): James Dean Brown (SLS), Paul Chandler (LLEA), Precy Espiritu (HILL), John Haig (EALL), Kathryn Klingebiel (LLEA), Kimi Kondo-Brown (Chair, EALL), John Mayer (HILL). The 2002-04 LLL assessment committee members are John Haig (EALL), Kimi Kondo-Brown (EALL), Melodee C. Metzger (LLEA), Jean Toyama (Chair, Associate Dean), Dina Yoshimi (EALL), Julius Soria (HIPLL), and Tao-Chung Yao (EALL). James Dean Brown continues to provide advice for this project.

TABLE OF CONTENTS

List of tables and figures.....	2
1. Defining terms and the scope of the 2002-04 LLL assessment project.....	3
2. What is the purpose and significance of the present assessment project?.....	5
3. How are we doing it?: An overview of five steps of the present assessment project.....	6
4. <u>Step 1</u> : Conduct an assessment survey to identify (a) key concrete and measurable learning outcomes and (b) preferred assessment tools and procedures (Fall, 2002-Spring 2003).....	7
4.1. Goal	
4.2. Survey participants	
4.3. Survey Instruments	
4.4. Survey Procedures	
4.5. Survey results	
4.5.1. Recommended learning outcomes	
4.5.2. Preferred assessment tools	
4.5.3. Preferred assessment procedures	
5. <u>Step 2</u> : Develop assessment instruments and procedures to assess one of the identified objectives in different languages for the 2003-04 pilot testing (Fall, 2003)	13
6. <u>Step 3</u> : Implement the instruments and procedures to assess the target objective in different languages (Fall, 2003)	15
7. <u>Step 4</u> : Compile and analyze the data from the participating language programs (Spring, 2004)	17
7.1. Example test item analysis to examine the effectiveness of the test	
7.2. Example performance analysis to estimate the degree of learning that occurred	
8. <u>Step 5</u> : Present the results of the data to the faculty members and plan for future actions (Spring, 2004)	21
9. Final remarks: What factors have contributed to the successful initiation and maintenance of the project to date, and how will we go beyond the present stage?	22
Appendix 1. Survey participants	25
Appendix 2. Survey results concerning learning outcomes	26
Appendix 3. Survey results concerning assessment tools	31
Appendix 4. Survey results concerning assessment procedures	34
Appendix 5. Language requirement assessment planning questionnaire	35

LIST OF TABLES AND FIGURES

- Table 1. Survey participants
- Table 2. Recommended learning outcomes for LISTENING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items
- Table 3. Recommended learning outcomes for SPEAKING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items
- Table 4. Recommended learning outcomes for READING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items
- Table 5. Recommended learning outcomes for WRITING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items
- Table 6. Recommended learning outcomes for CULTURE across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items
- Table 7. Possible assessment tools for measuring LISTENING skills
- Table 8. Possible assessment tools for measuring SPEAKING skills
- Table 9. Possible assessment tools for measuring READING skills
- Table 10. Possible assessment tools for measuring WRITING skills
- Table 11. Possible assessment tools for measuring CULTURAL KNOWLEDGE
- Table 12. Assessment procedures 1 (When should the assessment be conducted?)
- Table 13. Assessment procedures 2 (How often should we assess?)
- Table 14. Assessment procedures 3 (Who should be assessed?)
- Table 15. Assessment procedures 3 (Who should do the assessment?)
- Table 16. Test item analysis 1 (Item facility and distractor efficiency)
- Table 17. Test item analysis 2 (the *B*-index)

List of figures

- Figure 1. Direct and indirect data collection methods for assessment
- Figure 2. Steps for product-oriented approaches (Hammond, 1973: 168)
- Figure 3. Steps for the 2003-2004 LLL assessment project for the language component of the Manoa general education requirements
- Figure 4. Learning outcomes for Listening
- Figure 5. Learning outcomes for Speaking
- Figure 6. Learning outcomes for Reading
- Figure 7. Learning outcomes for Writing
- Figure 8. Learning outcomes for Culture
- Figure 9. Comparisons of test score distributions for the 201 and 202 groups

1. Defining terms and the scope of the 2003-04 LLL assessment project

Evaluation terms such as "assessment" and "tests" are often used with different definitions in the literature. In this report, I define these terms as follows with specific reference to the present context, i.e., assessing the Hawaiian/second language requirement at the University of Hawaii at Manoa (UHM).

Assessment: Erwin (1991) defines assessment as " the process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to increase students' learning and development" (p. 15).² Likewise, in the present report, assessment is defined as an ongoing process aimed at evaluating and improving student learning in Hawaiian/ second languages. Such evaluative processes necessarily involve:

1. Making the expected learning outcomes of the Hawaiian/second language requirement explicit,
2. Systematically collecting, analyzing, and interpreting data to determine the degree to which the actual student learning matches our expectations, and
3. Using such evaluation information to improve student learning.

In the preset context of assessment, "student learning" can be evaluated not only with reference to performance in Hawaiian/second languages but also with reference to attitudes toward the target cultures. However, for logistic reasons, this 2002-04 assessment activity has focused on the assessment of student performance in Hawaiian/second languages. In other words, the present project is best characterized as "a focused assessment project" with a specific focus on the assessment of certain performance outcomes in Hawaiian/second languages, which was evaluated primarily with the use of tests.

² Erwin, T. D. 1991. *Assessing student learning and development : a guide to the principles, goals, and methods of determining college outcomes*. Hoboken, NJ: Jossey-Bass.

Tests: As shown in Figure 1 below, the use of tests is only one of the many sets of tools that can be used in assessment activities. Based on the information obtained by the LLL assessment survey for instructors, which will be explained in detail later, we have decided to develop *criterion-referenced*, faculty-made achievement tests to obtain the kind data we wanted. The primary purpose of criterion-referenced tests is to measure the amount of learning that a student has accomplished on given objectives, which is very different from norm-referenced tests (e.g., Japanese proficiency test, SAT), which are designed to measure more global abilities of individual students, which are, in turn, interpreted with reference to all other students' abilities (Brown, 1996).³

<u>Direct data on learning</u>	<u>Indirect data on learning</u>
- Criterion-referenced achievement tests	- Student opinions and attitude surveys
- Norm-reference proficiency tests	- Exit interviews
- Capstone course (i.e., 202) evaluation	- Administrative data (e.g., class enrollment)
- Portfolio	- Alumni & employer surveys
- Curriculum syllabus analyses	

Figure 1. Direct and indirect data collection methods for assessment

³ Brown, J. D. 1996. Testing in language programs. Upper Saddle River, NJ: Prentice Hall.

2. What is the purpose and significance of the present assessment project?

According to the UHM general education descriptions, “. . . proficiency in Hawaiian or a second language is an integral part of the university’s mission to prepare students to function effectively in a global society to preserve and promulgate Hawaiian, Asian, and Pacific language, history, culture and provide students an education experience with an international dimension.” In order to achieve this goal, “before graduation all students must show competency at the 202 level.” However, the quoted description of the Manoa Hawaiian/second language requirement is very general. For example, we do not know what is meant by “202-level competency” in different language programs.

Thus, in order to make the assessment of the Hawaiian/second language programs an integral part of on-going curriculum activities within the College of LLL, first, we needed to identify concrete and measurable learning outcomes or objectives, which may be common across all programs. Once these objectives were identified, we then were in a position to develop appropriate assessment instruments for determining which objectives are being met and which ones are not. In other words, the present assessment project was intended to redefine future learning objectives for the core Hawaiian/second language programs and develop assessment instruments and procedures to measure the effectiveness of these programs in achieving these objectives. This type of cross-departmental assessment project had never previously been conducted within the College of LLL. In order to make the assessment of the Manoa Hawaii/second language programs a regular part of our teaching activities, we believed the present initiative would be crucial.

3. How are we doing it?: An overview of five steps for the present assessment project

There are various approaches available for doing a particular assessment project. The present assessment project falls into the category of "product-oriented approaches" (Hammond, 1973).⁴ Figure 2 below illustrates the steps for product-oriented approaches proposed by Hammond (1973: 168). Following this model, the present assessment project has been taking the steps illustrated in Figure 3. In the rest of the report, the details of each step will be described and discussed.

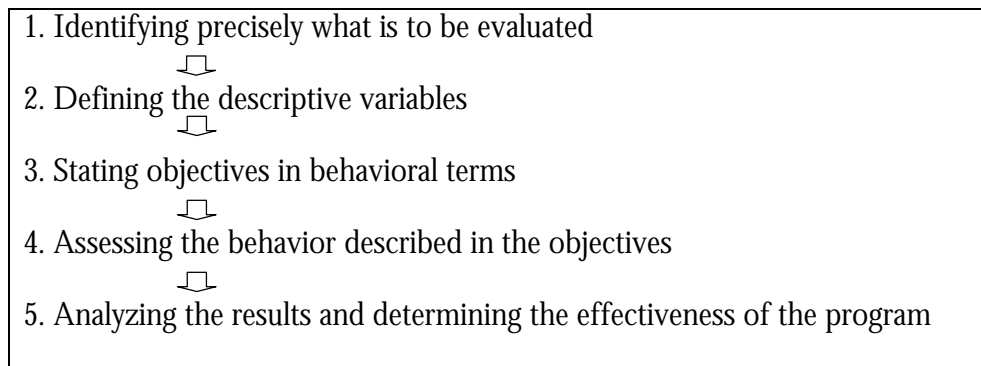


Figure 2. Steps for product-oriented approaches (Hammond, 1973: 168)

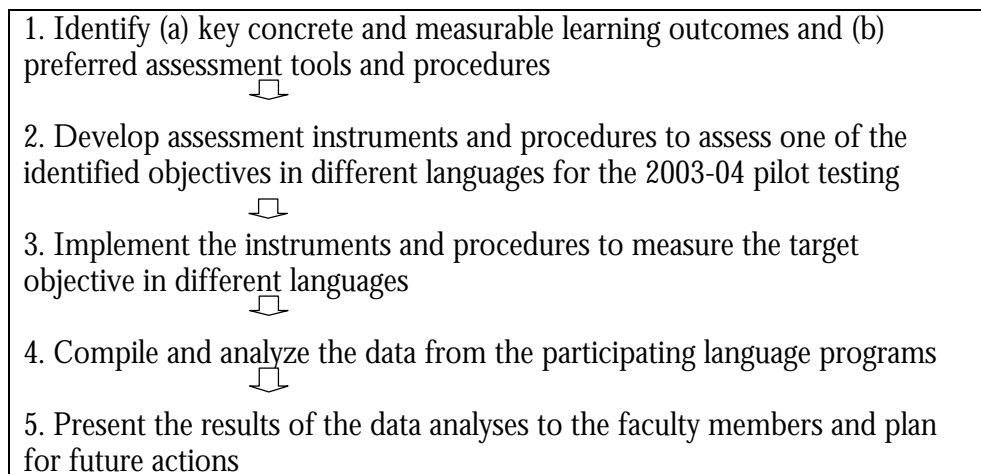


Figure 3. Steps for the 2003-04 LLL assessment project for the language component of the Manoa general education requirements

⁴ Hammond, R.L. 1973. Evaluation at the local level. In B. R. Worthen, & J.R. Sanders (Ed.), *Educational Evaluation: Theory and practice*. Worthington, OH: Charles A. Jones.

4. Step 1: Conduct an assessment survey to identify (a) key concrete and measurable learning outcomes and (b) preferred assessment tools and procedures

4.1. Goal

We conducted an assessment survey for the faculty members and graduate teaching assistants (TAs) teaching in the Manoa Hawaiian/second language programs in order to gather the cross-departmental information necessary for answering the following questions:

1. What should the learning outcomes for the two-year Hawaiian/second language requirement at UHM be?
2. What would the potential assessment instruments be to systematically measure each of the identified outcomes?
3. In conducting assessment using the identified instruments, what assessment procedures should be adopted?

4.2. Survey participants

All faculty members (professors, instructors, and lectures) and TAs teaching in the Hawaii/Second language programs were invited to participate in the study as volunteers. An assessment questionnaire was sent to a total of 192 faculty members and TAs in three departments (East Asian Languages and Literatures [EALL], Hawaiian Indo-Pacific Languages and Literatures [HIPLL], and Languages and Literatures of Europe and the Americas [LLEA]). Of those surveyed, 92 faculty members from the following 22 language programs responded to the survey (for a 47.9% return rate): Arabic, Cambodia, Chamorro, Classics (Greek and Latin), Chinese, Filipino, French, German, Hawaiian, Hindi, Japanese, Korean, Ilokano, Indonesian, Maori, Russian, Samoan, Spanish, Sanskrit, Tahitian, Thai, and Vietnamese (see Table 1 in Appendix 1 for more detailed background of the survey participants).

4.3. Survey instruments

I drafted the questionnaire primarily based on the information obtained through a preliminary electronic survey among program coordinators regarding the existing lower-division language programs (100-202 level) in terms of their objectives and teaching practices focused particularly on assessment. On the preliminary survey, I asked the following questions:

1. Do you have a document that lists learning objectives for students completing the two-year program in your language or for each level? If so, how can we get a copy?
2. Has your program developed any assessment instruments to measure the learning outcomes of students completing the two-year program in your language? If so, please describe them briefly. If not, what do you think possible assessment instruments might be for measuring the learning outcomes of students completing the two-year program in your language?

After I drafted the questionnaire based on the information obtained from the program coordinators, I circulated the draft among the members of the LLL Assessment Committee (which was formed based on departmental chairs' recommendations) for feedback.

The final version of the questionnaire had five sections. The first section asked the survey participants about their backgrounds (department, position, language(s) taught, and for which languages they are giving their opinions). The second section asked the participants to indicate the degree to which they agree with 42 statements as learning outcomes for students who complete a fourth-semester course in the language(s) for which they are giving their opinions (i.e., 5=strongly agree, 4=agree, 3=no opinion, 2=disagree, and 1=strongly disagree). These learning outcomes were listed in five categories: four language skills and cultural knowledge. At the end of each category, the participants were asked to provide any additional learning outcome(s) they would suggest for the category. The third section asked the participants to choose potential assessment instruments for measuring the identified student learning outcomes in the five categories mentioned above. This section also included a question asking about the participants' previous experiences with proficiency

testing (e.g., training in an oral proficiency interview, development of other types of proficiency tests, etc.). The final section asked the participants how they would like to conduct assessment using the instruments of their choice (e.g., when and how often should we assess, who should be assessed, and who should do the assessment). The entire questionnaire is attached as Appendix 5.

4.4. Survey procedures

A cover letter that explained the purpose and procedures involved in the survey and an anonymous LLL assessment questionnaire form were placed in the mailbox of each of the 192 faculty members and TAs in EALL, HIPLL, and LLEA in November, 2002. Before the questionnaire was distributed, the project proposal had been reviewed by the UH Committee on Human Studies and exempted from Department of Health and Human Services regulations. The survey participants were asked to return the questionnaire to the box labeled “LLL assessment survey” located in their main department office in a few weeks. They were also asked to make a check-mark by their name on the list when they returned the questionnaire so the investigator could keep track of the faculty who had responded to the questionnaire. By the survey deadline, approximately 40 faculty members and TAs had volunteered to participate in the survey and returned the questionnaire. Then, the Dean of LLL sent out a reminder letter and, by the second deadline, an additional 51 faculty members and TAs had returned the questionnaire, allowing for the acceptable 47.9% return rate ($N=92$).

4.5. Survey results

4.5.1. Recommended learning outcomes

Out of 92 returned questionnaires, responses to the section 2 (recommended learning outcomes) from four participants were not considered for the present analysis. The four participants were

faculty members teaching classic languages and, on their responses, they suggested their learning outcomes be evaluated separately because their learning objectives are considerably different from those of modern languages. Therefore, the remaining 88 responses were used for the present analysis.

Tables 2-6 in Appendix 2 summarize the results of participant responses to the section two. In this section, the participants were asked to indicate the degree to which they agree with 42 statements as learning outcomes for students who complete a fourth-semester course in the language(s) for which they are giving their opinions. Each table summarizes the results of one of the five categories: listening skills (Table 2), speaking skills (Table 3), reading skills (Table 4), writing skills (Table 5), and cultural knowledge (Table 6). Each of these tables shows means, standard deviations, and rankings of means (out of 42 items) for learning outcomes in respective category based on ratings given by all and individual department groups. In each table, the learning outcomes are sorted from the most recommended to the least by all participants. Figures 4-8 in Appendix 2 present visually the means of learning outcome ratings given by all participants for each of the five categories.

The tables and figures indicate that the following eight learning outcomes received the strongest recommendations across departments/programs:

1. Understand conversations about everyday experiences (e.g., school, work, interests, preferences)
2. Understand *factual* content of paragraph-length descriptions/narratives on familiar topics (e.g., recorded telephone instructions, announcements in public areas)
3. Perform a variety of “real-life” tasks in common social and transactional situations (e.g., shopping, making hotel reservations)
4. Sustain conversations/interviews about self, family, experiences, interests, and preferences
5. Understand fully paragraph-length texts dealing with personal and social needs such as personal letters, messages, and memos

6. Get main ideas from authentic everyday practical materials written entirely in the target language (e.g., menus, ads for products)
7. Meet practical writing needs and social demands by writing paragraph-length personal letters, messages, applications, and journals
8. Demonstrate understanding of holidays and traditions celebrated in the target culture

4.5.2. Preferred assessment tools

Out of the 92 returned questions, 80 participants actually responded to the section three (preferred assessment tools), and Tables 7-11 in Appendix 3 each summarizes the results of 80 participants' opinions concerning possible assessment tools for each of the five skill categories (i.e., listening, speaking, reading, writing skills and cultural knowledge).

Each of these tables shows raw frequencies, percentages, and rankings of preferences within the skill category based on ratings by all as well as individual department groups. As you will see in the following list of major findings, there seems to be strong interest in developing faculty-made achievement tests embedded in final exams across various language programs.

1. More than 75% of the participants chose faculty-made oral achievement test embedded in the final exam for measuring listening and speaking skills. More than 70% of LLEA participants also indicated that the use of available oral proficiency interview tests is also a possibility.
2. More than 75% of the participants chose "faculty-made paper-and-pencil achievement test embedded in the final exam" for measuring reading skills. Approximately 70% of the participants also chose this method for measuring writing skills. In addition, self-assessment questionnaire and portfolio assessment each was also a popular choice among HIPLL and LLEA participants, respectively.
3. Opinions about how to measure cultural knowledge seemed divided, although more than half of the participants chose report writing and its presentation as a possible method.

4.5.3. Preferred assessment procedures

Tables 14-17 in Appendix 14 each summarizes the results of participants' opinions concerning possible assessment procedures in terms of (a) how the assessment should be conducted, (b) how often we should assess, (c) who should be assessed, and (d) who should do the assessment. Each of these tables shows raw frequencies, percentages, and rankings of preferences within the skill category based on ratings by all as well as individual department groups. The following were the major findings regarding recommended assessment procedures.

1. More than half of the participants think that assessment should be conducted every semester, at the end of first- and second-year courses among all target students in all languages.
2. Opinions as to who should do the assessment seemed divided: Preferences were given to 'all faculty members as part of their work' (32.6%) and 'two faculty members from each department designated as assessment coordinators with a possible course release' (28.3%).

5. Step 2: Develop assessment instruments and procedures to assess one of the identified objectives in different languages for the 2003-04 pilot testing

In April 2003, I presented the results of the LLL assessment survey to all relevant administrators in the College of LLL (i.e., the Dean and Associate Dean as well as the Chairs of EALL, HIPLL, and LLEA) with the purpose of discussing strategies for continuing this college-wide assessment project in which the effectiveness of the existing two-year Hawaiian/second language programs will be systematically monitored in a manner that reflects the results of the present survey. At that meeting, it was first decided to form a new LLL Assessment Committee to undertake the next phases of this project. During the meeting, it was also decided to conduct a pilot assessment of selected two-year Hawaiian/second language programs (Chinese, French, Hawaiian, Ilokano, Japanese, Samoan, and Spanish) during the 2003-04 academic year.

In May 2003, via email, the new LLL Assessment Committee Chair announced the next phase of this continuing assessment project to the entire faculty in the involved departments and sent them the major findings of the LLL assessment survey. In Spring semester 2003, the committee discussed ways to implement the first pilot assessment during the semester. We decided to focus on one of the eight strongly recommended learning outcomes: "Understand fully paragraph-length texts dealing with personal and social needs such as personal letters, messages, and memos." This learning outcome proved to be the first or second most popular choice among the various language programs as the learning outcome to be assessed in the pilot assessment.

Then, the representatives of each language program agreed to develop a single set of multiple-choice reading comprehension test items for inclusion on second-year level final exams in order to assess this learning outcome. The committee specified common characteristics of the test to be developed across various language programs as follows:

1. Test format: one page test with a text of a few short paragraphs followed by five or six questions, which the students can finish in a short time (i.e., 10 minutes or so).

2. Prompt: A personal letter sent in early December or so from a native speaker of the target language who is visiting/living in Hawaii to her/his friend in the target country. The sender describes (a) what he or she did during the thanksgiving and (b) his or her plans for the Christmas holidays. As an achievement test, it should be designed with specific reference to the instructional content of a given program. In other words, the prompt must be consistent with the written materials used in the 101-202 course sequence. Where appropriate, additional items taught in the third level may also included for experimental/proving proposes.
3. Questions: Multiple-choice test with four options.
4. Administration: The test will be given to 202 and/or 201 students (depending on the courses offered in each language program) at the end of the semester (e.g., part of the final exam or during the last week of instruction).
5. Use of the test scores: Test scores will not be included in the student's final grade, but their performance may be taken into consideration as extra credit.

After the draft test developed by each language program was created, each program presented the test at a meeting of all representatives in order to ensure that the items were as similar as possible across all language programs. It must be noted that, since the political situation in each department was different, specific test writing procedures (e.g., who prepares the test) remained somewhat flexible. In the case of Japanese, a volunteer test writer who had years of experience in teaching the lower-division Japanese drafted the test, and then I worked with the test writer on revisions.

6. Step 3: Implement the instruments and procedures to measure the target objective in different languages

It turned out that five language programs (Chinese, French, Ilocano, Japanese, and Spanish) participated in the 2003-04 pilot assessment project. The ways the pilot testing was implemented were different from one language to the other because we needed to take the political situation of each language program into consideration. In the case of Japanese, we used the following procedures for the implementation of the pilot testing. The Japanese section head informed the Japanese faculty members of the present pilot assessment project during a section meeting, clarifying and explaining the nature of the present project (e.g., it is a part of assessing the Manoa two-year language requirement, and not for evaluating individual teachers/students) and how the data will be handled (e.g., tabulated and reported anonymously as grouped data). Then later, the Japanese section head sent a memo to all 201 and 202 instructors to seek volunteers to have their students participate in the pilot assessment. It turned out that 37.3% of all 201/202 Japanese students participated in the pilot testing. This participation rate was judged to be acceptable for the first year pilot testing.

The test was given in class during the final week of instruction. This procedure was preferred by the Japanese faculty rather than the other option of giving the test as part of the final exam. The tests were distributed to teachers in an envelop one or two day(s) before the scheduled administration. With each envelop, the following instructions for the participating instructors were given to maximize the validity of the obtained data:

1. Open this envelope on the day you administer the test.⁵
2. Allow exactly 10 minutes to complete the test
3. Encourage your students to do their best in order to have valid data for the program evaluation. Beyond-202 level items are included for probing purposes, so please encourage

⁵ One of the complaints in using criterion-referenced tests is that they will cause teachers to "teach to the test" (Brown, 1996, p. 7). In order to minimize this effect, I asked the teachers not to open the envelope until the day of the test administration.

your students to make their best guess when they do not know the answer or understand certain parts of the text without your help.

4. Students do not have to write their names on the test. However, if you wish to have scores for individual students, I need either their names or social security numbers.
5. After you finish, please return this envelope with ALL TESTS inside to my mailbox.
6. Please do not make a copy of this test.
7. If you have comments/suggestions about the test and/or special instructions for how I should analyze the data, please write them here.

7. Step 4: Compile and analyze the data from the participating language programs

Each language program was asked to send me the tabulated EXCEL data as an email attachment. In this way, I received data from 521 students in five language programs. Then, I performed (a) test item analysis to examine the effectiveness of the test and (b) performance analysis to estimate the degree of learning that occurred. What follows are example analyses given for one of the participating language programs.

7.1. Example test item analysis to examine the effectiveness of the test

In this language group, there were six items and each item had four choices. The first four items were ones that they expected their 202 students to be able to answer. The last two items (#5 & #6) included some third-level materials for experimental purposes. Thus, the item writers anticipated that the great majority of the 202 students could answer the first four questions, but not the last two questions. Therefore, the cut point of this test (that divided students who were at the mastery level from those who were not) was set at four points. In other words, if a student scored four or above, it was considered to be satisfactory performance.

Table 16 (next page) shows the results of one type of test item analysis. The first column of the table indicates the item number. The second and third columns indicate the item facility (IF) (or the item difficulty) of each question for the 201 and 202 groups. Each IF was estimated simply by calculating the percentage of students who answered the given item correctly. For example, for the first item, 71% and 77% of 201 and 202 students answered correctly respectively. The next four columns indicate the efficiencies of distractors. For example, for "item 1," 73% of the entire group of students chose the correct answer, "b" (marked by the star *), and 11%, 3%, and 12 % chose "a," "c," and "d," respectively. When the efficiency estimate for a given distractor is "0," it means that the particular distractor is not attracting any students at all, and should perhaps be modified.

Table 16. Test item analysis 1 (Item facility and distractor efficiency)

Item number	Item Facility (IF)		Distractor Efficiency (all students)				
	201	202	a	b	c	d	Missing
1	.71	.77	.11	.73*	.03	.13	.00
2	.73	.75	.74*	.05	.20	.00	.01
3	.50	.77	.07	.61*	.15	.17	.01
4	.43	.73	.17	.11	.55*	.15	.02
5	.45	.55	.11	.30	.09	.49*	.01
6	.26	.64	.13	.41*	.04	.40	.01

* correct answers

Table 17 below shows the results of another type of test item analysis using the "*B*-index." The *B*-index is "an item statistic that compares the IFs of those students who passed a test with the IFs of those who failed it" (Brown, 1996, p. 81). In other words, it indicates the degree to which a group of students who were at the Mastery level (i.e., those who scored four or above) outperformed the ones who were not at that level (i.e., those who scored three or less) on each item.

Table 17. Test item analysis 2 (the *B*-index)

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
IFpass	0.96	0.89	0.92	0.81	0.72	0.63
IFail	0.49	0.58	0.27	0.27	0.26	0.16
<i>B</i>-index	0.47	0.31	0.65	0.54	0.46	0.47

For example, the first column of Table 17 indicates that the IFs for "item 1" of mastery students ($n=74$) and non-mastery students ($n=68$) were .96 and .49, respectively, resulting in the *B*-index of .47 ($.96-.49=.47$). Naturally, the higher the *B*-index, the more effectively the particular item can be said to separate the students who passed the test from those who failed. With the present data, the *B*-indexes ranged from .31 to .65, which were considered satisfactory.

In terms of test reliability, however, this test was a little lower than expected, the K-R21 reliability index (i.e., a conservative estimate of the *phi* dependability index for criterion-referenced test, see Brown, 1996, p. 221) for the present reading test was .576, which is reasonably high for such a short instrument, but in order to improve the reliability of the test, adding more items is recommended for the next round of piloting.⁶

7.2. Example performance analysis to estimate the degree of learning that occurred

The mean and standard deviation for the 201 group were 3.1 and 1.6, respectively, and the mean and standard deviation for the 202 group were 4.2 and 1.5, respectively. A *t*-test performed on these two sets of scores suggested that the mean for the 202 group was significantly higher than that for the 201 group ($p < .000$). Figure 9 (next page) graphically compares the score distributions for the 201 and 202 groups. As you can see in the figure, while the scores for the 201 group were normally distributed, the scores for the 202 group were negatively skewed. Figure 9 also shows that, while 37.2% ($15.1+14+8.1=37.2$) of the 201 students performed at the mastery level (i.e., four points or higher), a large majority (77%) of the 202 students in this language program performed at that level. In other words, the overall performance of the 202 students was satisfactory, and more learning seemed to be demonstrated for the 202 students (who had [almost] finished their language requirements) than for the 201 students (who had not yet finished).

⁶ For example, the Spearman-Brown Prophecy formula (see Brown, 1996, p. 195) allows us to calculate what the reliability would be if six similar items were added for a total of 12 items, which, in this case, turned out to be .73.

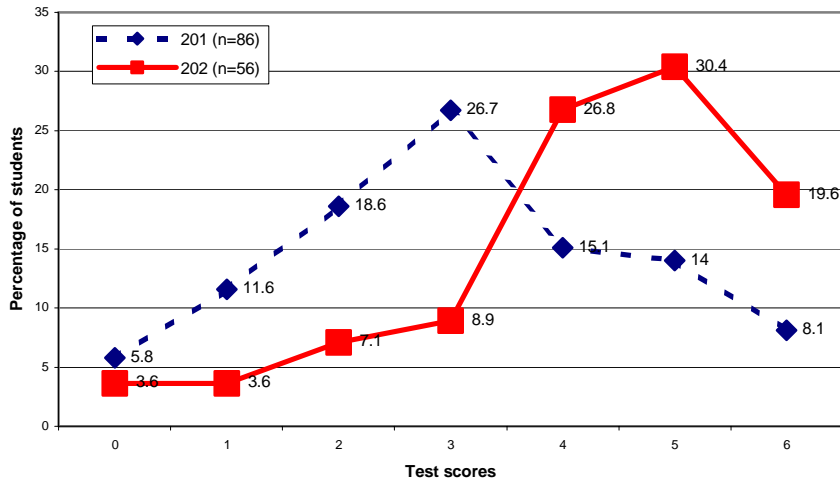


Figure 9. Comparisons of test score distributions for the 201 and 202 groups

8. Step 5: Present the results of the data to the faculty members and plan for future actions

In Fall 2003, at an LLL assessment committee meeting, I presented my analyses of the data from the first round of pilot testing to the representatives of the participating programs asking them to interpret the results and use their interpretations for the future improvement of assessment instruments. In the case of Japanese section, the section head discussed the analyses of the obtained assessment data with the faculty members at a Japanese-section meeting and encouraged them to continue our assessment efforts to monitor and improve the learning of our students.

For this semester (Spring, 2004), we have decided to continue focusing on the same objective and obtain the data using the revised tests, i.e., tests improved based on the previous year's test item analysis (e.g., adding a test item for a better reliability, revised distractors that do not attract any respondents, etc.). In addition, in order to avoid a practice effect, or "the effect of having taken a given test on the results of subsequent administrations of the same or similar tests" (Brown & Hilferty, 1986, p. 64), in the case of Japanese, the revised version (Form B) was revised by making minor changes to the original one (Form A) (e.g., changing certain lexical items such as names, places, and actions, and changing the order of options).⁷

⁷ Brown, J.D. , & A.G. Hilferty. 1986. The effectiveness of teaching reduced forms for listening comprehension. *RELC Journal*, 17.2.59-70.

9. Final remarks: What factors have contributed to the successful initiation and maintenance of the project to date, and how will we go beyond the present stage?

Active faculty participation and leadership in any assessment activity at a given educational institution does not seem to come easily for two reasons. First, the participants and leaders may have to deal with political issues commonly associated with assessment activities. No matter how good your intentions are and how enthusiastic you may be, assessment activities are not always viewed by teachers as a positive activity (Brown, 1995).⁸ On the contrary, the moment teachers hear the term "assessment" or "program evaluation" as a planned institutional activity, many of them may feel threatened or even get angry because they interpret such activities as "bureaucratic evil" demanding stricter teacher evaluation, loss of academic freedom, and/or increased workload.

Having said this, I would like to conclude this report by discussing several critical factors that have allowed us to initiate and maintain this assessment activity to date, and by explaining what we will need to do in order to keep it going in the future. First of all, it is important to remember that this project was initially funded by the 2002-03 University of Hawaii Assessment Fund. Without this funding, we probably would not have been able to start the project. Therefore, if the university considers assessment to be an important and regular part of our teaching activities, it would help if they would keep providing this kind of money for future assessment activities. Also, I hope the university or college will invest more in existing assessment activities like this one. For example, the funding for the present assessment activity lasted only one year (ended in July 2003), and since then, we have had to develop this activity without any further funding from the university or college.

Second, the initiation and continuation of the present project has been possible because our deans and chairs within the college have been so willing to get involved in the project. For example, after we had identified learning outcomes and were ready to develop instruments as the next stage of

⁸ Brown, J. D. 1995. Language program evaluation: Problems and solutions. In W. Grabe (Ed.), *Annual Review of Applied Linguistics*, XV (pp. 227-248). Cambridge: Cambridge University Press.

the project, our Associate Dean consented to serve as the chair of the LLL assessment group. Without her involvement, I fear that we might have had more difficulty keeping faculty members involved in the project. Already inundated with teaching, research, committee assignments, etc, most, if not all, faculty members would not be anxious to do this "extra" work. I am very much hoping for the continued involvement and support of our deans and department chairs to help keep this project alive.

Third, in the case of Japanese, the Japanese section heads assigned each year were both supportive of this project and helpful. For example, when we needed to seek volunteer test writers and participating teachers, the section head wrote email messages to all Japanese 201 and 202 teachers and encouraged them to participate. I am hoping the Japanese section head and the heads of all language programs in the coming years will continue to understand the project and be supportive of the activities involved.

Fourth, an acceptable number of 201 and 202 teachers actually volunteered to participate in the first round of pilot testing. Without their willingness to cooperate and participate in the testing, no assessment activity would be possible. I believe as long as we do our work well and teachers find that participating in the assessment activities is a positive experience, they will continue to cooperate. As a related issue, we need to discuss how best to incorporate a college-wide assessment activity like this into the existing second language curricula. For example, through the first round of piloting in Japanese, I have learned that testing time (in the present case, only 10 minutes!) is a critical factor in teachers' deciding whether or not they will let their students participate. With the present Japanese curricula, some 201 and 202 teachers have expressed hesitation because they say they cannot afford to spend even 10 minutes for an "extra" activity. If we want this kind of assessment activity to be expanded (e.g., 30 minutes instead of 10 minutes) and institutionalized, we need to establish some

kind of systematic procedure so the teachers will not feel that program evaluation activities are not part of their regular teaching duties.

Fifth, someone in any assessment team must be willing and able to compile and analyze the obtained data and report the results. Without faculty members who can do these tasks, this type of assessment activity would be difficult. For the 2003-04 pilot assessment, I was willing to do the work. It was not an unreasonable amount of additional service to the college, and in fact, I enjoyed the work most of the time. However, we need to keep in mind that, at this stage of the present project, we are only focusing on *one* of the eight learning outcomes with a small number of participants. The assessment activity documented in this report is only a start. If this project is meant to go beyond the current pilot stage and expand to an "acceptable" level as a sound assessment activity for our two-year Hawaiian/second language programs (including additional assessment tools to measure more outcomes and more participants in more languages), I would probably need an "assessment staff", made up of maybe a group of graduate students who are trained in testing, so that I can effectively manage and analyze the *large* amount of data (I hope!) and prepare reports to give feedback to individual language programs.

At this point, I am still hopeful that this project will continue in the coming years. Like any other educational activity, it is very hard to start a new activity, but collapse or abandonment of that activity can come quite easily unless we keep working to maintain it. I think our project has an excellent start, and if we keep improving our work, more teachers will recognize the value and usefulness of this curriculum activity.

Appendix 2: Survey results concerning learning outcomes

Table 2. Recommended learning outcomes for LISTENING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items

Item ID	Description	ALL (N=88)			EALL (n=26)			HIPLL (n=31)			LLEA (n=31)		
		M	SD	Rank	M	SD	Rank	M	SD	Rank	M	SD	Rank
LIS1	Understand conversations about everyday experiences	4.48	0.62	1	4.62	0.55	1	4.65	0.47	1	4.26	0.61	3
LIS2	Understand factual content of paragraph-length descriptions/narratives on familiar topics	4.25	0.65	6	4.27	0.58	6	4.33	0.64	15	4.19	0.63	6
LIS3	Understand <i>abstract</i> content of paragraph-length descriptions/narratives on familiar topics	3.43	1.06	41	2.87	1.21	36	3.84	0.83	39	3.55	0.85	40
LIS4	Get main ideas from every day TV news/reports of several paragraphs with visual support	3.88	1.07	23	3.15	1.30	31	4.35	0.69	14	4.07	0.76	14
LIS5	Figure out the meaning of unknown vocabulary in context through guessing strategies	4.03	0.72	12	4.02	0.84	8	4.23	0.54	22	3.87	0.70	24
LIS6	Get main ideas of movies, plays, & orally-read stories	3.76	1.10	27	3.48	1.27	23	4.13	0.89	30	3.61	0.92	35

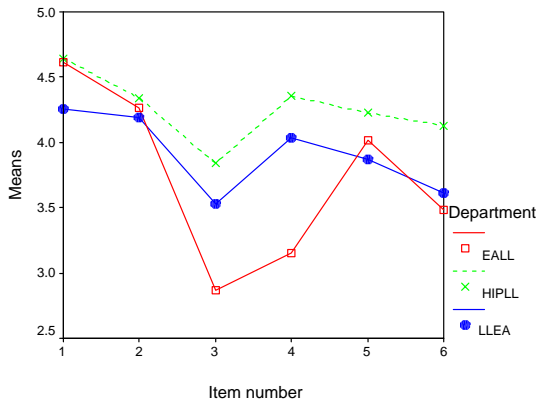


Figure 4: Learning outcomes for Listening

Table 3. Recommended learning outcomes for SPEAKING across departments/programs:
Summary of means, standard deviations, and rankings within all 42 items

Item #	Description	ALL (N=88)			EALL (n=26)			HIPLL (n=31)			LLEA (n=31)		
		M	SD	Rank	M	SD	Rank	M	SD	Rank	M	SD	Rank
SPK1	Perform a variety of "real-life" tasks in common social & transactional situations	4.42	0.76	3	4.62	0.72	2	4.48	0.74	5	4.23	0.65	4
SPK2	Sustain conversations/ interviews about self, family, experiences, interests, & preferences.	4.35	0.78	4	4.50	0.62	3	4.40	0.83	12	4.19	0.75	7
SPK3	Express opinions & feelings about passages/ stories read for class.	3.76	0.95	28	3.40	0.96	26	4.23	0.82	22	3.65	0.85	34
SPK4	Describe/narrate past & present experiences in paragraph-level discourse	3.98	0.93	17	3.88	0.83	15	4.32	0.77	16	3.77	0.99	28
SPK5	Describe/narrate future plans in paragraphs	4.02	0.94	13	3.92	0.90	13	4.26	0.83	20	3.90	0.98	22
SPK6	Give an oral presentation in paragraphs	3.96	1.11	18	3.50	1.28	21	4.32	0.91	16	4.00	0.94	16
SPK7	Perform short skits with some details about characters and scenes.	4.12	0.94	11	3.77	0.99	18	4.52	0.74	3	4.10	0.84	11
SPK8	Conduct interviews to gather info. for research on a topic of interest.	3.49	1.17	38	2.71	1.14	38	4.03	0.95	36	3.64	1.03	33
SPK9	Use language <i>creatively</i> by combining a variety of learned vocabulary expressions, & structures.	3.92	1.00	22	3.98	0.89	10	4.23	0.86	22	3.58	1.08	38
SPK 10	Initiate, sustain, & close a conversation using various communication strategies.	3.96	1.03	19	3.85	1.11	16	4.32	0.81	16	3.74	1.00	29
SPK 11	Successfully negotiate through simple misunderstandings	4.01	0.96	15	4.00	0.98	9	4.10	0.86	34	3.98	0.92	18
SPK 12	Seek and gather info. to solve problems	3.78	0.96	26	3.42	1.09	25	4.00	0.83	38	3.89	0.86	23
SPK 13	Use appropriate registers for various formal & informal situations.	3.63	1.06	34	3.46	1.16	24	3.74	1.00	41	3.74	0.93	29

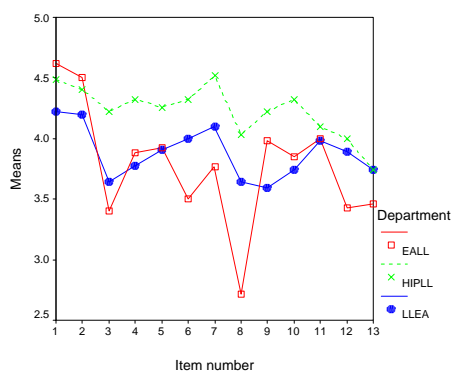


Figure 5: Learning outcomes for Speaking

Table 4. Recommended learning outcomes for READING across departments/programs: Summary of means, standard deviations, and rankings within all 42 items

Item #	Description	ALL (N=88)			EALL (n=26)			HIPLL (n=31)			LLEA (n=31)		
		M	SD	Rank	M	SD	Rank	M	SD	Rank	M	SD	Rank
RE1	Understand fully paragraph-length texts dealing with personal & social needs, i.e., personal letters, messages, & memos	4.26	0.85	5	4.29	0.96	5	4.48	0.70	6	4.08	0.75	13
RE2	Get main ideas from authentic everyday practical materials written entirely in the target language	4.43	0.65	2	4.38	0.60	4	4.45	0.65	9	4.52	0.61	1
RE3	Identify key facts & some supporting details in newspaper articles & short stories	3.74	1.23	29	2.67	1.25	39	4.26	0.97	20	4.23	0.66	4
RE4	Provide accurate, concise summary in English of a short reading passage	3.98	1.09	16	3.50	1.36	21	4.57	0.55	2	3.92	0.85	21
RE5	Form inter-textual connections between several readings or real-life events	3.49	1.11	39	2.60	1.17	40	4.13	0.74	31	3.60	0.79	37
RE6	Infer meaning from clues in a text or by 'reading between the lines'	3.45	1.11	40	2.86	1.37	34	3.97	0.81	38	3.47	0.84	41
RE7	Translate short passages written in the target language into clear & natural English	3.71	1.12	31	3.23	1.39	30	4.19	0.81	25	3.69	0.82	31
RE8	Figure out the meaning of unknown vocabulary in context through contextual guessing strategies'	3.93	0.84	20	3.83	1.10	17	4.16	0.67	29	3.87	0.61	24
RE9	Read texts written in the target language critically'	3.19	1.19	42	2.42	1.25	42	3.74	0.97	42	3.40	0.87	42

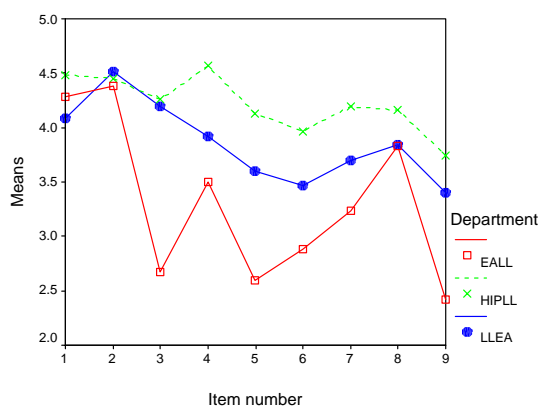


Figure 6: Learning outcomes for Reading

Table 5. Recommended learning outcomes for WRITING across departments/programs: Summary of means, standard deviations, and rankings within all 42 items

Item #	Description	ALL (N=88)			EALL (n=26)			HIPLL (n=31)			LLEA (n=31)		
		M	SD	Rank	M	SD	Rank	M	SD	Rank	M	SD	Rank
WR1	Meet practical writing needs and social demands by writing paragraph-length personal letters, messages, applications, & journals	4.25	0.96	6	4.13	1.24	7	4.47	0.71	8	4.23	0.69	5
WR2	Write paragraph-length <i>summaries</i> in the target language of reviewed reading or video material	3.79	1.16	24	2.98	1.31	33	4.42	0.70	11	3.94	0.86	19
WE3	Write paragraph-length <i>opinions</i> in the target language about reviewed reading or video material	3.68	1.12	32	2.88	1.26	35	4.19	0.77	25	3.94	0.75	19
WR4	Write narratives/descriptions of personal experiences, interests, and plans in several paragraphs	4.19	1.00	8	3.71	1.22	19	4.50	0.71	4	4.40	0.64	2
WR5	Write a report about the target culture in several paragraphs	3.63	1.23	35	2.60	1.23	41	4.17	0.88	27	4.10	0.80	12
WR6	Create short stories or skits with some details about characters and scenes in several paragraphs	3.56	1.23	36	2.73	1.38	37	4.10	1.06	33	3.79	0.70	27
WR7	Use language creatively by combining a variety of learned vocabulary, expressions, and structures	4.01	0.97	14	3.69	1.08	20	4.40	0.79	12	4.03	0.73	15
WR8	Use appropriate registers for formal and informal writing	3.54	1.12	37	3.31	1.33	28	3.81	0.88	40	3.58	0.99	39

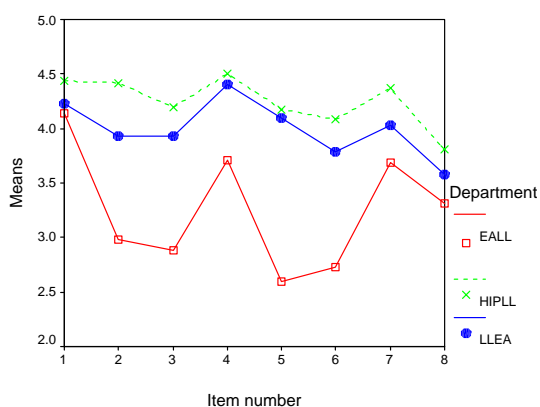


Figure 7: Learning outcomes for Writing

Table 6. Recommended learning outcomes for CULTURE across departments/programs: Summary of means, standard deviations, and rankings within all 42 items

Item #	Description	ALL (N=88)			EALL (n=26)			HIPLL (n=31)			LLEA (n=31)		
		M	SD	Rank	M	SD	Rank	M	SD	Rank	M	SD	Rank
CL1	Demonstrate understanding of holidays and traditions celebrated in the target culture	4.17	0.88	9	3.96	0.92	11	4.43	0.79	10	4.16	0.79	9
CL2	Demonstrate knowledge of traditional and contemporary sports, music, arts of the target culture	3.72	0.93	30	3.02	0.80	32	4.13	0.78	32	4.00	0.71	16
CL3	Demonstrate understanding of how culture is embedded in the use of target language	3.92	1.03	21	3.94	1.04	12	4.29	0.80	19	3.61	1.05	36
CL4	Identify major similarities and differences in social practices between their own culture and the target culture	4.16	0.91	10	3.92	0.86	14	4.48	0.66	7	4.13	0.96	10
CL5	Identify the ways the native and target cultures affect or influence each other	3.79	0.98	25	3.38	1.06	27	4.17	0.77	27	3.81	0.92	26
CL6	Understand the bi-directional nature of cultural sharing	3.65	1.02	33	3.25	1.20	29	4.03	0.74	35	3.68	0.91	32

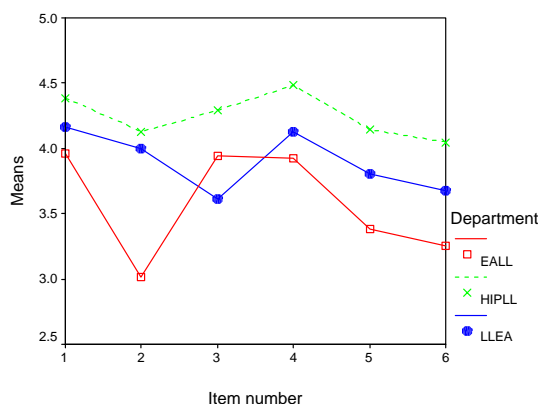


Figure 8: Learning outcomes for Culture

Appendix 3: Survey results concerning assessment tools

Table 7. Possible assessment tools for measuring LISTENING skills⁹

Item #	Description	ALL (N=80)			EALL (n=24)			HIPLL (n=27)			LLEA (n=29)		
		Fq	%	Rank	Fq	%	Rank	Fq	%	Rank	Fq	%	Rank
LIS1	Available Oral Proficiency Interview test (i.e., ACTFL OPI)	52	65.0	2	12	50	4	19	70.4	2	21	72.4	1
LIS2	Available Simulated Oral Proficiency Interview test (SOPI)	40	50.0	3	9	37.5	3	12	44.4	5	19	65.5	2
LIS3	Faculty-made oral achievement test embedded in the final exam	60	75.0	1	20	83.3	1	24	88.9	1	16	55.2	3
LIS6	Faculty-made paper-and-pencil achievement test embedded in the final exam	40	50.0	3	14	58.5	2	15	55.6	4	11	37.9	5
LIS7	Faculty-made computer-assisted achievement test	26	32.5	6	7	29.2	5	8	29.6	6	11	37.9	5
LIS8	Available paper-and-pencil proficiency test (Specify_____)	13	16.3	7	3	12.5	6	5	18.5	7	5	17.2	7
LIS9	Available computer-assisted proficiency test (Specify_____)	4	5.0	8	0	0	8	2	7.4	8	2	6.9	8
LIS10	Self-assessment questionnaire	36	45.0	5	7	29.2	5	17	6.3	3	12	41.4	4
LIS12	Other instrument (specify)	2	2.5	9	1	4.2	7	0	0	9	1	3.4	9

Table 8. Possible assessment tools for measuring SPEAKING skills

Item #	Description	ALL (N=80)			EALL (n=24)			HIPLL (n=27)			LLEA (n=29)		
		Fq	%	Rank	Fq	%	Rank	Fq	%	Rank	Fq	%	Rank
SP1	Available Oral Proficiency Interview test (i.e., ACTFL OPI)	53	66.3	3	13	54.2	3	19	70.4	3	21	72.4	1
Sp2	Available Simulated Oral Proficiency Interview test (SOPI)	38	47.5	5	10	41.7	4	11	40.7	6	17	58.6	3
Sp3	Faculty-made oral achievement test embedded in the final exam	62	77.5	1	23	95.8	1	25	92.6	1	14	48.3	4
Sp4	Write a report on a topic of interest & present it in front of an audience	50	62.5	4	14	58.3	2	19	70.4	3	17	58.6	3
Sp5	Write a short skit and perform it in front of an audience	55	68.8	2	13	54.2	23	23	85.2	2	19	65.5	2
Sp10	Self-assessment questionnaire	37	46.3	6	7	29.2	18	18	66.7	5	12	41.4	5
Sp11	Portfolio assessment	23	28.8	7	4	16.7	11	11	40.7	6	8	27.6	6
Sp12	Other instrument (specify_____)	4	5.0	8	2	8.3	6	0	0	7	2	6.9	7

⁹ In Tables 7-11, opinions of classical Greek and Latin instructors are not included. In addition, eight faculty members who left the entire section blank have been eliminated.

Table 9. Possible assessment tools for measuring READING skills

Item #	Description	ALL (N=80)			EALL (n=24)			HIPLL (n=27)			LLEA (n=29)		
		Fq	%	Rank	Fq	%	Rank	Fq	%	Rank	Fq	%	Rank
RE6	Faculty-made paper-and-pencil achievement test embedded in the final exam	61	76.3	1	18	75.0	1	20	74.1	1	19	65.5	1
RE7	Faculty-made computer-assisted achievement test	45	56.3	2	10	41.7	2	13	48.1	3	18	62.1	2
RE8	Available paper-and-pencil proficiency test (Specify_____)	14	17.5	5	2	8.3	6	6	22.2	5	6	20.7	5
RE9	Available computer-assisted proficiency test (Specify_____)	12	15.0	6	4	16.7	5	4	14.8	6	4	13.8	6
RE10	Self-assessment questionnaire	40	50.0	3	7	29.2	3	19	70.4	2	10	34.5	4
RE11	Portfolio assessment	28	35.0	4	5	20.8	4	12	44.4	4	11	37.9	3
RE12	Other instrument (specify_____)	2	2.5	7	0	0	7	0	0	7	2	6.9	7

Table 10. Possible assessment tools for measuring WRITING skills

Item #	Description	ALL (N=80)			EALL (n=24)			HIPLL (n=27)			LLEA (n=29)		
		Fq	%	Rank	Fq	%	Rank	Fq	%	Rank	Fq	%	Rank
WR4	Write a report on a topic of interest and present it in front of an audience	47	58.8	2	14	58.3	2	20	74.1	2	13	44.8	5
WR5	Write a short skit and perform it in front of an audience	43	53.8	3	11	45.8	3	18	66.7	4	14	48.3	3
WR6	Faculty-made paper-and-pencil achievement test embedded in the final exam	55	68.8	1	19	79.2	1	19	70.4	3	17	58.6	2
WR7	Faculty-made computer-assisted achievement test	35	43.8	6	9	37.5	4	12	44.4	6	14	48.3	3
WR8	Available paper-and-pencil proficiency test (Specify_____)	12	15.0	7	1	4.2	7	5	18.5	7	6	20.7	7
WR9	Available computer-assisted proficiency test (Specify_____)	7	8.8	8	1	4.2	7	3	11.1	8	3	10.3	8
WR10	Self-assessment questionnaire	39	48.8	5	7	29.3	6	21	77.8	1	11	37.9	6
WR11	Portfolio assessment	40	50.0	4	8	33.3	5	14	51.9	5	18	62.1	1
WR12	Other instrument (specify_____)	3	3.8	9	0	0	9	0	0	9	3	10.3	8

Table 11. Possible assessment tools for measuring CULTURAL KNOWLEDGE

Item #	Description	ALL (N=80)			EALL (n=24)			HIPLL (n=27)			LLEA (n=29)		
		Fq	%	Rank	Fq	%	Rank	Fq	%	Rank	Fq	%	Rank
CL1	Available Oral Proficiency Interview test (i.e., ACTFL OPI)	17	21.3	8	2	8.3	8	7	25.9	7	8	27.6	8
CL2	Available Simulated Oral Proficiency Interview test (SOPI)	13	16.3	9	1	4.2	9	5	18.5	9	7	24.1	9
CL3	Faculty-made oral achievement test embedded in the final exam	34	42.5	4	11	45.8	1	13	48.1	4	10	34.5	7
CL4	Write a report on a topic of interest & present it in front of an audience	42	52.5	1	9	37.5	3	15	55.1	2	18	62.1	1
CL5	Write a short skit and perform it in front of an audience	39	48.8	2	10	41.7	2	17	63.0	1	12	41.4	3
CL6	Faculty-made paper-and-pencil achievement test embedded in the final exam	33	41.3	5	9	37.5	3	9	33.3	6	11	37.9	5
CL7	Faculty-made computer-assisted achievement test	28	35.0	7	6	25.0	6	6	22.2	8	12	41.4	3
CL10	Self-assessment questionnaire	36	45.0	3	6	29.2	5	14	51.9	3	11	37.9	5
CL11	Portfolio assessment	33	41.3	5	7	25.0	6	11	40.7	5	16	55.2	2
CL12	Other instrument (specify)	3	3.8	10	1	4.2	9	0	0	10	2	6.9	10

Appendix 4: Survey results concerning Assessment procedures

Table 12. Assessment procedures 1 (When should the assessment be conducted?)¹⁰

When should the assessment be conducted?	All		EALL		HIPLL		LLEA	
	Fq.	%	Fq.	%	Fq.	%	Fq.	%
At the end of the fourth semester course	22	23.9	5	19.2	4	12.9	13	37.1
At the end of the second and fourth semester course	51	55.4	14	53.8	21	67.7	16	45.7
At the end of every semester	8	8.7	3	11.5	4	12.9	1	2.9
Other	3	3.3	0	0	1	3.2	2	5.7
No response	8	8.7	4	15.3	1	3.2	3	8.6
Total	92	100	26	100	31	100	35	100

Table 13. Assessment procedures 2 (How often should we assess?)

How often should we assess?	All		EALL		HIPLL		LLEA	
	Fq.	%	Fq.	%	Fq.	%	Fq.	%
Every semester	39	42.4	6	23.1	17	54.8	16	45.7
Every year	25	27.1	7	26.9	6	19.4	12	34.3
Every two years	19	20.7	9	34.6	7	22.6	3	8.6
No response	9	9.8	4	15.3	1	3.2	4	11.4
Total	92	100	26	100	31	100	35	100

Table 14. Assessment procedures 3 (Who should be assessed?)

Who should be assessed?	All		EALL		HIPLL		LLEA	
	Fq.	%	Fq.	%	Fq.	%	Fq.	%
All target students in all languages	55	59.8	9	3.5	24	77.4	22	62.9
All target students in a few languages from each department	1	1.1	1	3.8	0	0	0	0
A randomly selected "manageable number" of target students in all languages	25	27.2	11	4.2	4	12.9	10	28.6
A randomly selected "manageable number" of target students in a few languages from each department	2	2.2	0	0	2	6.5	0	0
Other	3	3.3	1	3.8	1	3.2	1	2.8
No response	6	6.5	4	15.4	0	0	2	5.7
Total	92	100	26	100	31	100	35	100

Table 15. Assessment procedures 3 (Who should do the assessment?)

Who should do the assessment?	All		EALL		HIPLL		LLEA	
	Fq.	%	Fq.	%	Fq.	%	Fq.	%
All faculty members as part of their work	30	32.6	5	19.2	13	41.9	12	34.3
One faculty member from each department designated as an assessment coordinator with a possible course release	14	15.2	4	15.4	4	12.9	6	17.1
Two faculty members from each department designated as assessment coordinators with a possible course release	26	28.3	9	34.6	7	22	10	28.6
External evaluators hired from outside campus	5	5.4	2	7.7	2	6.5	1	2.8
Other ¹¹	10	10.9	2	7.7	5		3	8.6
No response	7	7.6	4	15.4	0	0	3	8.6
Total	92	100	26	100	31	100	35	100

¹⁰ In Tables 14-17, opinions of classical Greek and Latin instructors are included.

¹¹ Departmental faculty assessment committee with several members, all faculty members with an assessment coordinator for each department, combination of faculty and external evaluator

Appendix 5: Language Requirement Assessment Planning Questionnaire

I. Background information

Your department (circle one) : EALL, HIPLL, LLEA Position : _____

Language(s) you teach this semester and/or taught before at UH : _____

II. Which of the following *fourth-semester* courses, are you giving your opinions for (mark all apply)?

- | | | |
|--|---|---|
| <input type="checkbox"/> Ancient Greek 202 | <input type="checkbox"/> Hawaiian 202 | <input type="checkbox"/> Latin 202 |
| <input type="checkbox"/> Arabic 202 | <input type="checkbox"/> Hindi 202 | <input type="checkbox"/> Maori 202 |
| <input type="checkbox"/> Balinese 202 | <input type="checkbox"/> Ilokano202 | <input type="checkbox"/> Russian 202 |
| <input type="checkbox"/> Cambodian 202 | <input type="checkbox"/> Indonesian 202 | <input type="checkbox"/> Samoan 202 |
| <input type="checkbox"/> Chinese 202 | <input type="checkbox"/> Italian 202 | <input type="checkbox"/> Sanskrit 202 |
| <input type="checkbox"/> Chinese 212 | <input type="checkbox"/> Japanese 202 | <input type="checkbox"/> Spanish 202 |
| <input type="checkbox"/> Filipino 202 | <input type="checkbox"/> Japanese 212 | <input type="checkbox"/> Tahitian 202 |
| <input type="checkbox"/> French 202 | <input type="checkbox"/> Korean 202 | <input type="checkbox"/> Thai 202 |
| <input type="checkbox"/> German 202 | <input type="checkbox"/> Korean 212 | <input type="checkbox"/> Vietnamese 202 |

III. Identifying learning outcomes: Please indicate to what degree you agree with the following statements as learning outcomes for students who complete the course(s) you marked above using the numerical scale 1-5 defined below.

1	2	3	4	5
Strongly disagree	Disagree	No opinion	Agree	Strongly agree

Learning Outcomes of Listening SkillsWhen the target language is delivered in standard dialect, **the students will be able to:**

- | | |
|-----------|---|
| 1 2 3 4 5 | Understand conversations about everyday experiences (e.g., school, work, interests, preferences) |
| 1 2 3 4 5 | Understand <i>factual</i> content of paragraph-length descriptions/narratives on familiar topics (e.g., recorded telephone instructions, announcements in public areas) |
| 1 2 3 4 5 | Understand <i>abstract</i> content of paragraph-length descriptions/narratives on familiar topics (e.g., a speech about a social issue) |
| 1 2 3 4 5 | Get main ideas from every day TV news/reports of several paragraphs when accompanied by visual support (e.g., weather report or sport news with information in chart) |
| 1 2 3 4 5 | Figure out the meaning of unknown vocabulary in context through guessing strategies |
| 1 2 3 4 5 | Get main ideas of movies, plays, and orally-read stories |

List here any additional learning outcome(s) you would suggest for listening skills:

Learning Outcomes of Speaking Skills

The students will be able to:

- 1 2 3 4 5 Perform a variety of “real-life” tasks in common social and transactional situations (e.g., shopping, making hotel reservations)
- 1 2 3 4 5 Sustain conversations/interviews about self, family, experiences, interests, and preferences
- 1 2 3 4 5 Express opinions and feelings about passages/stories read for class
- 1 2 3 4 5 Describe and narrate past and present experiences in paragraph-level discourse (e.g., recent travel experience)
- 1 2 3 4 5 Describe and narrate future plans in paragraph-level discourse (e.g., plans after graduation)
- 1 2 3 4 5 Give an oral presentation or demonstration of several paragraphs on a topic of interest (e.g., demonstrating how to make/use something, story telling)
- 1 2 3 4 5 Perform short skits with some details about characters and scenes
- 1 2 3 4 5 Conduct interviews to gather information for research on a topic of interest
- 1 2 3 4 5 Use language *creatively* by combining a variety of learned vocabulary, expressions, and structures
- 1 2 3 4 5 Initiate, sustain, and close a conversation using various communication strategies
- 1 2 3 4 5 Successfully negotiate through simple misunderstandings (e.g., misunderstanding of an order in a restaurant)
- 1 2 3 4 5 Seek and gather information to solve problems
- 1 2 3 4 5 Use appropriate registers for various formal and informal situations

List here any additional learning outcome(s) you would suggest for speaking skills:

Learning Outcomes of Reading Skills
--

The students will be able to:

- | | |
|-----------|--|
| 1 2 3 4 5 | Understand fully paragraph-length texts dealing with personal and social needs such as personal letters, messages, and memos |
| 1 2 3 4 5 | Get main ideas from authentic everyday practical materials written entirely in the target language (e.g., menus, ads for products) |
| 1 2 3 4 5 | Identify key facts and some supporting details in newspaper articles and short stories |
| 1 2 3 4 5 | Provide accurate, concise summary in English of a short reading passage |
| 1 2 3 4 5 | Form inter-textual connections between several readings or real-life events |
| 1 2 3 4 5 | Infer meaning from clues in a text or by “reading between the lines” |
| 1 2 3 4 5 | Translate short passages written in the target language into clear and natural English |
| 1 2 3 4 5 | Figure out the meaning of unknown vocabulary in context through contextual guessing strategies |
| 1 2 3 4 5 | Read texts written in the target language critically |

List here any additional learning outcome(s) you would suggest for reading skills:

--

Learning Outcomes of Writing Skills
--

The students will be able to:

- | | |
|-----------|--|
| 1 2 3 4 5 | Meet practical writing needs and social demands by writing paragraph-length personal letters, messages, applications, and journals |
| 1 2 3 4 5 | Write paragraph-length <i>summaries</i> in the target language of reviewed reading or video material |
| 1 2 3 4 5 | Write paragraph-length <i>opinions</i> in the target language about reviewed reading or video material |
| 1 2 3 4 5 | Write narratives/descriptions of personal experiences, interests, and plans in several paragraphs |
| 1 2 3 4 5 | Write a report about the target culture in several paragraphs |

- 1 2 3 4 5 Create short stories or skits with some details about characters and scenes in several paragraphs
- 1 2 3 4 5 Use language creatively by combining a variety of learned vocabulary, expressions, and structures
- 1 2 3 4 5 Use appropriate registers for formal and informal writing

List here any additional learning outcome(s) you would suggest for writing skills:

Learning Outcomes of Cultural Understanding

The students will be able to:

- 1 2 3 4 5 Demonstrate understanding of holidays and traditions celebrated in the target culture
- 1 2 3 4 5 Demonstrate knowledge of traditional and contemporary sports, music, arts of the target culture
- 1 2 3 4 5 Demonstrate understanding of how culture is embedded in the use of target language
- 1 2 3 4 5 Identify major similarities and differences in social practices between their own culture and the target culture
- 1 2 3 4 5 Identify the ways the native and target cultures affect or influence each other
- 1 2 3 4 5 Understand the bi-directional nature of cultural sharing

List here any additional learning outcome(s) you would suggest for cultural understanding:

IV. Identifying assessment tools: Check *all* assessment instruments you think could be adopted in order to measure the identified students' learning outcomes. For example, if you think the use of an available oral proficiency test could be used to measure the target students' listening and speaking skills, make check-marks as follows.

	Listening	Speaking	Reading	Writing	Culture
Available Oral Proficiency interview (i.e., ACTFL OPI test)	√	√	NA	NA	

	Listening	Speaking	Reading	Writing	Culture
Available Oral Proficiency Interview test (i.e., ACTFL OPI test) ¹²			NA	NA	
Available Simulated Oral Proficiency Interview test (SOPI) ¹³			NA	NA	
Faculty-made oral achievement test embedded in the final exam			NA	NA	
Write a report on a topic of interest and present it in front of an audience	NA		NA		
Write a short skit and perform it in front of an audience	NA		NA		
Faculty-made paper-and-pencil achievement test embedded in the final exam		NA			
Faculty-made computer-assisted achievement test		NA			
Available paper-and-pencil proficiency test (Specify _____)		NA			NA
Available computer-assisted proficiency test (Specify _____)		NA			NA
Self-assessment questionnaire (asking students to evaluate their own knowledge and skills)					
Portfolio assessment	NA				
Other instrument (specify)					
Other instrument (specify)					

Which of the following statements apply to you?

- () I am a certified OPI tester.
- () I was a certified OPI tester (certification expired in _____)
- () I have developed or used a proficiency test that may be useful for a two-year Hawaiian/second language program evaluation (explain _____)
- _____

¹² ACTFL OPI testing is available in such languages as Arabic, Cambodian, Chinese, French, German, Greek, Hindi, Indonesian, Italian, Japanese, Korean, Russian, Spanish, Thai, and Vietnamese.

¹³ The SOPI is a tape-mediated performance test designed to elicit samples that are similar to samples that would be elicited through an OPI. It is available in Arabic, Chinese, French, German, Japanese, Russian, and Spanish.

V. Identifying assessment procedures: In conducting assessment using the instruments of your choice, which of the following procedures would be desirable?

(Mark only *one* choice for each question)

1. When should the assessment be conducted?

- At the end of the fourth semester course (e.g., Spanish 202)
- At the end of the second AND fourth semester courses (e.g., Spanish 102 and 202)
- Other. Explain: _____

2. How often should we assess?

- Every semester
- Every year
- Every two-years
- Other: Explain _____

3. Who should be assessed?

- All target students in all languages
- All target students in a few languages from each department
- A randomly selected “manageable number” of target students in all languages
- A randomly selected “manageable number” of target students in a few languages from each department
- Other: Explain _____

4. Who should do the assessment?

- All faculty members as part of their work
- One faculty member from each department designated as an assessment coordinator with a possible course release
- Two faculty members from each department designated as assessment coordinators with a possible course release
- External evaluators hired from outside campus
- Other: _____

Thank you for answering all the questions!!!