

Issues in Placement Survey

James Dean Brown, Thom Hudson, Martyn Clark
National Foreign Language Resource Center
Department of Second Language Studies
University of Hawai'i at Mānoa

This paper reports on the results of a nationwide survey of 169 modern foreign language departments in the United States. The survey was conducted as the first part of a project to develop a manual of best practices in placement testing to be used in conjunction with a summer institute on placement testing. In addition to presenting raw data from the survey, five questions are addressed: (a) What skills are most often assessed for placement purposes, (b) Do programs tend to use commercially-produced or self-produced materials, (c) How do language programs validate their placement instruments, (d) How involved are teachers in the placement process, and (e) What issues would language programs like to see addressed in a placement handbook? Results and implications for the development of the placement testing handbook are discussed.

Introduction

As long as there are classification or selection decisions to be made, there will probably always be a need for some kind of placement testing. This can range from informal –such as the very informal process observed by one of the authors in which new students were met in the hallway and asked three questions; those who could answer all three were put in the upper level class and those who couldn't were put in the lower level class – to very extensive and thorough procedures involving a complete review of the students' language background and a current assessment of their language ability. Despite the widespread use of placement tests in modern language programs, however, there is a lack of research into what might be considered best practices in placement testing.

The survey reported here is part of a larger project targeting placement practices in language programs. The overall goal of the project is to develop a manual of best practices in placement testing that would be of use to those people involved in the placement process. Although in one sense good testing practices transcend specific testing purposes and any training in test development would be useful for placement testing, we wanted to tailor the manual as much as possible to concerns and issues common in assessing language for placement decisions. In order to begin the tailoring process, it was necessary to first get some information on what programs are currently doing for placement. A survey seemed like the most efficient way to accomplish this initial goal, and this paper reports on the development and results of that survey.

We recognized from the outset that we did not have the resources nor the need to produce a survey of placement practices that would allow us to state with confidence the precise percentage of programs using one procedure versus another or state definitively that statistically significantly more programs assessed reading than speaking. The types of test development procedures expounded in the testing literature would be useful across a range of placement procedures and good testing practices are at the heart of good placement testing practices. What we did need, however, was a kind of "slice of life" view of what was happening in a fair number of programs so that we could develop a manual that would discuss those good testing practices within the context of a recognizable (to the reader) placement process.

With this in goal mind, five research questions were posed:

1. What skills are most often assessed for placement purposes?
2. Do programs tend to use commercially-produced or self-produced testing materials?
3. How do language programs validate their placement instruments?
4. How involved are teachers in the placement process?
5. What issues would language programs like to see addressed in a placement handbook?

It was hoped that answers to these general questions would give a good starting point for thinking about how to develop the placement testing manual. If programs tended to focus on productive skills such as writing and speaking, for example, then a thorough discussion of the use of raters and rating scales for assessment would be necessary. On the other hand, if most programs were using selected-response tests, then focusing only on performance type assessments would be doing them disservice. By the same token, it would make no sense to pitch the handbook to language testing professionals if the people responsible for placement testing in most programs are teachers. This is not to say that the manual should support the status quo no matter what that is, but

This study was made possible under grant from the U.S. Department of Education (CDEA 84.299, P229A020002). However, the contents do not necessarily represent the policy of the Department of Education, the NFLRC, or the University of Hawai'i, and one should not assume endorsement by the Federal Government.

rather than for the handbook to be useful, it must address concerns of real programs while recognizing their constraints. Discovering the placement concerns of a number of real programs would obviously be very helpful in planning the manual.

Methods

Participants

Sampling. To distribute the paper-based survey, a list of names of language department chairs was rented from the Modern Language Association. The entire list consisted of 2589 names. From the outset, we were not planning to have a truly representative survey. If we were, we would ideally have to sample from the entire universe of colleges in the US. Instead, we were mainly interested in getting an idea of the types of placement procedures used in various schools around the country. For this more limited purpose, we decided to shoot for approximately 200 total responses (either paper or electronic) as a goal. Initially assuming a 20% response rate, this meant that we would need to send questionnaires to approximately 1,000 programs in total. To conserve resources should the response rate prove higher than anticipated, we decided to perform a smaller mailing first with the option of a second mailing should we not reach our goal of 200 responses. This meant that two lists of 500 names each would be needed.

Choosing the 500 names for each mailing was done through a semi-random process in which every 5th name from the MLA list was chosen starting from a randomly generated starting point. This procedure was repeated to generate the second list as well. It should be noted that the list itself is not a random sample in that non-MLA members are excluded and because MLA membership is not necessarily strictly proportional to various geographical populations or program types. We intentionally did not keep track of to which programs questionnaires would be sent because that would create a great administrative burden with no real benefit. The reader is referred to Table 2 for information on which geographical locations are represented by survey respondents.

Characteristics of the responding institutions. Because our survey deals with language programs, it is a little unclear how much of the information is respondent specific (i.e., the opinions of the individual actually filling out the survey) and how much is institution specific. Of the 169 responses, 109 were from public schools while 58 were from private schools. Not surprisingly, Spanish, French, and German were the three most common languages. Table 1 shows all of the languages reported and Table 2 shows the response distribution by state.

Although the surveys were initially mailed to department chairs, a cover letter instructed the chairs to pass on the questionnaire to the person in the department most involved in placement testing. The self-reported job description of the respondents can be seen in Table 3, with teacher being the most common answer.

Table 1
Languages taught in various programs

Language	Programs
Spanish	120
French	90
German	70
Italian	35
Japanese	32
Russian	32
Chinese	31
Spanish for Spanish speakers	18
Portuguese	13
Hebrew	9
Arabic	8
Korean	6
Vietnamese	5
Swahili	3
Thai	3
Dutch	2
ESL	2
Filipino	2
Irish	2
ASL	1
Burmese	1
Farsi	1
Modern Greek	1
Hindi	1
Indonesian	1
Norwegian	1
Ojibwe	1
Polish	1
Swedish	1
Tagalog	1
Urdu	1

Materials

The questionnaire used in this study was developed over a period of several months through discussions among the authors. Several earlier surveys of language programs in general (Rhodes & Branaman, 1999; Walker & Li, 2003) and placement testing in particular (Wherritt & Cleary, 1990) provided initial models for the development of the current questionnaire. During the first few discussions, it became apparent that the survey would only be able to provide a general window into the placement process. To truly understand and evaluate the placement process, knowledge of the context in which placement takes place is important. This means that ideally one would have information about the program's curriculum, goals, students, etc. Unfortunately, there was no feasible way to collect this type of information on any kind of national level. Because so much of the placement process is program specific (i.e., geared towards placing a certain population of students into a certain sequence of courses), a major challenge was to develop the questionnaire in such a way that it was general enough to cover many different types of

Table 2
Geographical location of programs

State	Responses
California	22
New York	14
Illinois	12
Minnesota	8
Pennsylvania	8
Iowa	7
Massachusetts	7
Michigan	7
Ohio	6
Wisconsin	6
Tennessee	5
Hawaii	4
Texas	4
Virginia	4
Connecticut	3
Florida	3
Georgia	3
Indiana	3
Kentucky	3
Louisiana	3
Missouri	3
New Jersey	3
Oregon	3
Washington, D.C.	2
Kansas	2
Nevada	2
Oklahoma	2
Rhode Island	2
Vermont	2
Arizona	1
Arkansas	1
Colorado	1
Delaware	1
Mississippi	1
New Hampshire	1
New Mexico	1
North Carolina	1
Puerto Rico	1
South Carolina	1
Washington	1

Table 3
Self-reported job description

Position	<i>N</i>
Language teacher	76
Foreign language department chair	57
School foreign language coordinator	16
Language resource center director	6
Testing coordinator	4
State foreign language coordinator	1
District foreign language coordinator	1

language programs while at the same time providing detailed information about the placement testing process.

Basic considerations. After several rounds of discussion, it was decided that the questionnaire would focus on six general areas with varying degrees of detail:

1. Basic program information. This is information about the location of the program, the program type, languages taught, number of students, and the types of courses offered. This information was collected to help with the categorizing and generalizing of results.

2. Language ability assessment. This section is designed to get information about the types of language ability assessment that occur in the program *other than* placement testing. This information was collected to give an idea of the continuity between placement procedures and other types of assessment in the program.

3. Placement test administration. This section asked questions about the details of placement test administration including how often the test is administered, who proctors the test, who scores the test, and how scores are transmitted to students. It was hoped that one use of this section would be to get an idea of how "hands-on" teachers are in the placement testing process.

4. Placement test content. This section dealt with the actual content of the placement test in terms of skills tested and the origin of the test materials.

5. Course placement procedures. Because placement test scores are not always the only part of the placement process, this section was designed to elicit information on how the final course placements were actually made.

6. Comments and contact information. Comments on what respondents would like to see in a placement manual was solicited. In order to facilitate potential follow-up sessions, contact information was collected from those participants willing to provide it.

For each of the basic sections, several questions were developed that would elicit the type of information we needed.

Determining the response format. While developing the specific questions for each section of the questionnaire, it became clear that a decision had to be made about how much leeway to allow the respondents. Survey questions can range from very tightly controlled to very open-ended, and there are advantages and disadvantages to each (Brown, 2001). There were essentially two competing factors, in our view. We wanted respondents to be able to give an information that they deemed necessary while at the same time constraining the types of answers that they gave to facilitate future analysis. We were also aware of the fact that as our target audience were people who tend to have many administrative things to juggle, we would need to cover a lot of ground in a very efficient fashion.

In the end, we opted for a very controlled set of questions (yes/no, multiple-choice) with an option of "other" for virtually every question. We also provided a space for comments at the end of each of the major sections of the survey to give respondents a chance to amplify their answers. This proved

Table 4
Survey responses by month

Type	March	April	May	June	July
Online	16	7	5	52	-
Mail ^a	21	56	8	2	3
Total	37	63	13	54	3

^a1st mailing March 17; 2nd mailing April 6

to be a decision not without some regret, however.

Developing the electronic version. Early on in the development process, we decided that the survey should also be available in electronic format on the World Wide Web. Once the paper-based version was ready, creating the web-based version was mainly a matter of translating the survey format into machine-readable format. Though we had investigated the use of commercially-available form generation software, we decided that a simple HTML page would be sufficient. By combining a simple HTML form hosted on a computer with a FileMaker Pro database, we were able to have the survey results be submitted directly into the database, thus simplifying the data collection process.

Procedures

The mailing. The mailing was conducted by one of the authors during the period of March to April 2004 through the University of Hawai'i post office. Each mailing piece consisted of a cover letter, a consent form, the questionnaire itself, and a business reply envelope. Assuming (perhaps optimistically) a 20% response rate, we decided to send 1,000 pieces split into two 500 piece mailings approximately two weeks apart. This way, if the response rate were much higher than anticipated, we could postpone or eliminate the second mailing. In the end, both 500 piece mailings were sent. The only difference between the two mailings was that the due date on the cover later was changed to reflect a slightly later date for the second mailing. Of the 1,000 pieces sent, approximately 10 surveys were returned as undeliverable because the addressee was no longer at that address or was deceased. Five surveys were returned as undeliverable as the address label had either fallen off or not been affixed initially.

At approximately the same time as the surveys were being mailed, we also placed announcements on several electronic LISTSERVs announcing the availability of an electronic version. The URL for the electronic version was also included in the cover letter of the mailing. The electronic version was accessible for approximately four months after which it was taken off line.

Return rate. Because the survey was available in both paper and electronic format and because even those people receiving a paper questionnaire had the option of completing it on-line, it is very difficult to judge the overall return rate. In all, 90 people returned the survey in the business reply envelope while 80 people completed the survey on-line. Table 4 shows the responses by month.

Table 5
Language ability assessment activities used in programs

Assessment Activity	Number of programs
Constructed-response tests	145
Oral proficiency interviews	131
Student presentations	131
Selected-response tests	120
Authentic activities	112
Translation exercises	80
Student portfolios	65
Student self-assessment	44
Extended writing or report*	5
Poems*	1
Debate*	1
Dictation*	1

*Write-in response

Results

Language Ability Assessment

Table 5 shows the breakdown of programs that reported using various types of assessment procedures to assess language ability in contexts *other than* placement situations. As can be seen from the table, programs reported using a rather wide range of assessment tools. As with all of the tables in this section, participants often marked more than one response while others left responses blank, thus the totals may be considerably more or less than the total number of programs responding.

Almost three-quarters of the respondents indicated that the teachers in their program were aware of the *Standards for Foreign Language Learning* when asked (*Yes* = 121, *No* = 42), but considerable fewer indicated that assessment practices in their program had changed because of this awareness (*Yes* = 77, *No* = 79). Use of proficiency scales was reported by a majority of programs (*Yes* = 105, *No* = 58) with the majority of scales being developed within the program either originally through discussions among program faculty (*n* = 34) or with reference to scales used in similar programs (*n* = 53). Other programs reported developing proficiency scales jointly with outside organizations (*n* = 15) or using completely externally developed scales (*n* = 13).

Placement test administration. A majority of programs reported using language placement tests when asked (*Yes* = 122, *No* = 42); of those programs, 87 reported using an locally developed test while 54 reported using a commercial or other externally developed test. (Note that some programs reported using both kinds). Details about test administration are shown in Table 6 through Table 11. Again note that many programs selected multiple answers on the questionnaire so the totals may exceed the number of programs reporting test use.

Table 6
Test frequency

Frequency	Responses
Only when needed	36
Several times a semester	30
Once a semester	20
Once a year to incoming students*	16
On demand/available on-line*	14
As needed*	5
Weekly*	1

*Write-in response

Table 7
Test registration process

Process	Responses
Mandatory participation	32
Department office	22
First-come, first-served	19
Admissions office	19
Testing office	13
Multimedia lab staff	9
Through appointment with advisor*	9
On-line*	7
Taken at orientation*	6
Test given during class time	5
Regionally delivered*	2

*Write-in response

Table 8
Test proctor

Proctor	Responses
Teachers	64
Testing office staff	21
Multimedia lab staff	16
Test delivered automatically/on-line*	13
Admissions office staff	11
Student help	11
Dept. Chair or staff*	5
Language coordinator*	1
LRC Director*	1
Placement director*	1

*Write-in response

Table 9
Test scoring

Scoring method	Responses
Teachers	66
Machine-scored	42
Scored automatically by computer*	14
Admissions office staff	9
Testing office staff	8
Multimedia lab staff	3
Student help	2
Dept. Chair or staff*	2

*Write-in response

Table 10
Communicating scores to students

Communication method	Responses
Teachers	52
Academic advisor	20
Scores are posted	20
Automatically generated/on-line*	14
Testing office staff	13
Admissions office staff	11
Dept. office staff*	8
1st year studies office*	4
Language coordinator*	2

*Write-in response

Table 11
Availability of retests

Times students can take test	Responses
Only once	69
Unlimited	27
Restest with special permission only	12
Not more than twice	9
Once per year*	2
Once per semester*	1
Once every three years*	1
Retests for suspicious scores*	1

*Write-in response

Placement test content

Tables 12 through 21 give some details about the contents and validation procedures used for both internal and external tests. Perhaps unsurprisingly, the oral skills and real-world skills were less often assessed than the traditional skills of reading, writing, and grammatical knowledge (see Table 12). One program also reported assessing cultural knowledge in addition to more linguistic skills.

Internally developed tests. Tables 13 through 18 refer to those tests developed locally by individual programs. Better

Table 12
Skills assessed for placement purposes

Skill	Responses
Reading	115
Grammar	109
Vocabulary	102
Writing	80
Listening	56
Speaking	42
Ability to complete a real world task	15
Cultural knowledge*	1
Personal interview*	1

*Write-in response

Table 13
Reasons for locally developing placement tests

Reason	Responses
To better control test contents	38
To better reflect ability range of students in program	31
Commercial tests unsuitable	31
No commercial test exists for language	20
Availability of funding/resources for development	19
As temporary measure only	8
To supplement external test	3
No cost/less expensive*	3
External assessments not well-known*	1
Create state-wide instrument for articulation*	1
Complement oral interview*	1

*Write-in response

Table 14
Responsibility for test creation

Test developer	Responses
Current/past teachers	77
Current/past administrator	5
Special committee	5
Language resource center	5
Individual test developer	4
Foreign language office*	1
Testing division*	1

*Write-in response

control over test content (including maximizing the match between student ability and test difficulty) and the unsuitability of commercially produced tests were the most often cited reasons for developing a test locally (Table 13). Three programs noted that internally developed tests incurred fewer costs than commercial tests.

As can be seen in Table 14, teachers bear most of the burden in developing tests at the program level, with test content coming from textbooks, course objectives, authentic and original material (Table 15). Two programs reported having the tests developed through the testing division or foreign language office. Perhaps because of their role in the test creation process, teachers' involvement in locally produced tests extends through all phases of the placement process, as seen in Tables 8 through 10.

Because teachers have other duties in addition to test creation, test revision seems to be a function of course content rather than natural test development per se (see Table 16 and 17). One program reported revising the cut scores each year, but not necessarily the test contents.

The validation process for internally developed tests (Table 18) included having current students take the test, matching the contents to course objectives, or review by language specialists. Some programs reported performing an item analysis, but those were in the minority. Some programs

Table 15
Primary source of test content

Source	Responses
Course objectives	40
Original materials	31
Course textbooks	28
Authentic materials	25
State/local content standards	11
MLA*	1
Proficiency guidelines*	1
Combination including national exams*	1
Past exams*	1

*Write-in response

Table 16
Test revision frequency

Frequency of test revision	Responses
When necessary	37
Once every several years	16
When conditions (personnel, financial) permit	11
Once or twice year	10
Never revised	10
With each new intake*	1
Only cut scores are revised*	1

*Write-in response

also reported that the review by language specialists was performed only when the test was first developed. One program reported that the test was not validated in the traditional sense, but that it seemed to be working fine.

Externally developed tests. Table 19 shows the number of respondents using various commercially produced tests. The Brigham Young University CAPE (Computer Adaptive Placement Exam) and the Advanced Placement (AP) subject tests were the most common commercially produced tests used by the respondents. Part of this is no doubt due to the

Table 17
Reason for revising test

Reason	Responses
To better reflect course contents	42
To address a deficiency in current test	28
To assess a skill not previously assessed	15
To make the test contents more timely	13
To lengthen or shorten the test	8
To prevent cheating	7
Demography of students*	1
More contextualization*	1
Improve validity, delivery, student-friendliness*	1
Upgrade/develop*	1
Prepare test for web delivery*	1

*Write-in response

Table 18
Validation method for internally developed tests

Validation method	Responses
Piloting on current students	57
Matching contents to course objectives	43
Review by language specialists	33
Item analysis	18
Not validated but seems to work*	1

*Write-in response

Table 19
Externally produced tests used by respondents

Test	Programs
AP subject test	38
CAPE (BYU)	21
Wisconsin Test	8
ACTFL OPI	7
CLEP	5
SAT II subject test	2
Michigan Test	1
SOPI	1
Japanese Proficiency Test	1
MLPA	1
AATG Test*	1
University of Oregon NFLRC STAMP*	1
College Board*	1
sras*	1

*Write-in response

availability of these tests for the most commonly taught languages. Widespread use in other programs and the consistency of the results were the most cited reasons for choosing the tests (Table 20), though some programs commented that the decision to use the test was out of their hands (i.e., a decision by the Dean) or that practical considerations, such as the lack of available staff for testing duties during the summer, were also considerations.

As with the internally developed tests, content review and

Table 20
Reason for choosing commercial test

Reason	Respondents
Widely used in other programs	21
Gives consistent results	18
Have always used it	8
Familiar to teachers	7
No other option	4
Ease of administration*	1
Cost*	1
Convenience*	1
Faculty not available in summer*	1
Dean's decision*	1

*Write-in response

Table 21
Validation method for externally developed test

Validation method	Responses
Review by instructors/department head	37
Piloting on current students	31
Matching contents to course objectives	7
Item analysis	2
Compared favorably to previous test*	1

*Write-in response

Table 22
Final course placement determination

Placement decision	Responses
Teacher's recommendation	64
Previous semesters of study	44
Advisor's recommendation	44
Student self-selection	40
Placement test scores	13
High school transcripts*	2
Combination of factors*	1

*Write-in response

piloting on current students were the most common ways of validating externally produced tests (Table 21). One program cited a high correlation obtained between a previously used test. Perhaps because of the inherent mismatch between a standardized test and an individual program's curriculum, matching the test content to course objectives was not used to the extent it was for locally produced tests.

Course placement

Regardless of whether or not a placement test is used, programs still must place their students into classes. Tables 22 through 24 highlight various parts of the placement process. Note that most respondents indicated a combination of procedures for any given facet of the process. It is interesting to note in Table 22 that very few programs placed students exclusively on the basis of test scores. Of course, the extent to which advisor and teacher recommendations are influenced by test scores is not seen. It is also important to note, as can be seen in Table 24, that only in a very small number of programs were course changes not allowed once instruction has begun. Some programs commented that changes of level were not allowed, but students were free to switch languages or drop language courses.

Comments and additional information

Of the 169 surveys received, 137 included contact information for the person who had filled out the survey while 32 people chose to remain anonymous. Additionally, 68 respondents were kind enough to include their thoughts and comments about both the placement process in general as well as give some thoughts on what the placement handbook should

Table 23
Determination of cut scores

Cut score setting method	Responses
Piloting on current students	59
Based on proficiency level	48
Test-maker's recommendation	31
No specific cut-scores*	5
Past experience*	3
Language faculty*	3
Colleague's recommendation*	2
Combination*	2
Students make final choice based on recommendation*	1
Tracking of placed students*	1

*Write-in response

Table 24
Class changes allowed after instruction commences

Course change policy	Responses
Changes allowed during add/drop period	112
Changes allowed only with special permission	37
Course changes not allowed	2
Change to other language only*	1
Changes allowed up to week six*	1
Class change only, no level change*	1
Change allowed if prerequisites met*	1

*Write-in response

cover. In general, comments could be classified into one of seven categories (with illustrative comments):

1. The philosophy of placement testing or other overarching concerns ("How passive knowledge or ability (reading, grammar rules, recognition of word meanings) sometimes gained over a period of years can be balanced with actual production in speaking and writing when determining placement level.")

2. The testing of a particular language skill ("Testing writing skills")

3. Appropriately defining levels of proficiency ("How to distinguish between students at 'high beginner' and intermediate level")

4. Administrative or practical issues ("How to find a reliable placement test that assesses all four skills and can be administered efficiently to a fairly large number of students.")

5. The placement of heritage students ("How to place heritage *speakers* with limited reading and writing proficiency")

6. On-line or computerized testing ("Specific information on web-based and/or other on-line testing, especially for levels beyond beginners")

7. General comments or encouragement about the survey ("I look forward to seeing that manual")

Table 25 shows the number of comments for each of the seven categories. Although some answers had elements of more than one category, the perceived main concern was used for classification purposes.

Table 25
Suggestions for the placement handbook

Category	Number of comments
Placement philosophy	15
Testing specific skills	14
Computerized placement	10
Practical issues	9
Heritage students	7
General comments	7
Defining levels	6

Discussion

Because individual departments are often charged with teaching multiple languages, the survey did not constrain respondents to describing the process for any individual language and, in fact, most respondents checked several boxes in response to the question of for which language the survey is valid. This multiplicity of languages coupled with the format of the survey which allowed for answers of "Other" for almost every question lead to a situation in which it became difficult to determine what weight to give any single answer. In other words, if a person had indicated that the survey covered French, Italian, and German, and that reading, writing, grammar, and speaking were assessed for placement purposes, it could be the case German placement was based primarily on reading, while French and Italian were based on the other skills. Though some respondents included marginal notes to the effect that a particular answer only covered a particular language, most of the surveys did not include such notation. Since there was no principled way to disambiguate the responses, it was impossible to get detailed language specific information (e.g. Japanese placement most often includes reading whereas French placement most often includes grammar). For this reason, no differentiation between placement procedures and languages can be made.

What skills are most often assessed for placement purposes?

Given the focus on the communicative use of the language in the pedagogical literature, one would expect programs to be especially interested in assessing their students' communicative ability. Certainly, the wealth of assessment procedures used in the programs in general (cf. Table 5) would indicate that programs have, for the most part, moved beyond a single-minded focus on the raw materials of the language, grammar, vocabulary, and so forth, and have begun to explore the students' ability to use the language.

Despite this trend in the instructional component of the program, placement tests are still fairly restricted in terms of the types of skills assessed. Reading, grammar, and vocabulary were the three most often tested skills and although there were a good number of programs that assessed speaking as well, the number was only half that of the more traditional skills. The lack of speaking assessment can be explained to a large extent by the difficulty of assessing speaking in any

type of large scale testing procedure. Some respondents indicated that although they would like to assess speaking, the constraints of the placement testing process effectively ruled out anything other than selected-response testing. In fact, many of the comments about the promise of on-line testing were specifically interested in the testing of oral proficiency. Though the lack of oral ability testing can be attributed to constraints rather than avoidance, it is a little harder to justify the widespread assessment of grammar for placement purposes in a communicative context.

Do programs tend to use commercially-produced or self-produced materials?

Despite the availability of commercial placement test, especially in the more commonly taught languages, a larger number of respondents reported using a locally produced test for placement. This could be partially a type of response bias in which programs using locally produced materials felt more inclined to respond to a survey on placement practices. Despite the use of locally produced test and the predominance of constructed-response test use in general, relatively few programs reported the use of item analysis in the development process. If teachers are mainly responsible for placement test construction, some basic knowledge of common test development practices should be beneficial.

How do language programs validate their placement instruments?

For the majority of the programs responding, test validation was essentially a process of giving the test to current students and looking at the results. Since piloting was the most common form of cut score setting, and since relatively few programs reported performing item analyses, it is likely the case that most programs considered a test valid if current students performed more or less as expected in terms of total scores. Although this is certainly an important quality of a placement test, a more thorough validation procedure would give the potential for a better placement process.

How involved are teachers in the placement process?

At almost all stages of the placement process, teachers do the bulk of the work. Except for registering students for the test, which seems to be the province of administrative offices, teachers are most often called upon for test administration, test scoring, and score reporting. In addition, in the case of internally produced tests, teachers are also the major players in test development. Of course, involvement does not necessarily mean investment, and it could very well be the case that many teachers see this aspect of their job as a chore rather than an academic endeavor on par with pursuing their own research interests.

What issues would language programs like to see addressed in the handbook?

In response to solicitations for handbook suggestions, many of the comments dealt with the distinction between testing to assess language ability in general and testing to ensure appropriate placement. Several comments addressed the need for institution specific placement procedures:

Definition – placement does not refer to achievement, rather to "placement". It is a matching instrument to match students' current ability to courses. Placement results are relevant to the particular institution and course selection available. (Private University)

–ending the practice of borrowing some placement instrument from other sites – stressing placement into a curriculum that is local – building a valid one based on changing curricular goals using alternative ways to place students and monitor these (Public University)

Other respondents stressed the notion of fairness with respect to the types of courses offered and the placement procedures used:

Fair assessment of students based on proficiencies which a program espouses. (Private University)

I would like to see learner perspectives on placement testing –both in terms of satisfaction, and in terms of how they perceive the testing? Do they think the tests and their outcomes were a good predictor of the kind of language and language tasks that were included in the course that they move into.(Public University)

In addition to concerns with the overall philosophy of testing, several respondents indicated that they would like to see the question of oral ability testing addressed. The difficulties of testing the spoken language and the potential for computerized tests to ease the burden seemed to be fairly common concerns:

If the placement exam includes speaking, how is the issue of *compensation* dealt with for the faculty who must evaluate the speaking proficiency? (Public University)

Quick and easy to administer oral placement? (Private University)

Options for computerized placement testing. Are there options to test production skills? (Private University)

I would like to know about online tests that can do a relatively quick evaluation of speaking ability. (Public University)

Other respondents raised concerns about the placement of heritage speakers as well as difficulties placing students in programs where there might be an incentive to do poorly on the test.

Distinguishing heritage language placement and establishing place-out procedures. (Public University)

How do departments control some students' deliberate poor scoring on the placement test in order to place into a level lower than their ability so that they may have a (perceived) easier course? (Private University)

Conclusions

The placement survey, though not perfect by any means, did seem to provide enough information about what is happening in various language programs around the country to give some direction as to the types of issues that should be covered in the placement handbook. Because teachers seem to be the people mainly responsible for the creation, administration, and evaluation of placement tests, it would make sense for the handbook to target that population rather than testing professionals (though the two are by no means mutually exclusive). This means that the handbook should provide basic information about general testing issues, such as reliability and validity, while also giving nuts and bolts instruction in item writing and item analysis. It is also possible to also have one or more sections of the manual devoted to slightly more technical issues for those teachers or administrators who have measurement experience. Because programs tend to test multiple skills, issues specific to the testing of those skills, such as the use of raters, need to be addressed. Also, as many programs expressed interest in computer based testing, issues pertaining to different test formats will need to be discussed, though the technical details of computerized assessment is probably beyond the scope of this manual.

Given those general parameters, a potential outline of handbook sections or chapters is presented below.

1. What is placement, articulation, curriculum?
2. How do programs tend to approach placement?
3. General testing issues: Validity, reliability, practicality
4. Determining what to test
5. Determining how to test
6. Developing selected-response tests
7. Interpreting test scores from selected-response tests
8. Developing constructed-response tests
9. Scoring and interpreting scores of selected-response tests
10. Setting cut-scores or guidelines
11. Maintaining test quality
12. Issues in testing specific skills
13. Issues in testing specific populations (e.g. heritage students)
14. Issues in using commercially available tests
15. Issues in test fairness

16. Introduction to computer-based testing

17. Issues in computer-based testing

18. Advanced statistical analysis

As the manual begins to take shape, it might be useful to follow-up with some of the programs that responded to the survey to get detailed information on how actual programs are dealing with some of the issues raised in the handbook.

References

- Brown, J. D. (2001). *Using surveys in language programs*. Cambridge University Press.
- Rhodes, N. C., & Branaman, L. E. (1999). *Foreign language instruction in the United States: A national survey of elementary and secondary schools*. McHenry, IL: Center for Applied Linguistics and Delta Systems Co., Inc.
- Walker, G., & Li, M. (2003). *Report on the Survey of East Asian Language Programs* (No. 11/30).
- Wherritt, I., & Cleary, T. A. (1990). A national survey of Spanish language testing for placement or outcome assessment at B.A.-granting institutions in the United States. *Foreign Language Annals*, 23(2), 157-165.

Appendix

Issues in placement survey

Issues in Placement Survey

Part One: Basic Program Information

Name of school / college / institution: _____

Location (state): _____ Type: Public Private

Level: High school College/university

Other: _____

Which language(s) does this survey cover? (Check all that apply)

- | | |
|-------------------------------------|---|
| <input type="checkbox"/> Arabic | <input type="checkbox"/> Korean |
| <input type="checkbox"/> Chinese | <input type="checkbox"/> Portuguese |
| <input type="checkbox"/> French | <input type="checkbox"/> Russian |
| <input type="checkbox"/> German | <input type="checkbox"/> Spanish |
| <input type="checkbox"/> Hawaiian | <input type="checkbox"/> Spanish for Spanish speakers |
| <input type="checkbox"/> Hebrew | <input type="checkbox"/> Swahili |
| <input type="checkbox"/> Indonesian | <input type="checkbox"/> Thai |
| <input type="checkbox"/> Italian | <input type="checkbox"/> Vietnamese |
| <input type="checkbox"/> Japanese | <input type="checkbox"/> Other: _____ |

How many students are in your program? _____

Is foreign language study required at your school?

Yes (# of semesters: _____) No Other: _____

What types of language courses and activities are offered in your program? (Check all that apply)

- Integrated skills
- Literature only
- Exploratory (general exposure to one or more language)
- Conversation only
- Advanced Placement
- Accelerated / intensive
- Content courses (history, math, science, etc.) taught in the foreign language
- Study abroad
- Homestay
- Other: _____

Comments (Basic Program Information): _____

Part Two: Language Ability Assessment

Check all of the following that your program uses to assess language ability in contexts other than placement:

- Oral proficiency interviews (students are interviewed in the language by the instructor)
- Student presentations (students prepare speeches or presentations which are assessed)
- Authentic activities (students complete real world tasks or role-play such tasks in the language)
- Student portfolios (a collection of material produced by the student)
- Student self-assessment
- Translation exercises
- Selected-response tests (multiple-choice, matching, etc.)
- Constructed-response tests (short answer, essay, cloze, dictation, etc.)
- Other: _____

Are the majority of teachers in your program aware of the national *Standards for Foreign Language Learning* (1996) and/or your state's version of the standards?

Yes No

Have foreign language assessment procedures in your program changed because of your awareness of the standards?

Yes No

Proficiency scales (such as the ACTFL proficiency guidelines) are sometimes used to express language ability in descriptive terms. For example, intermediate language ability might be defined in terms of the tasks that a learner at that level can be assumed to perform successfully (e.g. Buying a ticket to a movie or writing a simple letter to a friend). Does your program use any proficiency scale to describe the language ability of your students?

Yes No

If yes, which of the following best describes how this scale was developed?

- Internally, through discussions among program faculty
- Internally, but with knowledge of scales used in similar programs
- Jointly, with collaboration between our program and other programs / organizations
- Externally, by a group or authority outside of our program
- Other: _____

Comments (Language Ability Assessment): _____

Figure A1. Survey Page 1

Part Three: Placement Test Administration

Do you use a language placement test?

- Yes No (*Please skip to Part Five on page 4*)

How often is the test administered?

- Once a semester
 Several times per semester
 Only when needed
 Other: _____

How do students register for the test?

- First-come, first-served Multimedia lab staff
 Mandatory participation Testing office
 Admissions office Test given during class time
 Department office Other: _____

Who administers the test?

- Teachers
 Admissions office staff
 Testing office staff
 Multimedia lab staff
 Student help
 Other: _____

Who scores the test?

- Teachers
 Admissions office staff
 Testing office staff
 Multimedia lab staff
 Student help
 Machine-scored
 Other: _____

Who communicates the scores to students?

- Teachers
 Admissions office staff
 Testing office staff
 Academic advisor
 Scores are posted
 Other: _____

How many times may the students take the test?

- Only once
 Not more than twice
 Unlimited
 Retest with special permission only
 Other: _____

Comments (Placement Test Administration):

Part Four: Placement Test Content

Which of the following are explicitly assessed for placement purposes? (Check all that apply)

- Speaking
 Listening
 Reading
 Writing
 Grammar
 Vocabulary
 Ability to complete a real world task using a combination of skills
 Other: _____

Do you use an *internal* (locally produced) or *external* (commercially produced) placement test?

- Internal External Both

A. Internally developed tests

Who created the test?

- Current / past teachers
 Current / past administrator
 Special committee
 Language resource center
 Individual test developer
 Other: _____

Where does the test content primarily come from?

- Course textbooks
 Authentic material
 Course objectives
 Original materials
 State / local content standards
 Other: _____

How often is the test revised?

- Once or twice a year
 Once every several years
 When necessary
 When conditions (personnel, financial) permit
 Never revised
 Other: _____

If your test was revised, for which reason(s) was it revised?

- To prevent cheating
 To assess a skill not previously assessed (e.g., adding a listening component)
 To better reflect the current course contents
 To address a deficiency in the current test
 To make the test contents more timely
 To lengthen or shorten the test
 Other: _____

How was the test validated? (Check all that apply)

- Current students took it
- Language specialists reviewed it
- The contents were matched to course objectives
- An item analysis was performed
- Other: _____

What was the main reason for internally developing the test? (Check all that apply)

- No commercial test exists for the language
- Commercial tests were not suitable
- Funding / resources for development were available
- To better control test contents
- To reflect the ability range of students in the program
- To supplement an external test
- As a temporary measure only
- Other: _____

Comments (Internal Test):

B. Externally developed tests

Which externally developed test(s) do you use?

- | | |
|--|--|
| <input type="checkbox"/> CAPE (BYU) | <input type="checkbox"/> ACTFL OPI |
| <input type="checkbox"/> AP subject test | <input type="checkbox"/> SOPI |
| <input type="checkbox"/> CLEP | <input type="checkbox"/> Speaking Test (CAL) |
| <input type="checkbox"/> SAT II subject test | <input type="checkbox"/> Japanese Proficiency Test |
| <input type="checkbox"/> Wisconsin Test | <input type="checkbox"/> PASS |
| <input type="checkbox"/> Michigan Test | <input type="checkbox"/> MLPA |
| <input type="checkbox"/> Georgetown Test | <input type="checkbox"/> Other: _____ |

Do you require all test sections?

- Yes No Depends on student

What was the main consideration for choosing the test?

- Familiar to teachers
- Widely used in other language programs
- Have always used it
- Gives consistent results
- No other options
- Other: _____

How was the test validated for your program? (Check all that apply)

- Current students took it
- It was reviewed by instructors / department head
- The contents were matched to course objectives
- An item analysis was performed
- Other: _____

Comments (External Test):

Part Five: Course Placement

How is the final course placement made?

- Student self-selection
- Advisor's recommendation
- Placement test scores
- Teacher's recommendation
- Previous semesters of study
- Other: _____

How were cut scores determined for various courses?

- Test was administered to current students
- Test-maker's recommendation
- Based on proficiency level
- Other: _____

Can students change classes once instruction has begun?

- Yes, during add / drop period
- Only with special permission
- Course changes not allowed
- Other: _____

Comments (Course Placement):

Figure A3. Survey Page 3

Part Six: Comments and additional information

Which of the following best describes your position?

- State foreign language coordinator
- District foreign language coordinator
- Foreign language department chair
- School foreign language coordinator
- Language resource center director
- Language teacher
- Testing coordinator
- Other: _____

We are planning to produce a manual for placement testing in foreign language programs. Are there any areas that you would particularly like to see highlighted?

Please mail or fax the completed survey to:

University of Hawai'i at Manoa
National Foreign Language Resource Center
Issues in Placement Project
1859 East-West Road #106
Honolulu, HI 96822

Fax: (808) 956-5983

May we contact you for additional information if needed? If so, please fill in your contact information. (Your personal information will be kept confidential.)

Name: _____

E-mail: _____

Address: _____

Phone: _____

This is the end of the survey. Please be sure that you have filled out both sides of each page. Thank you for your participation!

Figure A4. Survey Page 4