

GAVILAN COLLEGE

Renewal Validation Report For The Gavilan English As A Second Language Objective Test And Oral Interview Performance Assessment



Submitted to

Dr. John Poggio
Director, Center for Educational Testing and Evaluation
University of Kansas
jpoggio@ku.edu

Dr. Doug Glasnapp
Professor, Psychology/Research in Education
University of Kansas
drg@ku.edu

Arnold Bojorquez
Matriculation Coordinator
California Community College Chancellor's Office
abojorqu@cccco.edu

Sally Montemayor
Matriculation Specialist
California Community College Chancellor's Office
smontema@cccco.edu

Prepared by
Terrence Willett
Director of Research
Gavilan Community College
5055 Santa Teresa Blvd
Gilroy, CA 95020
(408) 848-4852
twillett@gavilan.edu

April 15, 2005

Renewal Validation Report For The Gavilan English As A Second Language Objective Test And Oral Interview Performance Assessment

Introduction

For placement into English as a Second Language (ESL) classes, Gavilan Community College currently uses the provisionally approved Oral Interview Performance Assessment, the provisionally approved Gavilan English as a Second Language Objective test (GESL), and the CELSA instrument along with multiple measures such as amount of prior English instruction. The student begins with the Oral Interview with lower scoring students then taking the GESL and higher scoring students taking the CELSA. After the scoring of the objective test, the interviewer (an ESL faculty member) determines a final placement based on the Oral Interview score, the objective test score, and additional personal factors obtained via the interview process.

This report briefly recaps evidence provided last year for the initial approval report and provides preliminary evidence concerning validity, reliability, and disproportionate impacts not available previously. The report is organized by criterion and each component of the assessment is addressed for each criterion as appropriate. CELSA is an approved second party test that will be discussed when necessary but the focus of the evidence will be on the GESL, a locally developed objective test, and the Oral Interview, a locally developed performance assessment.

Background on Gavilan ESL curriculum

Gavilan College has a range of ESL classes that focus on writing/reading, listening, and speaking. Table 1 shows the Gavilan ESL curriculum with classes grouped by faculty determined skill level.

Table 1. Gavilan ESL curriculum (abridged for clarity).

Level	Class
Low Beginning*	ESL 501 Survival ESL I ESL 502 Survival ESL II ESL 503 Survival ESL III ESL 510A Basic ESL ESL 510B Basic ESL
High Beginning	ESL 521 Basic Listening and Speaking ESL 522 Fundamentals of Reading ESL 523 Fundamentals of Writing and Structure
Low Intermediate	ESL 531 Intermediate Listening and Speaking ESL 532 Reading and Vocabulary Development ESL 533 Development of Composition and Structure
High Intermediate	ESL 541 Intermediate Listening and Speaking ESL 542 Reading and Vocabulary Development ESL 543 Development of Composition and Structure ESL 546 Introduction to Computers for ESL Students
Advanced	ESL 552 College ESL Reading ESL 553 College ESL Composition ESL 555 Grammar Review ESL 556 Information Competency for ESL Students ESL 561 Communication Skills for College ESL 563 Composition and Reading

* The 501-502-503 sequence is equivalent to the 510A-510B sequence.

Overview of the GESL and Oral Interview

The GESL test is a timed 40 minute, 51 item objective test with a cover page containing instructions and three example problems. The Oral Interview is a performance assessment that consists of ten questions each of which is scored with a rubric on a four point scale from zero to three. The interview is designed to test listening and speaking skills and to gather personal data for use as “multiple measures”. Full time ESL faculty conduct the interviews one-on-one with students after a norming session to standardize scoring. Copies of the GESL test and Oral Interview forms were supplied with last year’s report and have not changed. The Oral Interview has a Form A and B that contain the same questions in slightly different order and are considered equivalent.

Validity and Fairness

Content Validity

Evidence concerning content validity was provided in last year's report and is abridged here for brevity. In general, the test items related well to class skills. The Matriculation Advisory Committee suggested that the GESL's coverage of skills at the ESL 520 level may not be adequate for reliable placement. In cases where a student's GESL score is in this transitional area, the faculty re-examine the Oral Interview results to determine if the CELSA should also be administered.

Test Bias

Evidence concerning content validity was provided in last year's report and is abridged here for brevity. Both the GESL and Oral Interview appeared free from bias according a sample of 24 representative ESL students.

Cut Scores

A committee of four current full time ESL faculty set initial cut scores for the GESL and Oral Interview. Students and instructors received consequential validity surveys to determine the adequacy of their placement, which is based upon a combination of scores on the GESL and Oral Interview and the multiple measures obtained during the interview process. Table 2 shows the definitions of ratings for instructors and students presented in the survey with each scale customized for the target audience. In general, instructor and student judgments appear to exceed the 75% agreement standard. Instructors felt that 69% of students were placed exactly correctly and 96% of students were placed reasonable accurately (Table 3). 81% of students felt that they were placed accurately (Table 4). In general, the instructors appeared to feel that more errors tended towards letting less prepared students into classes while students appeared to feel that more errors tended towards being placed too low. Due to a clerical error, students from the ESL520 level did not receive the survey. The ESL faculty plan to maintain the current cut scores and conduct another round of consequential validity surveys next semester to collect more evidence that will include the ESL520 level.

Table 2. Consequential validity rating scales for students and instructors.

Instructor Rating		Student Rating	
1	very overprepared, definitely should be in next level	1	I am overqualified
2	somewhat overprepared, perhaps should be in next level	2	I am well qualified
3	well prepared, should pass with reasonable effort	3	I am not completely qualified
4	somewhat underprepared, perhaps should be in previous level	4	I am not qualified at all
5	very underprepared, definitely should be in previous level		

Table 3. Instructor ratings of placement adequacy.

Instructor Rating	ESL501A		ESL501B		ESL502B		ESL510A		ESL510B		Total	
	Count	% within course	Count	% within course	Count	% within course	Count	% within course	Count	% within course	Count	% within total
1	0	0%	0	0%	0	0%	0	0%	1	10%	1	1%
2	0	0%	0	0%	2	25%	6	10%	2	20%	10	11%
3	1	100%	7	50%	6	75%	46	74%	6	60%	66	69%
4	0	0%	5	36%	0	0%	10	16%	0	0%	15	16%
5	0	0%	2	14%	0	0%	0	0%	1	10%	3	3%
Total	1	100%	14	100%	8	100%	62	100%	10	100%	95	100%

Table 4. Student ratings of placement accuracy.

Student Rating	ESL501A		ESL501B		ESL502B		ESL510A		ESL510B		Total	
	Count	% within course	Count	% within course	Count	% within course	Count	% within course	Count	% within course	Count	% within total
1	0	0%	4	29%	1	13%	4	6%	1	10%	10	11%
2	1	100%	9	64%	7	88%	49	79%	8	80%	74	78%
3	0	0%	0	0%	0	0%	2	3%	1	10%	3	3%
4	0	0%	0	0%	0	0%	2	3%	0	0%	2	2%
Total	1	100%	13	93%	8	100%	57	92%	10	100%	95	100%

Reliability and Errors of Measurement

The reliability of the GESL objective test was reported in last year's report and showed a split halves reliability of 0.93 and Cronhach's alpha of 0.92 indicating the test has high reliability. The standard error for the GESL objective test was reported in last year's report and appeared to be reasonable within placement levels to support the claim that the test is stable and reliable.

Reliability for the Oral Interview is provided by a norming session where raters review an audio taped interview and score according to a rubric. Raters discuss the scoring and then review a second taped interview so that all raters agree upon the proper scoring criteria. All raters are current ESL faculty. Last year's report did not include data on interrater reliability. Since that time, a sample of 23 oral interviews have been scored by two normed raters. The distributions of both raters' scores were approximately normal (the Kolmogorov-Smirnov statistics were 0.135 and 0.117 each with a lower bound p-value of 0.2). The two scores were highly correlated (Pearson $r = 0.876$, $p < 0.005$) and met the minimum threshold of 0.7 for interrater reliability. The ESL faculty plan to collect more data on interrater reliability next semester to ensure the quality of norming efforts.

Testing Special Groups

The college is prepared to provide alternative forms of all of our assessment instruments as necessary and appropriate to accommodate students with disabilities. Gavilan has a dedicated lab with specialized staff and equipment to assist visually and hearing impaired students, appropriate furniture for mobility impaired students, and extended time opportunities for verified learning disabled students.

Disproportionate Impacts

Last year's report indicated that within a pilot administration of the GESL there were no detectable disproportionate impacts by age or gender and that due to low numbers of non-Hispanics and disabled students, detection was not yet possible. Current data indicate similar patterns. In a

sample of students GESL placement level and age were not correlated (Spearman's $\rho = -0.006, p = 0.942, n = 139$) and neither were Oral Interview placement level and age correlated (Spearman's $\rho = 0.076, p = 0.373, n = 138$). No significant differences in placement level by gender were apparent for the GESL (Likelihood ratio (4) = 3.345, $p = 0.502$) or the Oral Interview (Likelihood ratio (5) = 4.884, $p = 0.43$). There are still insufficient numbers of non-Hispanics and disabled students to detect disproportionate impacts if they are present. Placement data will continue to be monitored for disproportionate impacts and more data becomes available.