

6-1994

An Analysis of the Relationship between Written and Oral Performances on the Spanish Proficiency Exam

Patricia A. Robitaille
The College at Brockport

Follow this and additional works at: https://digitalcommons.brockport.edu/ehd_theses

 Part of the [Curriculum and Instruction Commons](#), and the [Secondary Education Commons](#)

To learn more about our programs visit: <http://www.brockport.edu/ehd/>

Repository Citation

Robitaille, Patricia A., "An Analysis of the Relationship between Written and Oral Performances on the Spanish Proficiency Exam" (1994). *Education and Human Development Master's Theses*. 1155.
https://digitalcommons.brockport.edu/ehd_theses/1155

This Thesis is brought to you for free and open access by the Education and Human Development at Digital Commons @Brockport. It has been accepted for inclusion in Education and Human Development Master's Theses by an authorized administrator of Digital Commons @Brockport. For more information, please contact kmyers@brockport.edu.

AN ANALYSIS OF THE RELATIONSHIP BETWEEN WRITTEN AND
ORAL PERFORMANCES ON THE SPANISH PROFICIENCY EXAM

THESIS

Submitted to the Graduate Committee of the
Department of Education and Human Development

State University of New York

College at Brockport

in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Education

by

Patricia A. Robitaille

State University of New York

College at Brockport

Brockport, New York

June 1994

SUBMITTED BY:

Patricia A. Robitaille 7/6/94
Candidate Date

APPROVED BY:

Robert B. Riddle 7/6/94
Thesis Advisor Date

Kathy Strachan 6-3-94
Second Faculty Reader Date

Patricia E. Baker 7-7-94
Director of Graduate Studies Date

Table of Contents

I. Proficiency Testing in Perspective	1
II. Statistical Study	7
Method	8
Materials	9
Design	11
Results	12
III. Factors Affecting Relationship	17
IV. Significance of Study	24
V. References	28
VI. Appendices	33

Abstract

The purpose of this study is to investigate the relationship between written performance and oral performance on the Spanish Proficiency Exam which is administered to candidates in the Bilingual Multicultural Program at the State University of New York, College at Brockport. Due to the sequence of the exams, the researcher is interested in determining if students are more successful in the oral part if they have been successful in the written part. Factors that might account for differential levels of performance shown by the test scores are described and suggestions on the implementation of the testing procedure are proposed.

Proficiency testing, particularly oral proficiency testing, and its implications for teaching are frequent topics for foreign language educators and those concerned about the ability to compete globally. Magnan (1986) views the proficiency movement as having a major impact on the educational profession. Indeed, proficiency has become a “buzz word” that has captured the attention of many who are interested in foreign language testing and learning. As Omaggio (1983a) states, proficiency includes grammatical competence, sociolinguistic competence, discourse competence and strategic competence.

The language tests of the proficiency movement (Hewitt, 1993) have emerged to assess oral proficiency as well as writing, reading and listening skills. Omaggio (1983a) describes the oral test as one of the most difficult tests to create, schedule, administer and evaluate. Due to the complexity of assessing language attainment (Prabhu, 1990), the need for more research on language testing and the language learning process is greatly recognized by educators, test developers, policy makers and others interested in foreign

language learning process is greatly recognized by educators, test developers, policy makers and others interested in foreign languages. As a result, common standards, established by the ACTFL (American Council on the Teaching of Foreign Languages) Provisional Proficiency Guidelines and the ACTFL/ETS (Educational Testing Services) Oral Interview, have served to assess the proficiency of foreign language majors, teachers, and other professionals with a need for fluency in a second language (Schulz, 1986, p. 187). In addition, the ACTFL guidelines have helped to organize the curriculum for language teaching.

Although ACTFL Guidelines can be useful as an assessment tool (Henning ,1992) and offer advantages in the development of language testing, others (Stansfield, 1990) argue that there is a need for the development of language specific versions of the ACTFL Proficiency Guidelines where they do not currently exist. In this respect, these guidelines are provisional and they should be used with caution. The guidelines represent a “forward-looking” view (ACTFL

Guidelines) of the significant skills deemed valuable for the preparedness of foreign language teacher candidates.

Some researchers (e.g. , Magnan & Schulz, 1986) argue that tests of proficiency are global, summative measures that provide an assessment of the candidate's overall linguistic ability and are criterion referenced. Therefore, proficiency tests do not reflect one particular course of instruction nor do they evaluate what the student has learned. As a result, ACTFL Guidelines simply adhere to what the proficiency test is measuring, that being how much of the language is used by the person.

In addition to ACTFL , the organizations of MLA (Modern Language Association) and FSI (Foreign Service Institute) also have offered opportunities for discussion on issues of assessment testing and placement (Cleary & Wherritt, 1990,162). In fact, the ACTFL/ETS oral proficiency scale (Barnwell, 1991) represents an adaptation of the governmental Foreign Service Institute whereby the ACTFL/ETS adjusted the lower points of the FSI scale. Currently, the OPI (Oral

Proficiency Interview), developed by ACTFL, in conjunction with ETS , is the most well-known test to evaluate speaking skills of the foreign language learner (Caminero & Harlow, 1990, p. 489).

Symposiums relevant to proficiency testing, such as the LPA (Language Proficiency Assessment) Symposium held in 1981 (Schulz, 1986) also serve as forums to evaluate and exchange research, ideas and testing models necessary to improve proficiency testing. Likewise, some universities and schools perform studies to evaluate testing procedures and to develop language tests.

For example, Cleary et al (1990) conducted a national survey to determine the status of testing for placement and outcome assessment at The University of Iowa where FLAP (Foreign Language Assessment Project) served as their research project. In particular, FLAP focused on B.A. granting institutions that offered Spanish. The results of the study (Cleary et al, 1990) indicate that work in testing needs to include better instruments to assess language competency

outcomes necessary to meet teacher certification.

The University of Wisconsin conducted an experiment (Barhoum, 1989) comparing two proficiency testing models, the ACTFL/ETS Oral Proficiency Interview and another model developed by the University to determine which is more appropriate. Results (Barhoum, 1989) concluded that the ACTFL/ETS Oral Proficiency Interview was more suitable for use in the academic environment.

Often proficiency testing involves only oral proficiency, generally using the ACTFL Oral Proficiency Interview as demonstrated in the study at the University of Wisconsin. However, Meuser-Blinco and Villar (1993) directed a study at the University of Minnesota which compared results of the other skills, such as reading, writing and listening, in a proficiency based program versus a non proficiency program. The University of Minnesota is one of three programs in the United States with a proficiency based second language requirement (Meuser-Blinco et al., 1993) and its own foreign language proficiency test (Lange, 1987). The other two

programs are found at the University of Pennsylvania and the University of South Carolina.

The testing instrument, the CLA (College of Liberal Arts) Spanish Graduation Examination, provided the data to compare and analyze test scores. In this recent study of proficiency and non proficiency- based programs, Meuser-Blinchow et al. (1993) found that a proficiency based language requirement is an efficient means for students to achieve better communicative competence in all four modalities.

In addition to the previous academic institutions, other districts such as Pinellas County Florida and Connecticut (Stansfield, 1990) have developed their own foreign language tests which have generated much interest throughout the country.

As a result of the discussions regarding recent professional literature and studies relevant to proficiency testing, my interest focuses on the Spanish Proficiency Exam administered at the State University of New York, College at Brockport. Brockport University serves as a testing center

utilizing the Spanish and English Proficiency Exams as an exit outcome for candidates in the Master of Science in Education for Bilingual Multicultural Education program. Furthermore, Brockport University has a proficiency-based language requirement program.

This study will investigate the relationship between the written performance and oral performance on the Spanish Proficiency Exam at Brockport University and its implications for testing procedures. Due to the sequence of the exams, I am interested in determining if students are more successful in the oral part if they have been successful in the written part. I decided that the degree of relationship between the oral and written performances must be moderate. Therefore, the criterion of importance for this study was set at $r^2 = .50$.

The researcher believes and predicts that there will be a moderate to strong relationship between written and oral performances. Due to experience as a teacher and learner of foreign language, I have observed and been involved in the performance of oral and written skills. Therefore, as the

investigator of this study, I have a bias toward the predicted outcomes.

If it is assumed that a significant relationship between written and oral performance exists, then the administration of the oral part should occur only after success on the written part.

Method

Subjects

For this study, a random sample of 50 students' scores on the Spanish Proficiency Exam were selected from SUNY Brockport. Table 1 demonstrates these scores. The students involved, both graduate and undergraduate, are those seeking certification in the Bilingual Education Program. The students in this study represent various ages, race and sex.

Table I

Written and Oral Scores on Spanish Proficiency Examination

SUBJECT	WRITTEN SCORE	ORAL SCORE	SUBJECT	WRITTEN SCORE	ORAL SCORE
	X	Y		X	Y
1	25	15	26	50	70
2	26	16	27	50	75
3	28	19	28	50	60
4	28	25	29	51	69
5	28	30	30	51	72
6	30	35	31	52	67
7	32	28	32	53	75
8	32	32	33	54	86
9	34	48	34	54	68
10	35	49	35	55	63
11	35	45	36	55	72
12	36	53	37	56	76
13	37	52	38	56	81
14	38	54	39	57	77
15	38	50	40	58	72
16	39	53	41	60	95
17	40	55	42	62	91
18	41	51	43	63	86
19	41	59	44	63	80
20	42	42	45	65	86
21	45	70	46	67	90
22	45	68	47	68	95
23	46	45	48	69	90
24	48	66	49	69	91
25	49	70	50	70	95

Materials

The Spanish Proficiency Exam, both written and oral parts, was administered to the subjects. The students were tested and scored by university personnel. The Spanish

Proficiency Exam as administered comprises four separate tests which include the four modalities of reading, writing, listening and speaking. The reliability for writing and speaking is achieved through rater training.

The written performance on the Spanish Proficiency Exam represents the independent variable (X) for this study. Table 1 displays the range of scores from 0-70. The time allotted for the writing test is thirty five minutes. The format of the writing test consists of twenty one word sentence completions, ten grammatical questions, ten sentences to rewrite based upon previous sentences, six lines of scrambled words for the formation of sentences and a seven line dialogue.

The oral performance on the Spanish Proficiency Exam represents the dependent variable (Y). The range of the raw scores is from 0-100. Table 1 demonstrates this range. These scores are then translated to a scale from 1-5. For the purpose of this study, the researcher will be using the raw scores. The speaking test is ten to fifteen minutes in length.

The format of the oral exam is based upon questions, both formal and informal, posed by two interviewers. It is in the form of an interview patterned after the ACTFL OPI.

Both the written and oral examinations thus consisted of a number of different sections designed to elicit a variety of types of language from the examinee (Hall, 1993, p.30).

Design

The specific questions that will be addressed in this study include:

1. What is the mean of the written performance on the Spanish Proficiency Exam and what is the average around that mean?
2. What is the mean of the oral performance on the Spanish Proficiency Exam and what is the average around that mean?
3. As the written performance on the exam increases by one, what is the average change in the oral performance?
4. How strong is the relationship between the written performance and oral performance?

5. How much of the variation in the oral performance is explained by knowing the written performance? How much is unexplained? On a scatter plot of the sample pairs, is a straight line the best description of the trend provided by the data?

6. Using the standard deviation of the residuals, what is the average variation observed between the trend line and plotted data points?

7. For fifty people all having a written score near the mean score, what does the regression equation predict that the mean of the oral score will be?

8. For a group of fifty people with an average written score, the regression line predicts that 95% of them will have an oral score between what minimum and maximum scores?

In addition, there will be no statistically significant relationship between the written performance and the oral performance. Furthermore, this will be tested for chi-square $\chi^2 (4, N=50) = 9.49$ at a 95% confidence level.

Results

The criterion of importance established at the beginning of the study was $r^2=.50$. The coefficient of determination was found to be $r^2=.89$. Consequently, the relationship between written performance and oral performance was unusually strong as predicted by the researcher. Therefore, one's written performance has a strong effect on oral performance.

Furthermore, the correlation and regression analysis demonstrated that there was an unusually strong relationship between written performance and oral performance as a result of the coefficient of determination ($r^2=.89$). The strong relationship was further supported by Cramer's Phi ($\phi =.65$), (demonstrated in Appendix A), as well as the chi-square analysis ($\chi^2=42.02$). For interpretation, Appendix B shows the statistical work for determining the value of chi-square. Because the value obtained for chi-square in this study was higher than 9.49, the null hypothesis was rejected and a statistically significant relationship was found. Strong global relationships also existed in this study. However, the

coefficient of determination was the best predictor in determining the relationship.

Regarding questions 1 and 2, the data can be seen in Table 2 and will be discussed respectively.

Table 2

Summary of Statistics

VARIABLE	X
MEAN	47.52
VARIATION	12.85
RANGE	45
Ex	2376
Ey	
Exy	160975
Ex ²	121002
Ey ²	

In this study, the mean of the independent variable (X), the written performance on the Spanish Proficiency Exam, was 47.52 compared with 62.24 for the oral performance, the dependent variable (Y). The average variation around the mean for written performance was about 13 points compared with

about 22 points as the average variation around the mean of oral performance.

The written scores ranged from 25 to 70 while the oral scores ranged from 15 to 95. There was a much greater spread in the range of oral scores compared to the written scores.

As the written performance (X) increased by one point, on the average, the oral performance (Y) increased by 1.62 points.

Knowing the values for the written performance explained about 89% of the variation in the oral performance with about 11% of the variation being unexplained. Table 3 demonstrates the conditional means determined by the formula $Y^1 = a + b(x)$ where $b = 1.62$ and $a = -14.62$.

Table 3

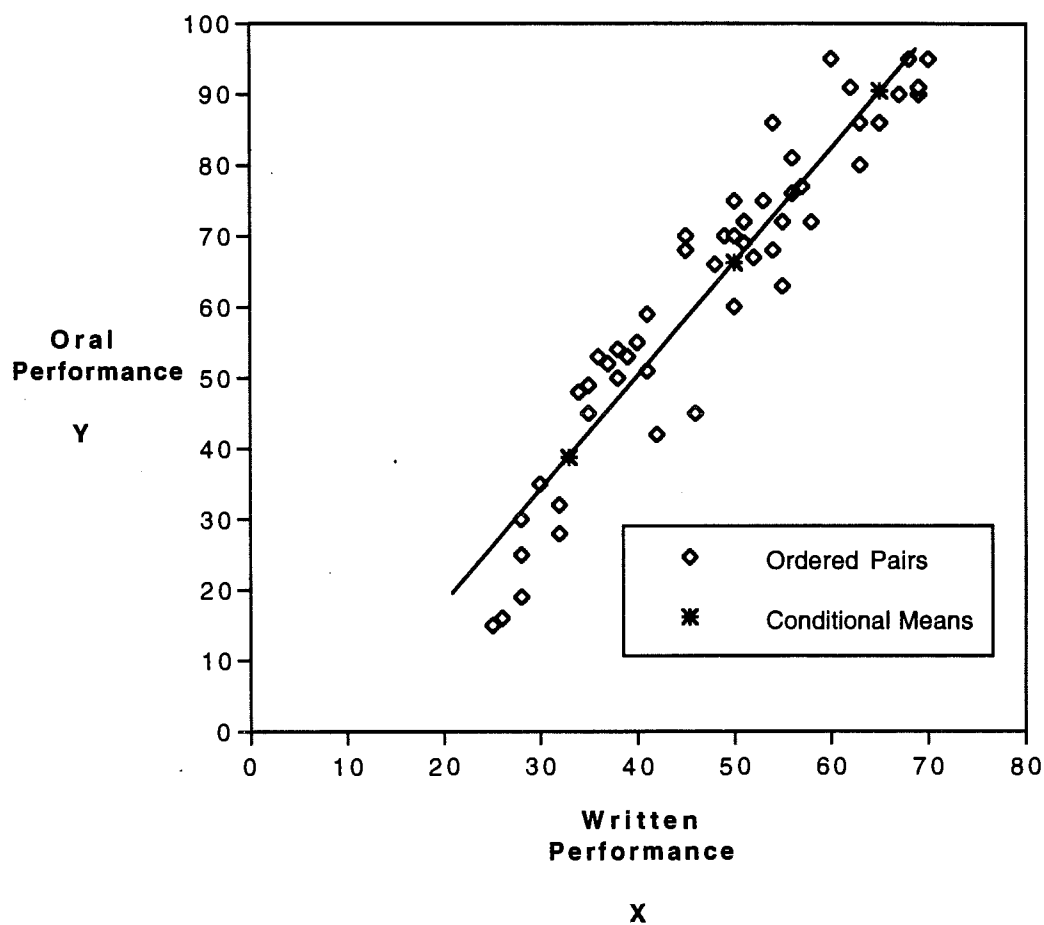
Conditional Means

WRITTEN PERFORMANCE X	ORAL PERFORMANCE Y
33	38.75
50	66.25
65	90.51

The scatter plot indicated that the slope is positive and that a straight line is the best description of the trend provided by the data. (Figure 1).

Figure Captions

Figure 1. A scatter plot of the relationship between written and oral performance.



Using the standard deviation of the residuals, the average variation observed between the trend line and the plotted data points was 7.23. Furthermore, for a group of 50 students all having a written score near the mean score, the regression line predicted that the mean of the oral score was 66.25. In addition, the regression line predicted that 95% of them would have an oral score between 68.33 and 64.17. This prediction interval was moderately small which reinforced the finding of a strong relationship between written performance and oral performance.

Discussion

The present findings of this study support my prediction that there is a strong relationship between written performance and oral performance. Explanations for such a strong relationship could be attributed to various factors but I will focus on methodology and the validity and reliability of the examination as causes for an unusually strong relationship and their implications.

To begin with, if one of the major goals of foreign

language learning is acquiring oral facility in the target language, then methods of instruction and testing must correlate with each other to achieve the goal (Robison, 1992, p. 493). With the arrival of the ACTFL Proficiency Guidelines and emphasis on proficiency, a great deal of attention has been devoted to oral testing yet inadequate attention has been paid to developing assessment strategies that test students in ways reflective of the methods used in the classroom. Few studies have looked at the effects of proficiency based teaching on student performance, classroom methods and materials, teacher expectations, and program administration.

If, indeed, proficiency is to become the dominant goal in foreign language programs, then we have to examine our curriculum and determine what changes need to be made in order to achieve this goal. It is important to consider, as Pennycook (1989) argues, that language teaching has undergone many transformations over the centuries. As a result, (Terrell, 1986), grammar is no longer the goal but rather a tool which can be used to achieve the goal of proficiency. However,

as we reexamine our curriculum and evaluate past teaching methods, we should see what has been effective so that the best of the past is not lost, but serves the purposes of the present (Rivers, 1968, p. 13).

In this respect, as we study those methods pertinent to the teaching of the four modalities, reading, writing, listening and speaking, Richards (1985) argues that the important issue is not what method to adapt; rather, how to develop procedures and activities that will enable objectives to be attained. Important factors that must be considered when implementing this curriculum are the nature of the instructors and the nature of the discipline to be learned (Schulz, 188).

According to Lado (1961) and Horwitz (1985), language teaching practices have been attributed to teaching habits of particular teachers and their own personal experiences in learning. Also, the trend or fashion which seemed prevalent at a particular moment in history influenced what was taught and the manner in which it was taught. These divergent attitudes about teaching have led to different instructional methods

which reflect language learning of students. As a result, the priorities in the teaching of the four skills (Rivers, 1968) are in very different order dependent upon the instructor and his own preconceived ideas about how languages should be taught (Horwitz, 1985).

However, regardless of individual teaching preference, there must be a balance in the instruction of writing, reading, listening and speaking so that one skill does not have more significance than the others as demonstrated in the past with reading and grammar as focal points of the foreign language curriculum. Of the four skills that are taught with supposedly equal emphasis, Terry (1989) argues that writing is perhaps the most poorly understood and the skill given the least attention. On the other hand, other investigators (Lado, 1961) argue that speaking skills are the least developed due to the complexity of assessing its attainment. In addition, speaking ability is the least practiced in the language field. As a result, without practice, one's speaking abilities will not improve. It is also important to realize that in the writing and

speaking skills, a foreigner will not achieve the same degree of mastery as a native speaker (Rivers, 1968, p.241).

Regardless of being a native or non native speaker, the skills of writing and speaking can be integrated to emphasize equal instructional time and practice to ensure that the methods of instruction correlate with the outcomes of proficiency testing. In fact, as Lado (1961) demonstrates, the ability to speak a language will greatly expedite and facilitate learning to write it. One who can speak a language well can also understand and read it well. As a result, one's written performance not only has an effect on oral performance but rather, oral performance greatly affects written performance.

The data collected in this study and its results also focus on the issues of validity and reliability of the Spanish Proficiency Exam with emphasis on the oral exam as a means to explain the relationship between written performance and oral performance. The fundamental problems with oral examinations are those of reliability (the consistency with

which different examiners mark the same test) and validity (whether or not a test assesses what it sets out to assess) (Hall, 1993, p. 24). Recent studies of the reliability of oral language exams have produced more positive results. In particular, the American Foreign Service Institute Oral Interview Test has provided research in which the problems inherent in the system do not include reliability among raters of the same performance (Wilds, 1975, p.33)

However, research (Lado,1968) indicates that oral proficiency in a foreign language has low reliability yet has validity. Likewise, written proficiency has varying degrees of validity but is of low reliability. Some practical problems that account for low reliability in both oral and written proficiency include the professional background of the rater (Shohamy, 1992a) and the need to develop different scales for different types of writing. Questions regarding the training and background of the rater need to be addressed as to determine the reliability of the test. Proper training of instructors who administer the tests is essential.

Similarly, a comparison between the format of the written exam and oral exam needs to be addressed regarding the issue of reliability. For example, in writing there is more time for the candidate to formulate and analyze the questions. On the other hand, in an oral exam the candidate must respond with less time thus not being able to analyze the questions as carefully. Therefore, the proficiency test should not demand that candidates demonstrate the same structural range in speech as in writing (Hall, 1993, p.35).

In addition to the reliability of both the written and oral tests, the validity of these tests needs to be addressed in a serious manner. The random scores of written and oral tests utilized in this study imply a very strong relationship which negates any fluctuation of scores thus deeming the Spanish Proficiency Exam both reliable and valid. It is obvious that realistic written and oral proficiency ratings were determined for this test administered at Brockport University. Furthermore, the raters of the test have professional training and are native speakers in Spanish. Therefore, the purpose of

the Spanish Proficiency Exam and this study have proven its validity.

Recommendations

Since the study yielded strong results, the following recommendations and suggestions are therefore indicated. To begin with, the testing procedure of the Spanish Proficiency Exam at SUNY Brockport needs to be reevaluated as a result of the study's findings. The data imply that the oral part of the exam should be administered to a candidate that has successfully completed the written part. It is evident that the written performance does predict the oral performance. Therefore, the candidate must meet certain criteria on the written part in order to be eligible for the oral part.

Secondly, in order for the candidate to fulfill the necessary criteria, the Department of Foreign Languages at Brockport University needs to make curricular changes designed to develop proficiency. In this respect, because proficiency varies from person to person and its development will depend on both the student's needs and prior experiences

(Terrell, 1986, p.184), the Department will need to implement those changes that reflect progress in a language according to proficiency (Millstone, 1983). Change can be made in teaching methods, textbooks or expectations. However, despite the emphasis on oral skills in the foreign language curriculum, many instructors still do not formally test Spanish proficiency of students on a regular basis (Omaggio, 1984). Therefore, the main purpose and rationale of the ACTFL Guidelines and the OPI is to create pressure on those teachers to upgrade the level of learning which will eventually lead to improvement of foreign language proficiency (Shohamy, 1992b).

Thirdly, in order for teachers, administrators and students to implement valuable information regarding teaching and learning for proficiency, Richards (1985) argues that the role of the test must be viewed as fundamental and vital in the educational process. Tests become sources of meaningful information about the improvement of foreign language. As educators, we need to utilize this information to develop new practices and theories which should be imparted to others.

Furthermore, it is the duty of everyone to keep abreast of new research, studies, methods, etc. that enhance the educational profession (Bennett, 1986). Finally, given the importance of proficiency testing and teaching, Caminero and Harlow (1990) state that there is a need for subsequent studies in order to track the progress of formal proficiency testing which will enable educators to monitor changes within the profession on this important topic.

In particular, my study's findings do indeed track the progress of proficiency testing at Brockport University and as result, benefit both students and faculty members in evaluating the administration of the Spanish Proficiency Exam. The study's results of a strong relationship clearly indicate that a change must be enacted in the administration of the exam. Such a change involves sequencing the exams in order to alleviate the financial cost of administering the expensive oral exam to a candidate more than once. Utilizing this study's findings to develop a new testing procedure would be very effective financially for Brockport University and its students

and faculty. Therefore, as a candidate in the Bilingual Multicultural Program at Brockport University, I believe that the purpose and results of my study have a significant impact on the financial aspect of the Spanish Proficiency Exam.

References

- ACTFL. (1986) ACTFL provisional proficiency guidelines.
Hastings-on-Hudson, New York: American Council on the
Teaching of Foreign Languages.
- Barhoum, K. (1989). Proficiency testing: Rating vs. profiling.
Foreign Language Annals, 23, 353-367.
- Barnwell, D. (1991). Proficiency testing and the schools.
Hispania, 74, 187-189.
- Bennett, R. (1986) . How much theory, how much practice?
Hispania, 69, 176.
- Caminero, R. & Harlow, L. (1990). Oral testing of beginning
language students at large universities: Is it worth the
trouble? Foreign Language Annals, 23, 489-498.
- Cleary, T. A. , & Wherritt, I. (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.
- Hall, C. (1993) The direct testing of oral skills in university
foreign language teaching. IRAL, 31(1), 23-38).

- Henning, G. (1992). The ACTFL oral proficiency interview: Validity evidence. System, 20, 365-392.
- Hewitt, C. B. (1993, March) . Assessment of student learning in foreign language. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.
- Horwitz, E. K. (1985) Using student beliefs about language learning and teaching in the foreign language methods course. Foreign Language Annals, 18(4), 333-340.
- Lado, R. (1961) . Language testing: The construction and use of foreign language tests. London: Longmans.
- Magnan, S. S. (1986). From achievement toward proficiency through multi-sequence evaluation. In James, C. J. (Ed.) , Foreign language proficiency in the classroom and beyond. (pp.117-145). Illinois: National Textbook.
- Meuser-Blinchow, F. & Villar, S. M. (1993). Proficiency requirement-based and nonproficiency requirement based language programs: How do students rate? Foreign Language Annals, 26(1), 49-62.

- Millstone, A. (1983, December) . Proficiency testing and programs at the university of south carolina. Paper presented at the annual meeting of Modern Language Association, New York, NY.
- Omaggio, A. C. (1983). Proficiency-oriented classroom testing. Language in education: Theory and practice. Washington D.C.: Center for Applied Linguistics.
- Omaggio, A. C. (1984). Methodology in transition: The new focus on proficiency. Modern Language Journal, 67, 333-340.
- Pennycook, A. (1989). The concept of method, interested knowledge and the politics of language teaching. TESOL, 23, 589-618.
- Prabhu, R. S. (1990) There is no best method-why? TESOL, 24, 161-176.
- Richards, J. C. (1985) . The context of language testing. Cambridge: Cambridge University.
- Rivers, W. (1968) . Teaching foreign-language skills. Chicago: University of Chicago.
- Robison, R. E. (1992). Developing practical speaking tests for

the foreign language classroom: A small group approach.

Foreign Language Annals, 25(6) 487-492.

Schulz, R. A. (1986) . From achievement to proficiency through classroom instruction: some Caveats. Hispania, 69, 187-192.

Shohamy, E. (1992a) . The effect of raters' background and training on the reliability of direct writing tests. Modern Language Journal, 76, 27-33.

Shohamy, E. (1992b) . Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. Modern Language Journal, 76, 513-521.

Stansfield, C. W. (1990). Some foreign language test development priorities for the last decade of the twentieth century. Foreign Language Annals, 23(5), 395-401.

Terrell, T. (1986) . Recent trends in research and practice: teaching Spanish . Hispania, 69, 193-202.

Terry, R. M. (1989) . Teaching and evaluating writing as a communicative skill. Foreign Language Annals, 22(1), 43-

54.

Wherritt, I. (1990) . Development and analysis of a flexible spanish language test for placement and outcome assessment. Hispania, 73(4), 1124-1129.

Wilds, C. (1975) . The oral interview test. In R. L. Jones & B. Spolsky, (Eds.) , Testing language proficiency (pp. 29-44) . Arlington, VA: Center for Applied Linguistics.

Appendix A

Cramer's Phi

$$\phi = [(X \text{ obtained}) / n * (k-1)]^{1/2}$$

$$\phi = [42.02 / (50 * (2))]^{1/2}$$

$$\phi = [42.02 / 100]^{1/2}$$

$$\phi = .65$$

Interpretation of Cramer's Phi

The ϕ obtained in this study was .65. This is higher than the .50 set as the criterion of importance. It indicates a very strong global relationship between written performance and oral performance. Thus, by categorizing variables, 89% of the variation in Y is explained by knowing X while 11% of variation remains unexplained. Furthermore, the global relationship is as reliable as individual predictions.

Appendix B

Chi Square Table

CELL	fo	fe	(fo-fe)	(fo-fe) ²	(fo-fe) ² /fe
R1 C1	12	4.76	7.24	52.42	11.01
C2	5	6.8	-1.8	3.24	0.48
C3	0	5.44	-5.44	29.59	5.44
R2 C1	2	4.76	-2.76	7.62	1.6
C2	12	6.8	5.2	27.04	3.98
C3	13	5.44	-2.44	5.95	1.09
R3 C1	0	4.48	-4.48	20.07	4.48
C2	3	6.4	-3.4	11.56	1.81
C3	13	5.12	7.88	62.09	12.13
				$\chi^2=42.02$	

Appendix C

Contingency Table

Y Oral Performance \ X Written Performance	< 51 C ₁	51 to 73 C ₂	> 73 C ₃	Row Totals
< 41 R ₁	12 4.76	5 6.80	0 5.44	17
41 to 54 R ₂	2 4.76	12 0	3 5.44	17
> 54 R ₃	0 4.48	3 6.40	13 5.12	16
Column Totals	14	20	16	50