Refereed Articles

An Examination of English as a Second Language Assessment Tools Utilized by Adult Literacy Providers in Pennsylvania

Rhonda S. Dennis

Abstract

This study examines the perceptions of English as a Second Language (ESL) adult literacy program teachers and administrators regarding the adequacy of ESL assessment tools utilized by their programs. Although the literacy providers who responded to the survey seemed to agree that they were generally satisfied with the performance of their assessment tools, their responses also gave strong indications that within specific purposes of use, these tools were weak in the assessment of several skill areas. These areas included English, writing, ELT Documents No. 11 reading, speaking, vocabulary development, and the identification of difficulties with English grammar, pronunciation, student language goals, and educational background information.

Introduction

Interest in the quality of assessment tools utilized by English as a Second Language (ESL) adult literacy programs has been evident at the federal, state, and local levels. At the 1995 Pennsylvania Adult Education Midwinter Conference, ESL assessment was one of the major topics of discussion among ESL adult education providers. Information about the quality of assessment tools available and utilized by programs was shared and debated. Discussions seemed to centralize around three basic concerns: 1) the quality of ESL assessment tools available, 2) the appropriateness of tools to the programs utilizing them, and 3) guidelines or standards for assessment set at the state and federal levels.

Rhonda S. Dennis is an Instructor of English as a Second Language and Director of the Learning Resource Center at Wilson College, Chambersburg, PA.

Locally, ESL adult education programs seek assessment tools that best serve their program needs. They want to be confident that their assessment tools are valid, appropriate for the characteristics of their student populations, reliable, fairly easy to interpret, and relevant to their program objectives (Henning, 1987). They need to be able to provide learners in their programs with accurate information on "what they need to learn to meet their goals and how they are making progress toward achieving these goals" (Alamprese, 1995, p. 20).

Additional criteria, including federal and state mandates, guidelines, and indicators for program quality in adult basic educationmust be considered for ESL adult literacy programs in Pennsylvania. For example, adult education programs receiving federal or state funds provide to the Pennsylvania Bureau of Adult Basic and Literacy Education (ABLE) assessment data for adults enrolled in their programs (Fluke, 1992). In order to make such reports, these programs need assessment tools that will collect the kind of data state and federal agencies are requiring.

At the state level Pennsylvania's ABLE Bureau is conducting a project, titled Project Equal, that has as its goal assisting local education providers with this collection task and includes as one of its objectives "strengthening [providers'] skills in selecting and administering appropriate learner assessment instruments and in interpreting assessment results for guiding learners and managing programs" (Alamprese, 1995, p. 21). In addition, indicators of program quality within the areas of assessment may be established and serve as a comprehensive model for measuring the performance of adult basic and literacy programs in the Commonwealth (Alamprese, 1995; Shaw, 1993).

In order for Pennsylvania programs to meet the needs of their learners and the quality standards set at federal and state levels, the assessment tools utilized by these programs need to serve successfully each specific program. One way to determine whether or not the tools meet these requirements is to conduct an evaluation of the assessment tools utilized by these programs. This study addresses these concerns by investigating whether or not ESL adult literacy providers in Pennsylvania, as a whole, believe that their assessment instruments are appropriate to their program needs and serve specific purposes of assessment.

Review of the Literature

There are several methods for evaluating assessment tools. Most sources have focused on empirical methods of evaluation—comparing statistically test scores and drawing conclusions from these data (Baker,

1989; Carroll, 1980; Harris, 1969; Hill & Parry, 1989; Oller, 1979; Oller & Perkins, 1978; Palmer & Bachman, 1980). Other sources, and usually including empirical evidence, as well, provide critiques or reviews of available tools and focus on the testing components and processes in nonspecific settings (Alderson, Krahnke, & Stansfield, 1987; Behrens, 1983; Carder, 1990; Johnston, 1987). Resource guides that provide general information on tests for adult education providers are also available (Zellers, 1986).

While these methods of evaluation are valid, necessary, and beneficial, not much consideration has been given to measuring the appropriateness of the assessment tools according to the specific needs of the programs in which they are utilized. According to Hoekje and Linnell (1994), "It is no longer acceptable to have language tests which are statistically valid and reliable but inauthentic in their tasks" (p. 122). Brown (1987) states that "ultimately, validity [of tests] can only be established by observation and theoretical justification" and that we have to "ask questions that give us convincing evidence that a test accurately and sufficiently measures the testee for the particular purpose or objective of the test" (p. 221). According to Brown, "In tests of language, validity is supported most convincingly by subsequent personal observation of teachers and peers" (p. 222). Specifically, the perceptions of the administrators and teachers who work closely with the students whose language proficiency is measured by these tests need to be examined.

Research Methodology

Data for this study were gathered on a questionnaire using Likert scales by asking subjects to respond to various statements by selecting one of five responses ranging from strongly agree to strongly disagree (Babbie, 1989; Henning, 1987). The 24 statements were designed to measure the attitudes of ESL program administrators and teachers regarding the ESL assessment tools utilized by their programs.

During the fall, 1993, the questionnaire was distributed to 93 adult literacy providers throughout Pennsylvania. These providers enrolled at least 15 students each in a program of ESL instruction. A total of 35 programs responded to the questionnaire, a response rate of 38%.

Most of the programs responding served immigrants, and more than half of the programs served refugees, foreign students, and spouses of foreign students. A smaller percentage of the programs served United States citizens. In addition, most of the programs provided classroom

instruction, more than half provided individual tutoring, and a small percentage of the programs provided small group and computer-assisted instruction.

Purpose of Assessment Tools

ESL adult literacy and education programs that provide instruction to adult non-native speakers of English in individualized and group settings utilize a wide variety of proficiency tests and assessment tools for several purposes. These purposes can be organized into three categories: place-ment, diagnosis, and progress evaluation (Alderson et al., 1987; Bell & Burnaby, 1984; Hughes, 1989; Short, 1993).

* Placement: When a student enters an ESL literacy program, an assessment of the student's learning needs helps to place that student in the ESL class or program that will best address these needs.

* Diagnosis: By diagnosing the specific language areas in which the student needs help, an appropriate individual education plan can be created and implemented.

* Progress evaluation: Periodic evaluation of the student's progress can ensure that appropriate teaching methods and approaches are being used with the student, and can document the student's language development.

Some adult literacy programs utilize their assessment tools for all three testing purposes; others utilize them for only one or two. Some programs utilize more than one assessment tool, possibly a separate tool for each testing purpose. The pattern adopted depends on each program's specific functions and needs. For example, a program may utilize a particular proficiency test that would provide enough information about the current English abilities of a learner to place that student in an appropriate class or program. Then, for the purpose of diagnosis, the program might conduct an alternative method of assessment, such as an oral interview, to gather further information about the educational back-ground and immediate ESL learning needs of the student. Finally, a program might use the same proficiency test that was used for placement to evaluate the student's progress in the program. Often this evaluation is conducted both at certain periods throughout the learner's participation in the program and at the time of exit from the program. However, every program is unique in its testing needs, and the assessment tool or tools utilized by each program will reflect this diversity across programs.

According to the survey responses in this study, 88% of the programs utilized their tools for student placement, 65% of the programs utilized their tools for progress evaluation, 50% of the programs utilized their tools for diagnosis, and 6% of the programs reported other uses.

Methods of Assessment

The 32 programs responding to the survey reported using a total of 23 different assessment tools and combination of tools. Standardized tests that are used more frequently include the Basic English Skills Test (BEST; Alderson et al., 1987), the Tests of Adult Basic Education (TABE), the Adult Basic Learning Examination (ABLE), the Wide Range Achievement Test (WRAT; Wegener, 1993), and the Comprehensive English Language Test (CELT).

Some adult literacy programs utilize alternative methods of assessment. These methods include observation and review of work, portfolios, student self-evaluations, performance-based tasks (Short, 1993), Cloze testing, and oral interviews (Zook, 1992). In addition, some assessment tools are created by the programs themselves. Of the 32 respondents in the study, 7 reported using an assessment instrument developed in house. Tools created in house are usually designed around the immediate assessment needs, instructional activities, and student characteristics of a particular program. These tools may use one or more of the methods of assessment noted above.

Analysis

Table 1 displays responses regarding program satisfaction with assessment tool performance for the purpose of placement. Most respondents strongly agreed or agreed (82%) with the statement that their assessment tools provided the necessary information for placing students accurately into the appropriate program or level. This satisfaction was expressed for several component areas within student placement. For example, most respondents strongly agreed or agreed that their tools accurately measured student abilities in English listening comprehension (73%). In addition, most of the programs surveyed strongly agreed or agreed that their tools provided individualized information about a student's language learning needs (83%) and information necessary for identification of immediate survival and functional needs for speaking English (65%).

Table 1

Questionnaire Results - Placement

The assessment tool utilized by our program*:	N	(1)	(2)	(3)	(4)	(5)	Mean	SD
Provides the necessary information for placement into the appropriate program level.	34	26%	56%	3%	15%	0%	2.059	0.952
Accurately measures the student's present abilities to speak English.	34	12%	47%	6%	23%	12%	2.765	1.281
Accurately measures the student's present English writing abilities.	34	3%	32%	3%	38%	24%	3.471	1.261
Accurately measures the student's present English reading abilities.	35	9%	46%	11%	31%	3%	2.743	1.094
Accurately measures the student's present abilities for English listening comprehension.	33	15%	58%	6%	15%	6%	2.394	1.116
Provides individualized information about the student's language learning needs.	35	23%	60%	3%	14%	0%	2.086	0.919
Identifies the student's immediate survival and functional needs for speaking English.	34	24%	41%	6%	18%	12%	2.592	1.354

*Note: Scale: (1) = Strongly Agree; (2) = Agree; (3) = No Opinion; (4) = Disagree; (5) = Strongly Disagree

However, some respondent dissatisfaction was noted. For example, just over half of the programs were satisfied with the performance of their

assessment tools in measuring abilities in speaking English (59%) and reading English (55%). A total of 62% of the programs disagreed or strongly disagreed that their assessment tools accurately measured student English writing abilities.

Table 2 displays the responses for levels of satisfaction within progress evaluation and the individual components of the progress evaluation process. These responses indicate that the programs consider the assessment tools weak in this area.

According to the respondents, a total of only 50% strongly agreed or agreed that their assessment tools provided the necessary information for accurate evaluation of student linguistic progress. In addition, apparent dissatisfaction with several component areas within progress evaluation was revealed. For example, more of the respondents strongly disagreed or disagreed that their assessment tools accurately evaluated student progress in writing English (65%), speaking English (51%), and reading English (44%). A total of only 48% of the respondents strongly agreed or agreed that their tools provided an accuratemeasurement of student progress in vocabulary development. The strongest component within this area was English listening skills; 61% of the respondents strongly agreed or agreed that their tools accurately measured this component.

Table 3 displays the responses regarding program satisfaction with assessment tool performance for the purpose of diagnosis. Just over half the programs strongly agreed or agreed (53%) that their tools provided the necessary information for accurate diagnosis of a student's English proficiency. This percentage was higher than for those programs that strongly disagreed or disagreed (38%). In addition, a majority of programs reported that their assessment tools provided the necessary information for assignment of appropriate instructional materials (71%).

However, as in placement and progress evaluation, the responses regarding specific component areas for diagnosis did suggest weak areas. For example, just over half the programs strongly agreed or agreed that their assessment tools correctly identified the grammatical areas in which a student has the most difficulty (53%). More of the respondents strongly disagreed or disagreed that their assessment tools correctly identified errors or difficulties students may have with English pronunciation (66%) and that they identified student English language goals and interests (50%).

Some of the results within the area of diagnosis were mixed. For example, 41% of the programs strongly agreed or agreed that their assessment tools provided the necessary information for development of

Table 2

Questionnaire Results - Progress Evaluation

The assessment tool utilized by our program*:	N	(1)	(2)	(3)	(4)	(5)	Mean	SD
Provides the necessary information for accurate evaluation of the student's linguistic progress.	32	3%	47%	9%	38%	3%	2.906	1.058
Accurately measures student progress in writing English.	34	3%	21%	12%	41%	24%	3.618	1.155
Accurately measures the student's progress in speaking English.	33	9%	36%	3%	42%	9%	3.061	1.248
Accurately measures the student's progress in reading English.	34	6%	32%	18%	41%	3%	3.029	1.058
Accurately measures the student's progress in English listening comprehension.	33	6%	55%	3%	36%	0%	2.697	1.045
Accurately measures the student's progress in English vocabulary development.	33	3%	45%	6%	39%	6%	3.000	1.118

*Note: Scale: (1) = Strongly Agree; (2) = Agree; (3) = No opinion; (4) = Disagree; (5) = Strongly Disagree

accurate student individual education plans (IEP), while 41% of the programs disagreed with this statement. In addition, 46% of the respondents strongly disagreed or disagreed that their assessment tools provided student educational background information, while 45% strongly agreed or agreed with this statement.

Table 3

Questionnaire Results - Diagnosis

The assessment tool utilized by our program:	N	(1)	(2)	(3)	(4)	(5)	Mean	SD
Provides the necessary information for accurate diagnosis of the student's English proficiency.	34	6%	47%	9%	32%	6%	2.853	1.132
Accurately identifies the student's English language goals and interests.	34	12%	32%	6%	35%	15%	3.088	1.334
Provides student educational background information.	33	9%	36%	9%	30%	16%	3.061	1.298
Provides the necessary information for assignment of appropriate instructional materials.	34	21%	50%	6%	21%	3%	2.353	1.125
Provides the necessary information for development of an accurate individual education plan (IEP) for each student.	34	6%	35%	18%	41%	0%	2.941	1.013
Accurately identifies the grammatical areas in which the student has the most difficulty.	34	12%	41%	9%	35%	3%	2.765	1.156
Correctly identifies errors or difficulties the student may have with English pronunciation.	33	9%	18%	6%	42%	24%	3.545	1.301

*Note: Scale: (1) = Strongly Agree; (2) = Agree; (3) = No opinion;
(4) = Disagree; (5) = Strongly Disagree

Discussion

Although the programs that responded to the survey seemed to agree that they were generally satisfied with the performance of their assessment tools, their responses to questions regarding individual components of the tools (assessing specific skill areas or seeking specific information) suggested otherwise. The responses gave strong indication that specific areas of these tools were lacking in performance and were less than appropriate to meet the needs of the programs utilizing the tools. For the purpose of placement, the deficient component is the assessment of English writing. For progress evaluation, the deficient components include the assessment of progress in writing English, reading English, speaking English, and vocabulary development. Finally, the deficient components in the assessments used for the purpose of diagnosis include the identification of difficulties with English grammar, pronunciation, student language goals and interests, and student educational background information.

Literacy providers, state agencies, and test writers can all use this information to their advantage. For example, literacy providers can become more aware of current assessment practices and examine their own assessment activities by using this study as a model. Individual programs can ask questions such as those displayed in the research questionnaire to evaluate the performance of their own tools. In addition, these programs could go beyond this study and seek other information, such as reasons behind the dissatisfaction that occurs. Some answers to such an inquiry may be shared among adult literacy providers, but others would be unique to individual programs. Context specific information thus acquired would not only help further identify problems, but it would also provide insight for correcting them.

Furthermore, state agencies could also benefit from this study; it could be considered when collecting data for developing standards and quality indicators of assessment. The components of assessment outlined in the study could serve as a model for these purposes. Finally, test creators could use the data collected for this study as a message from ESL adult education providers regarding what these programs seek from assessment tools in order to meet their program objectives and assessment needs. The results indicate that assessment tools used frequently by adult literacy programs are lacking in some essential assessment areas. Further research into the reasons for these perceptions could also be beneficial in developing ESL testing instruments in the future.

Conclusion

The purpose of this study was to determine whether or not providers of ESL adult education programs were satisfied with the assessment tools they use. The findings are crucial to ESL adult literacy program structures and evaluations, especially within the areas of language proficiency assessment and testing. Quality performance and service from these programs is expected in order to ensure continuing funding and support to these agencies. Quality service to the learners being served by these types of programs is most critical.

As demonstrated in this study, it is important to consider the attitudes and perceptions of the actual tool users, in this case ESL adult literacy teachers and administrators, in order to ensure a more complete understanding of the appropriateness of tools for specific uses. Once programs have accurately identified these areas, they can then either adapt their tools to better match their needs or obtain tools which are more appropriate to the specific programs. While some programs and assessment tools may be appropriately matched and working effectively, it is obvious from this study that others are not. Further research is needed to determine why this is occuring and what additional steps can be taken to change this situation.

References

Alamprese, J.A. (1995). Assessment of adult student performance. In T. Reiff (Ed.), *The Pennsylvania adult basic and literacy education staff handbook: 1995 edition* (pp. 20-22). Lancaster, PA: New Educational Projects.

Alderson, J.C., Krahnke, K., & Stansfield, C. (Eds.). (1987). *Reviews of English language proficiency tests.* Washington, DC: Teachers of English to Speakers of Other Languages.

Babbie, E. (1989). *The practice of social research* (5th ed.). Belmont, CA: Wadsworth.

Baker, D. (1989). *Language testing. A critical survey and practical guide.* Kent, U.K.: Edward Arnold.

Behrens, S. (Ed.). (1983). *From the classroom to the workplace: Teaching ESL to adults.* Washington, DC: Center for Applied Linguistics.

Bell, J., & Burnaby, B. (1984). *A handbook for ESL literacy*. Toronto, ON: OISE Press.

Brown, H.D. (1987). *Principles of language learning and teaching* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Carder, M. (1990). Assessing the ESL student. In E. Murphy (Ed.). *ESL: A handbook for teachers and administrators in international schools*. Bristol, PA: Multilingual Matters.

Carroll, B. (1980). *Testing communicative performance*. Elmsford, NY: Pergamon Press.

Fluke, D. (1992). Testing and assessment in Pennsylvania ABLE programs. In T. Reiff (Ed.), *The Pennsylvania adult basic and literacy education staff handbook: 1992 edition* (p. 20). Lancaster, PA: New Educational Projects.

Harris, D. (1969). *Testing English as a second language*. New York: McGraw-Hill.

Henning, G. (1987). *A guide to language testing: Development, evaluation, and research*. Cambridge, MA: Newbury House.

Hill, C., & Parry, K. (1989). Autonomous and pragmatic models of literacy: Reading assessment in adult education. *Linguistics and Education, 1*, 233-283.

Hoekje, B., & Linnell, K. (1994). Authenticity in language testing: Evaluating spoken language tests for international teaching assistants. *TESOL Quarterly, 28*(1), 103-126.

Hughes, A. (1989). *Testing for language teachers*. Cambridge, U.K.: Cambridge University Press.

Johnston, P. (1987). Assessing the process and the process of assessment in the language arts. In J.R. Squire (Ed.), *The dynamics of language learning*. Urbana, IL: National Conference on Research in English / ERIC Clearinghouse on Reading & Communication Skills.

Oller, J.W. (1979). *Language tests at school. A pragmatic approach.* London: Longman

Oller, J., & Perkins, K. (1978). *Language in education: Testing the tests.* Rowley, MA: Newbury House Publishers.

Palmer, A., & Bachman, L. (1980). Basic concerns in test validation. In J.C. Alderson & A. Hughes (Eds.), *Issues in language testing* (ELT Documents No. 111, pp. 135-151). London: The British Council.

Shaw, M. (1993). Indicators of program quality. In T. Reiff (Ed.), *The Pennsylvania adult basic and literacy education handbook for program administrators: 1993 edition.* (pp. 28-29). Lancaster, PA: New Educational Projects.

Short, D. (1993). Assessing integrated language and content instruction. *TESOL Quarterly, 27*(4), 627-656.

Wegener, S. (1993) The four most commonly used assessment evaluations in Pennsylvania for adult students. In T. Reiff (Ed.), *The Pennsylvania adult basic and literacy education handbook for program administrators: 1993 edition* (pp. 48-49). Lancaster, PA: New Educational Projects.

Zellers, R. (1986). *A resource guide of tests for Adult Basic Education teachers.* Harrisburg, PA: AdvancE, Pennsylvania Department of Education.

Zook, D. (1992). ESL for adults. In T. Reiff (Ed.), *The Pennsylvania adult basic and literacy education staff handbook:1992 edition* (p. 32), Lancaster, PA: New Educational Projects.