



## VALID ASSESSMENT OF ENGLISH LANGUAGE ABILITY OF BACCALAUREATE HONORS STUDENTS

*Dr. Douglas E. Trimble*

Director of Assessment and Institutional Research,  
Forman Christian College (A Chartered University), Lahore – Pakistan



Dr. Trimble earned his B.A. in Psychology from Messiah College and his M.A. and Ph.D. in Social Psychology from Temple University in Philadelphia, Pennsylvania, USA. He taught undergraduate psychology at Northwestern College from 1996-2003 and served as Assessment Coordinator. For 5 years he taught both undergraduate and graduate psychology at Eastern University, near Philadelphia. He also was the Coordinator of Student Learning Assessment at Eastern University.

Since August 2008, he has been Director of Assessment and Institutional Research at Forman Christian College (A Chartered University). He has also been appointed Professor of Psychology and teaches Cognitive Psychology with Lab and Research Methods in Psychology. His research interests include the psychology of religion, altruism, prejudice, organizational commitment, and sleep deprivation. He is also interested in issues of reliability and validity of outcome assessment, student engagement, and student satisfaction.



## VALID ASSESSMENT OF ENGLISH LANGUAGE ABILITY OF BACCALAUREATE HONORS STUDENTS

*Dr. Douglas E. Trimble<sup>1</sup>, Shajeel Imran Khokhar<sup>2</sup>*

<sup>1</sup>Director of Assessment and Institutional Research,

<sup>2</sup>Research Associate in Assessment and Institutional Research  
Forman Christian College (A Chartered University), Lahore - Pakistan

### ABSTRACT

#### Purpose

This study examines the validity of an online English assessment test for placing students accurately in Baccalaureate (Honors) English classes at a university in Lahore, Pakistan. Assessment takes place at different levels - individual students, courses, department, and the institution as a whole - so valid techniques are employed appropriately. The university's 3-year cycle of assessing student learning outcomes is reviewed. As part of this program of ongoing assessment, incoming freshmen were administered an online assessment of their English ability offered through the College Board's ACCUPLACER suite of tests ([www.accuplacer.org](http://www.accuplacer.org)). They took a 20-question multiple-choice English grammar test and wrote a 300-600 word essay on one of five general topics that was computer-graded. ACCUPLACER scores determined the level of English class for which students subsequently registered.

#### Methodology

A month into the semester, English faculty were surveyed about the appropriateness of the placements. They were asked whether each of their students was correctly placed in their class or whether they should have been placed in an easier course or in a more challenging class. After the conclusion of the semester, student grades in their English classes were analyzed to see if students succeeded in the class in which they were placed.

#### Findings

According to the faculty survey, 90% of students were accurately placed in the appropriate level class. Analysis of course grades showed that 90% of students earned a grade of C or higher in the course in which they were placed; 65% earned an A or B in the course.

#### Research Limitations

The study is limited as a quasi-experimental design due to the lack of appropriate control groups. A survey was not done the previous year to establish a baseline for satisfaction

among faculty for English placement prior to implementing ACCUPLACER. Also students were not randomly assigned to an experimental and control group. This case study allows us to draw conclusions about the sample but not about other methods of placement into English courses.

## **Value**

ACCUPLACER offers a valid, affordable, and efficient method for assessing student English language ability for course placement. This study models a method of validity checking that is vital for assuring high-quality educational programs.

**Keywords:** ACCUPLACER; quality assurance program; English placement; assessment of student learning; student learning objectives; validity

**Paper Type:** Case study

## **PURPOSE**

One function of quality assurance is to evaluate and improve the validity of student assessments. This study examines the validity of an online English assessment test for placing students accurately in university Baccalaureate (Honors) English classes. Given that students enroll in university with varying degrees of preparation, some come with excellent English skills and others come with less experience. Accurate placement into courses is imperative for the benefit of the student, classmates, and instructor. Students who are not well prepared need additional remediation and are not served by taking a class too advanced for them. They slow down the class and can be a burden to the instructor. Likewise, students who are required to take a class below their ability are delayed in their progress towards attaining a degree and might place pressure on the professor to proceed faster than needed by their peers. So for all parties involved, accurate placement is of great advantage. The current study documents how an English language assessment is embedded within a campus-wide network of student learning assessments. This is a case study of the validity of using ACCUPLACER ESL tests in a Pakistani university.

### **a) Assessment at Forman Christian College (A Chartered University)**

Forman Christian College (FCCU) has a diverse assessment program that is overseen by the Office of Assessment and Institutional Research (AIR). The director of this office is the chair of the Educational Assessment Committee that functions in a similar manner to quality assurance cells in other universities. The purpose of assessment is to document what our students have learned and to adjust programs in such a way to improve learning at the institution (Suskie, 2004). The focus of assessment is not on what faculty do or on the curriculum that is offered – rather, it focuses on students and how they develop. At Forman Christian College, student assessment occurs at many different levels: individual students, courses, departments, and institution-wide.

i) **Assessment of Individual Students**

On an individual basis we use assessment for admissions (Intermediate Board exams; A-levels; National Achievement Test), course placement (ACCUPLACER which is described in detail later in the paper), course grades, competency exams and comprehensive exams. The AIR office closely monitors the reliability and validity of the five competency exams students must pass before graduating: Spoken English, Written English, Written Urdu, Quantitative Skills, and Information Technology. The spoken and written components of the Urdu and English exams are double-marked by two faculty members independently using reliable rubrics (c.f., Stevens & Levi, 2005). Comprehensive exams are given in each discipline during the spring of a student's senior year. Graduating seniors are required to clear an externally-marked exam that has been specifically prepared to reflect learning of the curriculum at FCCU.

ii) **Departmental Assessment**

Academic programs make claims about what their students should know and be able to do at the time of graduation. A major purpose of assessment is to assure that these claims are fulfilled. If they are not, then departments make appropriate adjustments to rectify these deficiencies. Departmental assessment is a cyclical process that follows the following steps:

- (1) Define student learning objectives for program;
- (2) Induce student learning through teaching courses and providing co-curricular programs;
- (3) Assess student learning outcomes through various methods;
- (4) Adjust program – curriculum, pedagogy, or both – to better achieve desired student outcomes; and,
- (5) Re-assess to demonstrate improvement or the need for further adjustment.

Each department has specified about 6-10 students learning objectives that have been published in the university catalogue. They specify the student learning outcomes that are desired for all students upon graduation from the program. The objects are specific enough to be measured and written as student-oriented outcomes (i.e., what should students be able to do upon completion of the program?). They are not written about what the department intends to do or teach but rather what the student will gain from the program in terms of knowledge, skills, and values that are important to practitioners in that discipline. These objectives cover the following six requisite areas:

- (1) Content and theory of the discipline
- (2) Methodology and skills (disciplinary research methods and professional skills, such as writing, speaking, and data analysis)
- (3) Application of knowledge to the real world
- (4) Critical thinking (create, analyze, and evaluate learning in the discipline)
- (5) Values (use knowledge and skills in an ethical manner that reflects the values of the discipline and the Core Values of FCCU: integrity, excellence, respect for human dignity, discipline and accountability, fairness and justice, service, and community)
- (6) Future (What vision does the program have for its graduates in terms of professional involvement, employment or graduate school?)

Each academic department on campus is in the process of developing an Educational Assessment Plan in conjunction with the Educational Assessment Committee. These plans have four required elements:

- (1) student learning objectives;
- (2) assessment techniques;
- (3) budget of estimated costs associated with intended assessment methods; and,
- (4) action plan: what, who, when of administering assessments.

Over a period of three years, academic departments thoughtfully plan, assess, reflect, and report their findings. After feedback from administration they begin the three-year cycle again:

- Year 1: Develop an Educational Assessment Plan.  
Year 2: Assess the program.  
Year 3: Report results to Educational Assessment Committee. This committee prepares an Effectiveness Report for the Deans Council and Board of Studies.
- Year 1: Respond to Effectiveness Report by making warranted changes in the program and by submitting a revised assessment plan.  
Year 2: Assess the program.  
Year 3: Report results.

The advantage of a three-year cycle over annual reporting of assessment results is two-fold: (a) Departmental faculty members have time to think carefully about their department's goals, to develop valid assessments, and to draw considered conclusions; and (b) it allows a period of time to see if changes in the program have had the desired effect. Assessment reporting on an annual basis has the potential to be too rapid for needed deliberation and for modifications to the program to show benefit.

iii) **Institutional Assessment**

The AIR Office conducts assessments that transcend department. By aggregating individual assessments, institutional decisions can be made that are informed by data. For example, by analyzing cumulative grade-point average (CGPA), intermediate board exam results, and National Achievement Test (NAT) scores, prediction models can be developed for student success. This past year, we were able to determine that, at FCCU, a combination of first year intermediate board result and NAT, equally weighted, was as predictive of first year CGPA as was the composite intermediate board exam score that combines first and second year results. This allowed us to make provisional admissions decisions in the spring without having to wait for the second year board results to be announced during the summer.

We also conduct surveys to assess institutional effectiveness. For example, recent alumni are surveyed six months after graduation to see how they have fared in the market place. It has been especially necessary to make sure that employers and programs of higher studies acknowledge the equivalence of the Baccalaureate (Honors) degree with the conventional Master's degree. Both require 16 years of education and have been recognized by the Higher Education Commission (HEC) as providing the same level of preparation.

Another survey is given to graduating seniors to assess their satisfaction. This survey gives the administration a broad-range gauge of how students feel about topics that are important for institutional planning. One a 154-item survey, graduating seniors rated how satisfied they were in 22 categories: Overall Satisfaction, Accounts, Administration, Admissions, Advising, Campus Atmosphere, Campus Facilities, Career Services, Communication, Core Values, Courses, Discipline, Faculty, Financial Assistance, Food Service, Harmony, Hostels, Individual Attention, Registration, Security, Skills, and Student Societies and Activities. We found that 84% of graduating seniors indicated that they were satisfied, with a mean of 4.15 on a 5-point scale. The areas of greatest satisfaction (mean above 4.2) related to the beautiful campus, the development of academic skills (confidence, leadership, and communication in English), and religious harmony. Other areas in which graduating students were satisfied were in the key areas of educational delivery (excellent courses and high quality faculty), vital operations (accounts, admissions, hostels, administration, campus security, and communication) and individual attention given to students. We also found areas in which we need to improve student satisfaction, such as advising, food service, and career services.

**b) ACCUPLACER**

ACCUPLACER is an online suite of tests that have been designed by the College Board for undergraduate course placement ([www.accuplacer.org](http://www.accuplacer.org)). The College Board also offers the SAT, the Advanced Placement (AP) Exams, and the College-Level Examination Program (CLEP). ACCUPLACER tests are offered in English (Sentence Skills; Reading Comprehension; and WritePlacer), English as a Second Language (ESL Language Use; ESL Listening; ESL Reading Skills; ESL Sentence Meaning; and WritePlacer ESL) and mathematics (Arithmetic, Elementary Algebra, and College Level Math). All tests are multiple choice question (MCQ) format except the WritePlacer and WritePlacer ESL which are essay tests. ACCUPLACER's strength is in identifying students who need remediation in English, mathematics, or both and placing them in appropriate classes (Mazzeo, 2002; Nunley et al., 2000).

**i) Computer Adaptive Testing**

The MCQ tests in ACCUPLACER are computer adaptive, which means that each student is administered questions at his or her ability level, thus enabling the program to zero-in precisely on the students' English grammar ability. The first question that the student attempts is of medium difficulty (approximately 50% probability of getting it correct). If is answered correctly, the next question given to the student is of higher difficulty. Conversely, if it is answered incorrectly, the program will select an easier question for the student. This process continues for 20 questions on most ACCUPLACER MCQ tests. Through this process, ACCUPLACER is able to estimate the number of questions that a student would get correct if he or she had been administered 120 questions, thus a students' total is reported out of 120 possible points. This saves students time and frustration in attempting questions that are too difficult for them.

**ii) Validity**

A number of empirical studies have documented the validity of ACCUPLACER data. Factor analyzing ACCUPLACER scores, Cole et al. (2003, as cited by Cole et al., 2005) found a two-factor structure that corresponded to mathematics ability and English ability. The two-factor solution fit better than a single, general-ability factor. James (2006) found that ACCUPLACER was a valid means of placing students in remedial English classes. After taking the Reading Comprehension and Sentence Skills tests, 61.5% of her sample would have been correctly placed by earning a final grade of A, B, or C in English. James found a higher success rate (77.3%) for mathematics courses. James did not directly measure writing because she did not use WritePlacer, so she speculated that her predicative validity might

have been compromised for writing courses. But Sullivan and Nielsen (2009) questioned the benefit of a writing sample beyond that which ACCUPLACER's Reading Comprehension and Sentence Skills provide. They found a .7 correlation between the composite of Reading Comprehension and Sentence Skills and with a locally developed and marked essay, thus making the essay of marginal additional value beyond the MCQ test.

**c) Automated essay grading**

Automated grading of essays has recently become popular due to the development of excellent technologies to do the grading, their reliability and validity, and their ease and speed of use (Dikli, 2006). ACCUPLACER has been using WritePlacer since 2003 to automatically grade essays (Johnson, 2003). WritePlacer and WritePlacer ESL use the computer program Intelligent Essay Assessor (IEA) to instantly grade essays (College Board, 2010). A sophisticated algorithm compares each essay to a bank of 500 previously-graded essays for structure and development of ideas. Using the statistical technique of Latent Semantic Analysis (LAS), the program is able to rate the coherence of the essay's content, so LSA is very well suited to assess the content of an essay (Miller, 2003). The overall score given by Intelligent Essay Assessor to an essay is based on the content of the essay, its style, and mechanics (Landauer et al., 2003). Intelligent Essay Assessor also has mechanisms to flag "unusual content, plagiarism, and other forms of counterfeiting" (p. 299). IES is not without its critics: McGee (2006, as cited by Brown, 2010) found that IEA was "insensitive to sequence of sentences, factual inaccuracies and jumbled word order and mangled syntax" (p. 282). Yet, Landauer et al. reported a correlation between IEA and a human at .81, which is equivalent to the .83 correlation they found between two humans grading the same essays. Miller (2003) reported .83 human-IEA correlation for the *holistic* essay score (content, performance, and style). It was even higher when separately assessing the *content* of the essay ( $r = .90$ ), which is the main strength of LSA. The human-IEA correlations for *performance* ( $r = .68$ ) and *style* ( $r = .66$ ) were lower but acceptable. Thus IEA seems to be mimicking relatively closely human grading, which is itself flawed; yet, this is the standard to which computer assessment ought to be held.

## METHODOLOGY

**d) Sample**

Two thirds of the students were male and most (81.8%) were enrolled as Bachelor of Arts or Bachelor of Science, with 9.6% BS Business and 8.6% Bachelor in Information Technology/Bachelor in Computer Science. The specific type of instruction before enrolling was as follows: Intermediate (87%), A-level (7.7%), Western style high school (GPA, 3.1%), transfer from another university (1.5%), or DAE (technical diploma, 0.7%). About three quarters (76.6%) came from English medium colleges or

schools before enrolling at FCCU, 23% from Urdu medium institutions, and 0.4% from other mediums.

The vast majority of the sample was from Pakistan (98%) with the rest from USA (5 students), Turkey (3), Australia (2), Canada (2), Saudi Arabia, South Korea, and United Kingdom (1 each). Most of the students came from homes where English was not the primary household language: Urdu (71.2%), Punjabi (16.9%), English (3.8%), Pashto (3.6%), Seraiki (1.8), and less than 1% each were Balochi, Burushaski, Khowar, Potwari, Shina, Sindhi and Turkish.

**e) Instrument**

Students were administered the ACCUPLACER ESL Language Use test which assesses English grammar. Each test is composed of 20 MCQs that are computer adaptive and cover five content areas:

- (1) Nouns, Pronouns, Pronoun Case Structure
- (2) Sentence Structure
- (3) Subject-Verb Agreement
- (4) Adverbs, Adjectives, Verbs
- (5) Subordination/Coordination

The ACCUPLACER ESL Language Use test requires students to fill in the blank with the correct word or phrase. A second question format requires them to merge two sentences into one grammatically correct English sentence. The following skills are covered: subject-verb agreement, verb tenses, irregular verb forms, appropriate verb forms in structures, noun-noun agreement, noun forms, pronouns, modifiers, comparatives, prepositions, connectives, parallelism, and sentence fragments/run-ons (College Board, 2010). Scores on ESL Language Use have an estimated alpha reliability of .87 (College Board, 2003).

Students also wrote a 300-600 word essay on one of five randomly-assigned general topics in WritePlacer ESL. Topics are ones that anyone can write on given their life experience. Students had one hour to complete the essay, which was required to be written on the computer. Although James (2008) found that the assigned topic generally did not affect the students' scores on WritePlacer, no similar research has been done on the topics assigned through WritePlacer ESL.

**f) Procedure**

Upon matriculation, students completed two ACCUPLACER assessments of their English language ability. After completing demographic questions, they took the ESL Language Use test and then they wrote an essay for WritePlacer ESL.

ACCUPLACER scores were used to determine the level of English class for which students subsequently registered (see Table 1). Low scoring students placed into

three remedial classes (Reading Comprehension, LANG 101; Basic Writing Skills, LANG 102; Vocabulary Skills, LANG 104) or just one (LANG 102) if their scores were better. Students with sufficient English ability for university-level work placed into a required ENGL 101 (Writing and Grammar) class. However, students demonstrating high ability were able to skip ENGL 101 and take the second required English class (Advanced Writing Skills, ENGL 103).

**Table 1. Course Placements as a Function of Scores on ESL Language Use and WritePlacer ESL**

WritePlacer ESL	ESL Language Use (grammar)		
	0-54	55-99	100-120
0	LANG 101, 102, 104		
1	LANG 101, 102, 104		
2	LANG 101, 102, 104		
3	LANG 101, 102, 104	LANG 102	ENGL 101
4	LANG 101, 102, 104	ENGL 101	ENGL 101
5	LANG 101, 102, 104	ENGL 101	ENGL 103
6	LANG 101, 102, 104	ENGL 101	ENGL 103

English faculty were surveyed about the appropriateness of the placements. They were asked whether each of their students was correctly placed in their class or whether they should have been placed in an easier course or in a more challenging class. All faculty teaching LANG 101, LANG 102, LANG 104, ENGL 101 and ENGL 103 were sent a survey 2-4 weeks into the semester, which read, “Please complete [this survey] after you have had an opportunity to assess formally (e.g., quiz or essay) or informally all the students’ English ability.” They rated each of their students as to whether he or she was correctly placed in the class, not prepared for the class, or over prepared for the class. Eighteen faculty members completed 63 surveys for the students in their courses (most faculty members taught more than one course and up to five courses).

Student grades in their English classes were analyzed to see if students succeeded in the class in which they were placed. The remedial English program was reengineered to create two levels (100 and 110) of integrated English instruction emphasizing reading, writing, listening, and speaking. Cut scores were modified the following year based on the faculty survey (see Table 2).

**Table 2. Revised Course Placements as a Function of Scores on ESL Language Use and WritePlacer ESL**

WritePlacer ESL	ESL Language Use (grammar)				
	0-54	55-83	84-99	100-110	111-120
0	LANG 100				
1	LANG 100				
2	LANG 100	LANG 110			
3	LANG 100	LANG 110	ENGL 101		
4	LANG 100	LANG 110	ENGL 101	ENGL 103	

5	LANG 110	ENGL 101	ENGL 103
6	LANG 110	ENGL 101	ENGL 103

## FINDINGS

According to the faculty survey, on average 90% of students were accurately placed in the appropriate level class.

*Table 3. Placement Accuracy by Course*

	LANG 101	LANG 102	LANG 104	ENGL 101	ENGL 103	Overall
N	205	282	162	213	50	912
% Correctly placed	95	88	99	83	88	90
% over prepared	5	12	1	7	0*	7
% not prepared	0*	0*	0*	10	12	3

\* LANG 101, 102, and 104 are entry-level courses, so students could not be “not prepared” for these courses. Likewise, ENGL 103 is the highest course for which first year students are eligible, so they could not be “over prepared” for this course.

These rates compare favorably with those reported by Corso (2006, as cited by Brown, 2010). Corso reported 17-26% misplacements versus 1-17% in the current study.

Analysis of course grades showed that 86.2% - 95.8% of students earned a grade of C or higher in the course in which they were placed; the weighted mean success rate was 89.9%. Using an A or B in the course as the criterion for success, 51.3% - 87.2% succeeded; on average, almost 2/3 met this even more stringent definition of success (weighted mean = 64.9%).

*Table 4. Likelihood of Course Success as a Function of ACCUPLACER Placement*

Course	N	% A or B	% A, B, or C
Reading Comprehension (LANG 101)	118	78.0	95.8
Basic Writing Skills (LANG 102)	304	60.3	88.4
Vocabulary Skills (LANG 104)	121	51.3	86.2
Writing and Grammar (ENGL 101)	196	75.5	91.8
Advanced Writing Skills (ENGL 103)	47	87.2	95.7
Weighted Mean		64.9	89.9

In their meta-analysis, Mattern and Packman (2009) reported the results of one study that used Sentence Skills and WritePlacer, which are the non-ESL equivalent of ACCUPLACER assessments ESL Grammar Use and WritePlacer ESL, respectively. The present findings are more than 20 percentage points better than that study ( $n = 1704$ ), which had a 69.0% success rate (C or better in the course).

The correlations between final course grade was not significant (see Table 5). The mean weighted correlations that Mattern and Packman (2009) reported in their meta-analysis were uniformly small.

Table 5. Correlation between Course Success (C or higher) and ACCUPLACER Test Scores

<u>Forman Christian College</u>			<u>Mattern &amp; Pakman (2009) Meta-Analysis</u>			
<u>Test(s)</u>	<u>N</u>	<u>Mean Weighted Corr.</u>	<u>Test(s)</u>	<u>N</u>	<u>k</u>	<u>Mean Weighted Corr.</u>
ESL Lang. Use	983	-0.03	Sentence Skills	12485	21	0.13
WritePlacer ESL	983	-0.01	WritePlacer	3408	8	0.13
ESL LU + WP ESL	983	-0.04	SS + WP	1704	1	0.15

N = sample size; k = number of studies in the meta-analysis; ESL = English as a Second Language; LU = Language Use; WP = WritePlacer; SS = Sentence Skills

Thus ACCUPLACER scores do not strongly predict course outcomes. This is however due to the multitude of factors that contribute to success in a university English course beyond pre-term preparation, such as student motivation, course difficulty, validity of instructor's assessments, life events, and random error variance. A better source of validity comes from the faculty survey that indicated that 90% of students were correctly placed and from the 90% who successfully completed their assigned course.

## RESEARCH LIMITATIONS

The study is limited by its quasi-experimental design that lacks appropriate control groups. A survey was not done the previous year to establish a baseline for satisfaction among faculty for English placement prior to implementing ACCUPLACER. Also students were not randomly assigned to an experimental and control group. This case study allows us to draw conclusions about the sample but not about other methods of placement into English courses.

Another limitation is the survey data that faculty members submitted about the appropriateness of course placements. Although we might take pleasure in the finding that 90% of students were correctly placed, an alternative hypothesis is that acquiescence response bias affected the survey responses. In fact, 49% of the 63 surveys indicated that all the students were in the correct class; thus, students either were placed very accurately or faculty were not exercising appropriate discrimination in evaluating their students' readiness. Most of the faculty, however, did show some degree of discrimination with 72% of the 18 faculty members rating at least one of their students as placed incorrectly. This seems to show that most were thoughtfully completing the survey.

## VALUE

Assessment of student learning is best taken in the context of an overarching program to monitor institutional effectiveness. This includes not only the academic program but also

such areas as the accounts department, admissions, advising, athletics, career services, food service, hostels, library, physical plant, religious life, security, and student affairs. For institutions that have inculcated a culture of assessment, monitoring quality becomes second nature and seen as an intrinsically valuable activity.

The model of assessment at Forman Christian College can serve as a case study in effective assessment that is driven by clearly-stated learning outcomes and valid assessment techniques, which in turn inform educational practices for continually improving student learning. Within this framework, ACCUPLACER offers a valid, affordable, and efficient method for assessing student English language ability for course placement. This study models a method of validity checking that is vital for assuring high-quality educational programs. When ACCUPLACER is used for course placement as part of a larger institutional assessment program it can contribute to the quality of undergraduate instruction.

## REFERENCES

- 1) Brown, G. T. L. (2010), "The validity of examination essays in higher education: Issues and responses", *Higher Education Quarterly*, Vol. 64, pp. 276–291.
- 2) College Board. (2003), *ACCUPLACER OnLine Technical Manual*. The College Board, New York, NY.
- 3) College Board. (2010), *ACCUPLACER Program Manual*, The College Board, New York, NY.
- 4) Cole, J. C., Motivala, S. J., Khanna, D., Lee, J. Y., Paulus, H. E., and Irwin, M. R. (2005), "Validation of single-factor structure and scoring protocol for the health assessment questionnaire-disability index", *Arthritis & Rheumatism*, Vol. 53, pp. 536–542.
- 5) Dikli, S. (2006), "An overview of automated scoring of essays", *Journal of Technology, Learning, and Assessment*, Vol. 5 No. 15, available at <http://www.jtla.org> (accessed 23 September 2010).
- 6) James, C. L. (2006), "ACCUPLACER OnLine: Accurate placement tool for developmental programs?", *Journal of Developmental Education*, Vol. 30 No. 2, pp. 2-8.
- 7) **James, C. L. (2008)**, "Electronic scoring of essays: Does essay topic matter?", *Assessing Writing*, Vol. 13, pp. 80-92.
- 8) Johnson, S. (2003), "The write-stuff software", *Discover*, Vol. 24 No. 6, pp. 26.
- 9) Landauer, T., Laham, D., and Foltz, P. (2003), "Automatic essay assessment", *Assessment in Education: Principles, Policy & Practice*, Vol. 10 No. 3, pp. 295-308.
- 10) Mattern, K. D. and Packman, S. (2009), "Predictive validity of ACCUPLACER scores for course placement: A meta-analysis" (Research Report S. 2009-2). New York: College Board.
- 11) Mazzeo, C. (2002), "Stakes for students: Agenda-setting and remedial education", *The Review of Higher Education*, Vol. 26, pp. 19–39.
- 12) Miller, T. (2003), "Essay assessment with latent semantic analysis," *Journal of Educational Computing Research*, Vol. 29, pp. 495-512.

- 13) Nunley, C. R., Shartle-Galotto, M. K., and Smith, M. H. (2000), "Working with schools to prepare students for college: A case study", *New Directions for Community Colleges*, Vol. 2000, pp. 59-71.
- 14) Sullivan, P. and Neilsen, D. (2009), "Is a writing sample necessary for 'accurate placement'?", *Journal of Developmental Education*, Vol. 33 No. 2, pp. 2-4, 6, 8, 10-11.
- 15) Stevens, D. D. and Levi, A. J. (2005), *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning*. Stylus, Sterling, VA.
- 16) Suskie, L. (2004), *Assessing Student Learning: A Common Sense Guide*, Anker, Boston, MA.