## Title

What's in a Test? ESL and English Placement Tests in California's Community Colleges and Implications for US-Educated Language Minority Students



## Authors

Llosa, Lorena
Bunch, George

## Publication Date

2011-01-25


NEW YORK UNIVERSITY

# WHAT'S IN A TEST?

## ESL and English Placement Tests in California's Community Colleges and Implications for US-Educated Language Minority Students

**Lorena Llosa and George C. Bunch**

**January 2011**

Report prepared for the William and Flora Hewlett Foundation

## About the Authors

**Lorena Llosa, PhD,** is project Co-investigator and assistant professor of education at New York University.

**George C. Bunch, PhD,** is project Principal Investigator and assistant professor of education at the University of California, Santa Cruz.

## Contacts

**Lorena Llosa**
Department of Teaching and Learning
New York University
82 Washington Place, 7th Floor Pless
New York, NY 10003
(212) 998-5485
lorena.llosa@nyu.edu

**George C. Bunch**
Education Department
University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95064
(831) 423-8483
gbunch@ucsc.edu

## Suggested Citation

Llosa, L., & Bunch, G.C. (2011). *What's in a test? ESL and English placement tests in California's community colleges and implications for US-educated language minority students*. Report prepared for the William and Flora Hewlett Foundation. Available at http://escholarship.org/

## Associated Reports

Bunch, G.C., Endris, A., Panayotova, D., Romero, M., & Llosa, L. (2010). *Mapping the terrain: Language testing and placement for US-educated language minority students in California's community colleges*. Report prepared for the William and Flora Hewlett Foundation. Available at http://escholarship.org/

Bunch, G.C., Endris, A., & Kibler, A. (Forthcoming, Spring, 2011). Innovative practices for Generation 1.5: Meeting the needs of US-educated language minority students in California's community colleges. Will be available at http://escholarship.org/

## Advisory Board

**Patricia Gándara**, Professor of Education and Co-director of The Civil Rights Project/Proyecto Derechos Civiles, University of California, Los Angeles

**Kenji Hakuta**, Lee L. Jacks Professor of Education, Stanford University

**Linda Harklau**, Professor, Teaching Additional Languages (TESOL & Foreign Language Education), University of Georgia

**Bayinaah Jones**, Executive Director of Research and Institutional Effectiveness, San Jose/Evergreen Community College District

**Mike Kirst**, Emeritus Professor of Education and Business Administration, Stanford University

**Mark Wade Lieu**, Former President, Academic Senate for California Community Colleges, and Professor of English as a Second Language, Ohlone College

**Kathleen "Kitty" Moriwaki**, Vice President, California Community College Assessment Association and Matriculation Assessment Coordinator and ESL Instructor, City College of San Francisco

**Rosa G. Pérez**, Former Chancellor, San Jose/Evergreen Community College District

**David Plank**, Executive Director, Policy Analysis for California Education (PACE)

**Guadalupe Valdés**, Bonnie Katz Tenenbaum Professor of Education, Stanford University

**Andrea Venezia**, Senior Research Associate, WestEd

## Acknowledgments

We are grateful for the help of many in the initiation and completion of our research and the preparation of this report. Our work would have been impossible without major financial support from the William and Flora Hewlett Foundation. Pamela Burdman, our initial program officer at the William and Flora Hewlett Foundation and now Senior Project Director at WestEd, first encouraged us to build on a small pilot study to do a more comprehensive profile of language testing and placement relevant to language minority students in California community colleges. We have appreciated Pam's unwavering enthusiasm and support throughout the project and, upon Pam's departure from Hewlett, Denis Udall's interest and continued support.

Ann Endris, project manager and research specialist at the University of California, Santa Cruz, made significant contributions to the conceptualization of this project, provided helpful feedback throughout the research and writing process, and provided the logistical support to keep us moving forward. We are also grateful for the assistance of our research assistants at New York University: Kimberly Woo, Cecilia Guanfang Zhao, and Beth Clark-Gareca. At UC Santa Cruz, research assistant Erin Maravilla helped with the formatting of the report.

We are particularly grateful for the assistance of our advisory board members, listed on the previous page. Representing deep and broad experience and expertise related to the testing, placement, and education of language minority students in community colleges, board members engaged substantively in our work, meeting with us, reading drafts of our report, and providing critical guidance at several important stages of our research and writing.

Finally, this research would have been impossible without the input of faculty members, counselors, staff, and administrators in the California community college system. Although we cannot identify these individuals by name due to promises of anonymity, we are grateful to participants at colleges throughout the system who agreed to be interviewed for this research; district and systemwide officials who provided background information and context; and participants at numerous conferences throughout the state who responded to our early findings and helped us better understand the issues at hand.

In acknowledging all of the help, support, and guidance we have received from others, it is also important to clarify that the views we express in this report do not necessarily reflect those of the Hewlett Foundation, our advisory board, or the research participants. Furthermore, we alone are responsible for any shortcomings associated with the report.

## Table of Contents

# List of Tables and Figures

**TABLES**



**FIGURES**

## INTRODUCTION

This report focuses on ESL and English placement tests used in California community colleges and the implications of using these tests with students from language minority backgrounds who have attended US K-12 schools and seek to continue their studies in community colleges. US-educated language minority (US-LM) students, who vary in their level of bilingualism in general and English proficiency in particular, are sometimes called "Generation 1.5" because they have linguistic backgrounds and characteristics that differ from native speakers of English and from international students, recent immigrants, and older adult immigrants (Harklau, Losey, & Siegal, 1999; Roberge, 2002; Roberge, Siegal, & Harklau, 2009).[1] The larger population of community college students speaking languages other than English include those who were born in the US to immigrant parents, immigrants who arrived at some point during their K-12 schooling, immigrants who arrived in the US as young adults after secondary schooling in their countries of origin, older immigrants who are returning to school to learn English or pursue vocational training, and students on international visas planning to return to their home countries after studying in the United States. Although there is no official classification of students from language minority backgrounds in community colleges as there is in the state's K-12 schools, it has been estimated that students from immigrant and language minority backgrounds collectively represent over 25% of the 2.5 million California community college student population (Woodlief, Thomas & Orozco, 2003). Given the fact that 40% of California's K-12 students comes from homes where languages other than English are dominant, and one in four K-12 students is classified as an English Learner (EL)[2], it is likely that a sizable portion of community college students are US-LM students.

Focusing on the constructs and characteristics of the placement tests most commonly encountered by US-LM students in California's community colleges, the purpose of this report is both to inform decisions made at local colleges regarding what tests to use for US-LM students and to contribute to policy discussions at the state- and system-levels regarding placement testing. Specifically, we address the following questions:

1. What ESL and English placement tests do US-educated language minority students most commonly encounter in California community colleges, and what are these tests designed to measure?

---

[1] Given the tendency of the term Generation 1.5 to emphasize students' linguistic and experiential deficits, we prefer to use the more descriptive term *US-educated language minority students* (see Bunch et al, 2010).

[2] In California K-12 schools, students who come from homes where English is not the dominant language and who are determined to be in need of language instruction before they can engage in mainstream academic instruction in English are classified as English Learners (ELs).

2. What are the implications of using these tests for the placement and instruction of US-educated language minority students in California community colleges?

This report accompanies a broader description of the testing and placement policies and practices relevant to US-LM students in California's community colleges (Bunch, Endris, Panayotova, Romero, & Llosa, 2011).

In community colleges, the language testing and placement process is one of the first aspects of higher education encountered by language-minority students transitioning from U.S. high schools. This process represents high stakes for students' academic trajectories, as it determines whether students have access to credit-bearing English courses required for graduation or transfer to four-year institutions or will be assigned to developmental English or English as a Second Language (ESL) courses that often do not earn credits toward a degree or transfer. Ideally, the placement process identifies what students are able to do in English and steers them toward the instructional environment that holds the most promise for them to complete their academic goals.

Yet, despite the stakes involved in community college placement testing, research is only beginning to examine the placement process, the tests that are at its core, and how this process affects language minority students in particular. Focusing on the general student population, several recent reports have highlighted limitations concerning the effectiveness of community college remedial education generally (e.g. Bailey, 2009; Bailey, Jeong, & Cho, 2010) and the most commonly-used placement tests in particular (Hughes & Scott-Clayton, 2010). Recent research has also revealed that insufficient information is available at many colleges for students, faculty, and staff regarding the placement tests themselves and how they are used in the placement process (see Bunch, 2008, 2009; Bunch & Panayotova, 2008; Bunch et al., 2011; Venezia, Bracco, & Nodine, 2010).

Selecting appropriate tests, deciding on cut scores, and evaluating the appropriateness of the decisions made are complex issues requiring time, expertise, and research capacity, yet community colleges often have limited resources to engage in these activities effectively. Although testing and placement practices at some colleges are the result of deliberative processes by faculty and matriculation personnel, at other colleges there is much less awareness about the nature and uses of the particular placement tests used (Bunch et al., 2011). At some colleges, ESL and English instructors cannot name the placement test at their college, at others they provide conflicting information, and at others even matriculation staff are unaware of the content or nature of the tests used. Meanwhile, decisions regarding what tests each college should use are often based more on logistical and financial bases than on decisions regarding what should be tested. Perhaps because of the perceived technical nature of tests and testing (Bachman & Palmer, 1996), or because of a general "faith" in tests to do their job (Valdés & Figueroa, 1994), stakeholders often delegate responsibility for test selection and use decisions to

a single test director at a college, or even in some cases to outside consultants, and are not themselves informed about the tests used in their own college and what those tests measure (Bunch et al., 2011).

Meanwhile, many stakeholders, including students, faculty, staff, counselors, administrators, and policymakers have misperceptions about what placement tests are designed to measure. Many believe that language placements tests can measure overall levels of English proficiency, but do not realize that "English proficiency" is not an agreed upon construct. Instead, the meaning of language proficiency depends on how a particular test or testing program has chosen to define it. What language proficiency a test measures also depends on how a particular college chooses to use these tests. There are a number of decisions that are made locally at each college that can affect the constructs that a placement test measures. These decisions include, among others, choosing which tests in the battery to administer; deciding which students should take which tests (or allowing students to decide), making decisions regarding how computer-adapted tests "branch" from one test or subtest to another based on responses to background questionnaires and performance on individual tests; selecting essay topics and whether the writing is timed or untimed; and setting cut-scores. All of these local decisions about how to use the test, along with the characteristics and constructs of the tests themselves, can affect the inferences that can be made about test takers' language ability. This in turn has implications for the placement and instructional decisions that are made on the basis of that information.

The relationship between what a test measures and its usefulness for placement purposes is further complicated when the test taker population is comprised of US-LM students, who have both unique and diverse characteristics. US-LM students, by virtue of the fact that they are in the process of developing English in addition to a home language, share some characteristics with more recently-arrived immigrants, older adults, and international students, the traditional population served by community college ESL programs. On the other hand, because US-LM students have learned English primarily by using it in naturalistic contexts in US communities and schools rather than by studying the formal features of the language in second-language grammar classrooms abroad, there are also differences between US-LM students and more traditional ESL students (Valdés & Figueroa, 1994). In some cases, students' perceived second language "errors" may actually be features of non-prestige dialects of English. US-LM students' challenges with academic language and literacy may be quite similar to those faced by the monolingual, English speaking population.

We begin by overviewing the testing and placement process that US-LM students undergo in California's K-12 and community college systems. Then, we provide a brief framework for understanding language testing and why it is important to consider what tests measure and how they are used. This framework provides a background against which the subsequent description of the placement tests can be discussed in order to raise awareness and

foster conversation regarding best ways to assess and place language minority students in order to ensure that they receive the most appropriate instruction to facilitate their academic success. Next we describe and compare the constructs and characteristics of the ESL and English placement tests most commonly used in California community colleges to highlight the various ways in which the construct of English proficiency is operationalized. Finally, we discuss implications for the use of these assessments as part of a process for making placement decisions and how the tests and the placement process might affect language minority students in particular.

## Language Minority Students: Testing and Placement in California's K-12 Schools

In California K-12 schools, 25% or 1.6 million students are considered English learners (EL) (Gándara, Maxwell-Jolly & Driscoll, 2005). In this context a formal system exists that identifies English learners in order to provide them with appropriate language services. Figure 1 illustrates this process: When a parent first enrolls her child in the district, she completes a home language survey. If the language spoken at home is English, the child is classified as an English Only (EO) student. If the language is other than English, the student is tested for language proficiency using the California English Language Development Test (CELDT). If the student is rated as proficient, she is designated as Initially Fluent English Proficient (IFEP or FEP). If not, she is classified as an English learner and placed into 1 of 5 English Language Development (ELD) levels. Thus, an EL is a student whose home language is not English and who is identified as not having sufficient English proficiency to successfully participate in a mainstream English program. As the student gains proficiency, she will move through the ELD levels until she is Reclassified as Fluent English Proficient (RFEP). An RFEP student is expected to have sufficient mastery of English to succeed in mainstream courses and to meet academic standards in grade-level content areas in English. Reclassification decisions are made locally by school districts and are typically based on multiple indicators that address the following areas: 1) basic language proficiency standards as measured by the CELDT, 2) teacher evaluation, 3) academic achievement standards as measured by the California Standards Tests in Language Arts, and 4) consent or notification of a parent or guardian (California Department of Education, 2009a).

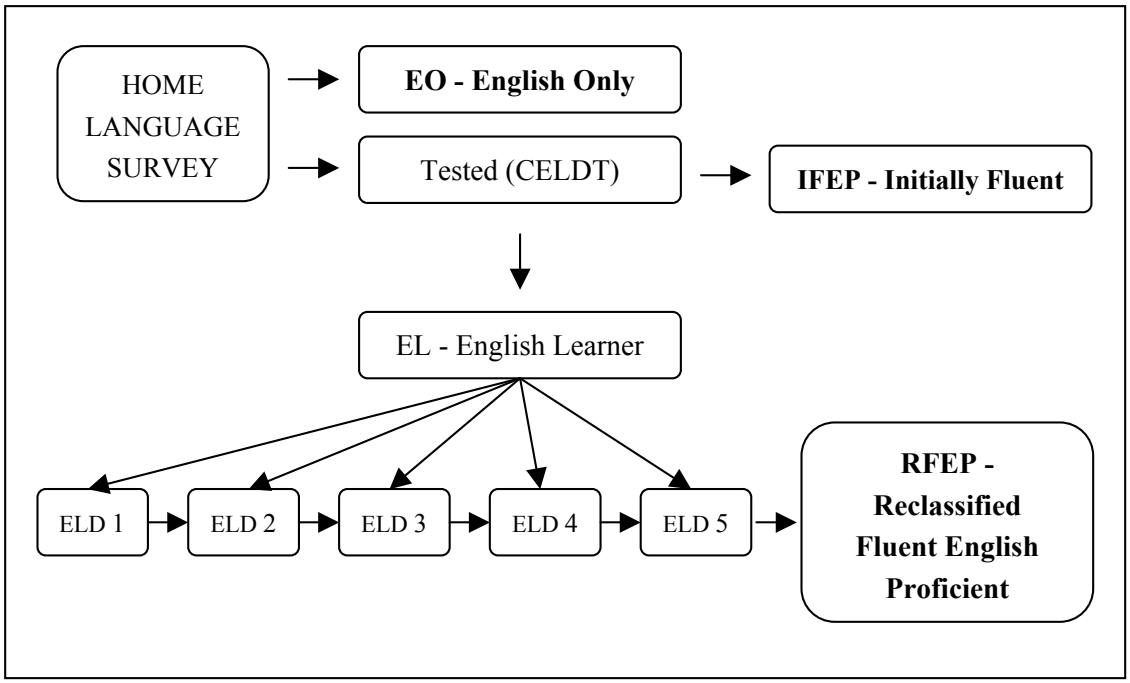


*Figure 1*. The Language Classification Process in K-12 California Schools

At the center of this process for determining whether a language minority student is classified as an EL or deemed ready for mainstream classrooms is the CELDT. The CELDT is a standardized test designed to assess the English proficiency of ELs in California. The CELDT is aligned to the California English Language Development (ELD) standards. All students whose home language is not English according to the home language survey must take the test within 30 calendar days after they are enrolled in a California public school for the first time to determine if they are ELs. The CELDT must also be given once a year in the fall to ELs until they meet the criteria for reclassification as Fluent English Proficient.

Following the organization of the California ELD Standards, the CELDT covers four skill areas: Listening, Speaking, Reading, and Writing. Results are reported according to five proficiency levels—beginning, early intermediate, intermediate, early advanced, and advanced—for each skill area and overall. Appendix A provides a detailed description of the characteristics and constructs assessed by the CELDT as well as a comparison between the CELDT and the ESL community college placement tests. This comparison sheds important insights into the level of alignment in terms of language expectations across these two educational contexts.

The CELDT, however, is just one of the many tests that language minority students encounter in K-12. Language minority students are also subject to the California Standards Tests (CSTs), given to all students, used to measure their mastery of state standards. In spring 2009 in California, all students were required to take the following CSTs: English–language arts (ELA)

for grades two through eleven; Mathematics for grades two through nine; Science for grades five, eight, and ten (life science); and History–social science for grades eight and eleven (U.S. history). In addition, all students must also take and pass the California High School Exit Exam (CAHSEE) in order to receive a high school diploma (California Department of Education, 2009b).

Although the goal is to reclassify ELs as soon as they are ready, many language minority students are still classified as EL by the time they finish high school. Many of these students, those still considered ELs and those already reclassified, turn to the community college system to continue their education. As discussed earlier, these US-LM students, sometimes referred to as Generation 1.5, often do not "fit" the traditional profile of students that community colleges typically serve: they are neither monolingual native speakers of English nor are they international students or ESL adult education students. Depending on their prior education, age of arrival, and other factors, they may be fully proficient in English or they may be at different levels of English language development.

In the context of higher education, the language characteristics of these US-LM students, or Generation 1.5, are usually discussed in comparison to the language characteristics of international students, since for both groups English is not their first language. In general, Generation 1.5 students are believed to have stronger speaking and listening abilities in English and greater knowledge of American colloquialisms, popular culture, and the US educational system than international students. Generation 1.5 students are also believed to have weaker metalinguistic knowledge of grammar than international students given that they acquired English through immersion in casual settings rather than through formal language instruction (Harklau, Losey, and Siegal, 1999).

These characterizations of the language of Generation 1.5 or US-LM students reflect trends for groups of students but may not apply to any one individual student. In fact, Valdés & Figueroa (1994) caution against grouping these students "using one or two key variables, such as first language learned or language spoken in the home as criteria" (p. 18). The authors have described US-LM students as "circumstantial bilinguals" who, unlike elective bilinguals, need to acquire the majority, prestige language in order to succeed in US academic settings. Valdés & Figueroa have also highlighted the complexity of attempting to characterize this population:

> " . . . individual bilingualism that results from the real use of and experience with two languages is highly complex and variable. Although at the macrolevel one may be able to generalize about group tendencies or experiences, at the microlevel one cannot make assumptions about the relative strengths and proficiencies of a bilingual's two languages based on one or two factors about his background and experiences. Factors such as: (a) language spoken in the home, (b) age of arrival in the United States, (c) first language spoken, and even (d)

language used most frequently can predict little about a bilingual's relative strengths in each language" (p. 20).

Thus, while it is important to highlight some general characteristics that distinguish US-LM students from other groups of community college students, it is also necessary to keep in mind the fact that the label encompasses a wide range of students with various linguistic and other characteristics and is by no means a homogenous group.

## Language Minority Students: Testing and Placement in California's Community Colleges

In contrast to the K-12 system described above, where there is a clear mechanism for identifying, testing, and placing language minority students into ESL or mainstream programs, the mechanisms for identification, testing, and placement of this population in California community colleges vary markedly from one college to the next (Bunch et al., 2011). At the community college level, K-12 designations (e.g. ELL vs RFEP), are rarely used to inform placement, either because that information is unavailable or because community college personnel are not familiar with it. In many community colleges students self-identify as ESL or not-ESL and take the ESL placement test or the English placement tests based on their self-identification. Students are then placed in an ESL or an English program depending on the placement test they chose to take.

In other colleges, students are tracked to one test or the other often without their knowledge (Bunch et al., 2011). Some colleges give students a questionnaire at the beginning of the test to gather multiple measures information but also to decide which test (ESL or English) students will take based on their responses to questions about their language and educational background. Many colleges take advantage of the commercially-available placement batteries' branching mechanisms to make decisions about identification and testing of language minority students. In some colleges, students are identified and tested based on their performance on an initial test, not based on their background. In these colleges all students are given the same initial test, typically one of the ESL tests, and based on their performance they are branched into more ESL tests or the English tests.

In a report aiming to "map the terrain" regarding the variability in placement and testing policies and practices relevant to language minority students in California community colleges, Bunch et al. (2011) discuss the overall placement and matriculation process, including measures besides test scores used to determine student placement, that leads to placement decisions. The one element in the placement process that is consistent across the vast majority of colleges is the use of placement tests.

Most California community colleges use commercially available tests to place students into their ESL and English course sequences. In fact, according to the California Community College Consultation Council Task Force on Assessment (2008), 86 of the then-109 colleges used commercially developed tests for ESL placement, 14 used a non-writing homegrown test, and 8 used some form of self-assessment. Meanwhile, 31 colleges used a writing sample to inform placement, often in conjunction with another test. Of those who used commercially developed tests, 47 used the Combined English Language Skills Assessment (CELSA), 22 used the Computer-adapted Placement Assessment and Support Services (COMPASS), 17 used ACCUPLACER, and 3 colleges used other commercially developed tests. Similarly, in Fall 2006, 80 colleges were using a commercially developed placement test for writing placement and 91 colleges were using these tests for reading placement into their English course sequence. The most commonly used commercial placement tests were ACCUPLACER (37 colleges reported using it for writing and 46 for reading); COMPASS (22 for writing and 23 for reading) and the CTEP (15 for writing and 16 for reading) (Consultation Council Task Force on Assessment, 2008).

Given the prevalence of placement tests in the placement process across colleges, and the predominant use of commercially available tests in particular, this report focuses on the characteristics and constructs assessed by the three most commonly used commercially available ESL and English placement tests. However, the issues addressed in the detailed discussion of how these tests differentially define and measure the construct of English proficiency are very much relevant to homegrown tests and writing samples used in many colleges.

## What's in a Test?

Contrary to common perceptions that language tests can accurately measure the test taker's "language proficiency" as a whole, tests are designed for specific purposes, and different kinds of tests will measure different kinds of abilities test takers have to use language in different contexts and for different purposes. What a test measures, also known as the test *construct*, is critical because it provides the basis for interpreting test scores. The decisions made on the basis of those score interpretations are equally important. According to Bachman and Palmer (1996) language tests are given because certain decisions need to be made. In the case of the tests reviewed in this report, community colleges administer these tests to make decisions about course placements. Bachman and Palmer explain that if we want to use the scores from a language test to make interpretations about individuals' language ability, and make various types of decisions, we must be able to demonstrate how performance on that language test is related to language use in specific situations other than the language test itself. They refer to a situation outside of the assessment itself in which test takers need to use language as a target language use (TLU) domain. A TLU domain includes the tasks, outside of the assessment itself, that test takers will need to perform using language. In the case of

community college placement tests, the TLU domain to which interpretations might be generalized could be ESL courses, English courses, general education requirements or courses required for a particular major or professional certificate.

Bachman & Palmer (1996) explain that it is important to consider what tests measure (constructs), and how they do so (characteristics), because these directly relate to two important qualities of test usefulness: construct validity and authenticity. Construct validity pertains to the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores. Authenticity refers to the correspondence between the characteristics of the test task and the characteristics of the TLU domain. As Figure 2 shows, a test taker's performance on a language test will depend on two main factors: their own language ability and the characteristics of the task (e.g. whether the test is multiple choice vs. constructed response; oral interview vs. tape-mediated). Assuming clear and reliable scoring criteria and procedures (also considered a task characteristic), the score obtained should accurately reflect the test taker's test performance. According to Bachman and Palmer (1996), the "test scores are to be interpreted appropriately as indicators of the ability we intend to measure with respect to a specific domain of generalization" (p. 21). The construct definition and the characteristics of the test task are of critical importance to the interpretation of test scores.

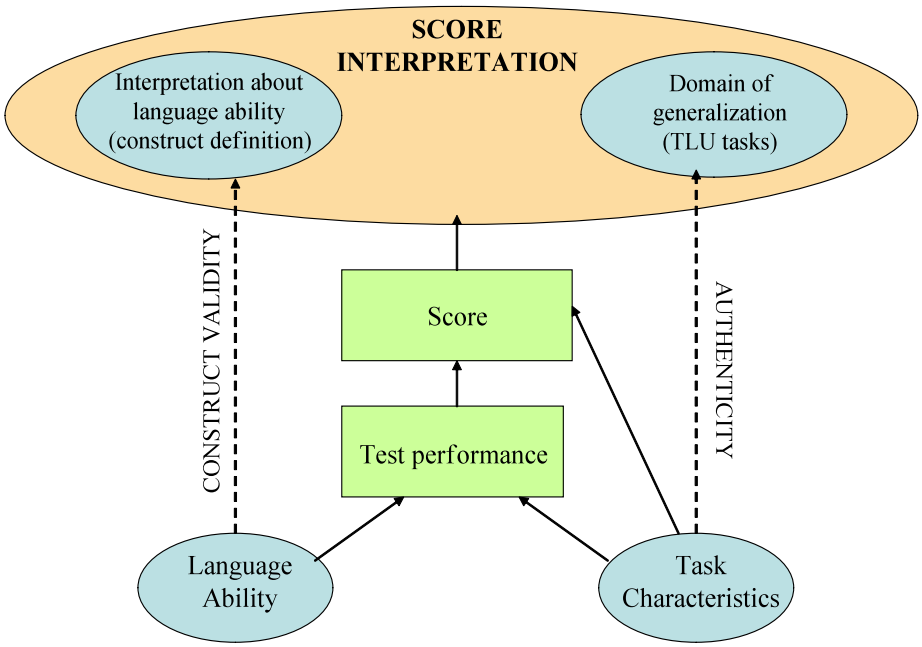


*Figure 2*. Score Interpretation (Bachman & Palmer, 1996)

The higher the correspondence between the language and the task characteristics assessed by the test and the language and the task characteristics in the TLU domain, the better and more useful the interpretations that we make on the basis of test takers' performance will be and the more generalizable to the TLU domain. By examining what the placement tests measure, that is, the aspects of language ability that they measure, community colleges can then examine the correspondence between the constructs of the tests and the constructs focused on in their courses and thus the extent to which scores based on these tests will generalize to student performance in these course. This is important because language ability is a complex construct and no test can measure it all.

Bachman and Palmer's (1996) model of language use presented in Figure 3 illustrates the complexity of language use.

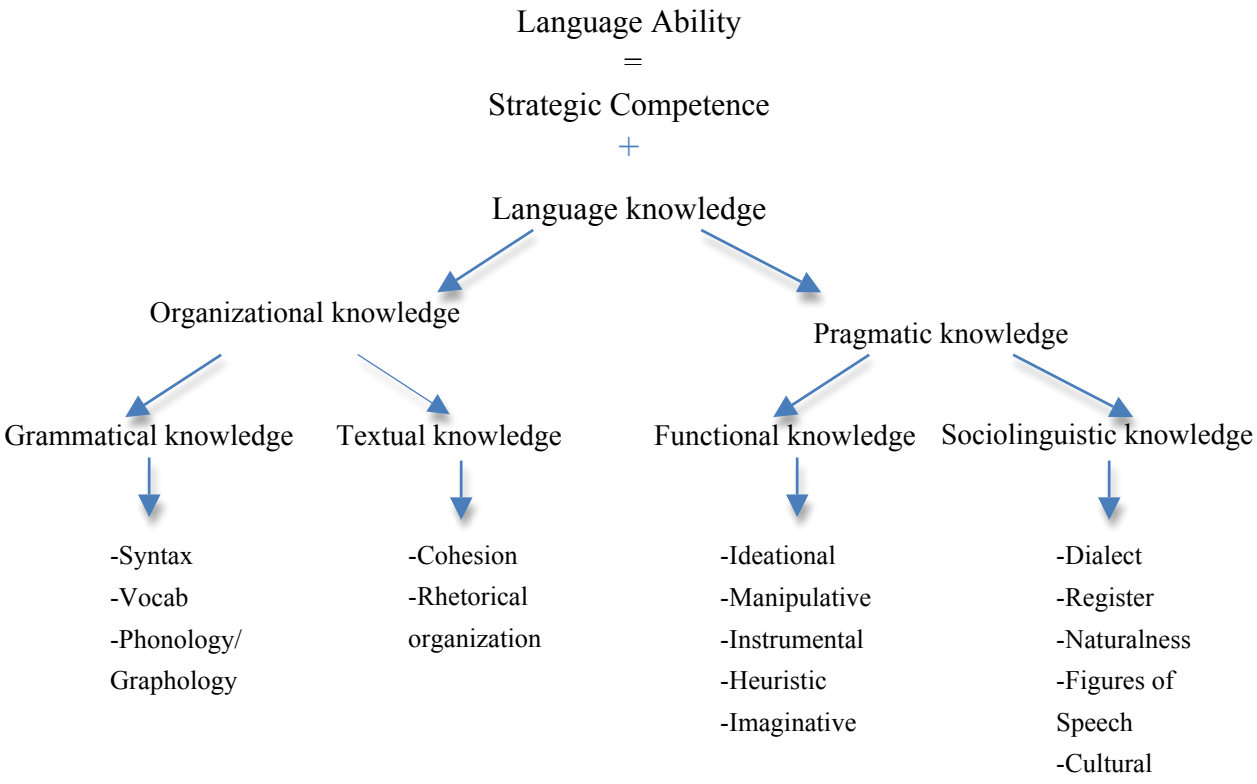


*Figure 3*. Language Ability (Bachman and Palmer, 1996)

Bachman and Palmer (1996) explain that language use involves both language knowledge and strategic competence. Language knowledge consists of two broad categories: organizational knowledge and pragmatic knowledge. Organizational knowledge, which includes knowledge of grammar and rhetorical organization, refers to the control of the formal elements of language. Pragmatic knowledge, which includes functional and sociolinguistic knowledge, enables language users to create or interpret discourse by relating sentences to their meaning, to the intentions of language users, and to relevant characteristics of the language use setting. As Valdés & Figueroa (1994) conclude from their review of various models of language ability including an earlier version of this one (see Bachman, 1990), knowing a language "goes much beyond simplistic views of good pronunciation, "correct" grammar, and even mastery of politeness. Knowing a language and knowing how to use a language involves a mastery and control of a large number of interdependent components and elements that interact with one another and that are affected by the nature of the situation in which the communication takes place" (p. 34).

As the next sections will show, each commercially developed test represents a very specific definition of the construct of language ability. Colleges using these placement tests

should be aware of what aspects of language ability these tests measure because the interpretations made about students' language ability will be specific to the way in which these tests have defined the construct of language ability and the tasks that they use to measure it.

## CALIFORNIA'S COMMUNITY COLLEGE PLACEMENT TESTS

## AND WHAT THEY MEASURE

This section reviews and compares the constructs and characteristics of the most commonly used commercially developed ESL and English placement tests: CELSA, CTEP, ACCUPLACER and ACCUPLACER ESL, COMPASS and COMPASS ESL. A large number of publicly available documents and websites were examined to gather information about the tests. For the CELSA, we had access to a complete package that included the two forms of the test, the test administrator's manual, and the technical report, provided by its publisher, the Association of Classroom Teacher Testers (Association of Classroom Teacher Testers). The two COMPASS batteries and the two ACCUPLACER batteries are both computer-adaptive and thus there is no fixed form of the test. ACT, Inc. and the College Board, the respective publishers of these test batteries, have informative websites and documents with detailed descriptions of all of the tests as well as sample items. The CTEP does not have an official website, so material for this review was gathered from Southwestern College (http://www.swccd.edu/Pdfs/CTEPSampleFormforStudents.pdf). We also contacted ACT and The College Board to request the technical manuals for the COMPASS and ACCUPLACER batteries but were unsuccessful.

Although the Consultation Council Task Force on Assessment survey identified COMPASS and ACCUPLACER English and ESL among the most commonly used commercially-available test batteries, in actuality only the Reading and Grammar tests in the batteries are used for placement purposes in California community colleges. The Listening and Essay Writing tests tend not to be used. This may be because of additional costs associated with the use of these tests and because the Chancellor's list of assessments approved for statewide use do not include the ESL listening and writing tests associated with either ACCUPLACER or COMPASS. Thus, in this report, we focus on the Reading and Grammar tests in the placement batteries, but we have also included descriptions and comparisons of the Listening and Writing tests in Appendix B for those interested.

In the following section, we review the most commonly used ESL placement tests and compare them to each other. In the subsequent section, we review and compare the most commonly used English placement tests. In each section, we first provide a general description of the assessments, including their purpose, intended population, and overall structure. Next, we present the main findings resulting from the comparisons of the tests. For those interested, this summary of findings is followed by a detailed description and comparison of the tests/batteries organized by language skills (e.g. reading, and grammar). Finally we compare and contrast the ESL tests with the English tests.

## ESL Placement Tests

The CELSA, COMPASS ESL, and ACCUPLACER ESL are the three most commonly used commercially developed tests used in California community colleges for ESL placement. The CELSA was developed for the purpose of placing adult learners of English as a second language into different language ability levels in ESL courses. Based on CELSA results, students are placed into seven levels of proficiency from low beginning to advanced plus (http://www.assessment-testing.com/cfaqs.htm). Despite having been originally developed for the adult community college population, it is the most widely used placement test for ESL programs due to its low cost and ease of administration. Some colleges use it in conjunction with a writing sample. The CELSA consists of two parallel forms with 75 multiple-choice items each. According to the Association of Classroom Teacher Testers website, the CELSA can be used with students in open enrollment adult education, community college, university, and high school English as a second language or foreign language programs (i.e., adults and high school students in programs using ESL/EFL materials) (http://www.assessment-testing.com/cfaqs.htm).

Unlike CELSA, both the COMPASS ESL, published by ACT, Inc. and ACCUPLACER ESL, published by the College Board, are test batteries consisting of different component tests, which can be administered individually or collectively. Except for the essay writing tests, they are computer-adaptive, Internet-delivered, and computer-scored. According to the ACT website, "ACT designed the COMPASS English as a Second Language (ESL) placement tests to help postsecondary institutions quickly and accurately assess incoming ESL students' English language ability levels and place them into appropriate ESL courses" (ACT, 2007, p.3). The test taker population identified by COMPASS ESL is "incoming ESL students." The COMPASS ESL battery is composed of four tests: Reading, Listening, Writing (E-Write Essay), and Grammar/Usage. However, only the Reading and the Grammar/Usage tests are used for placement in California's community colleges.

According to the ACCUPLACER ESL's current website, its purpose is "to place students of limited English proficiency in appropriate language courses. ACCUPLACER ESL was developed as an assessment tool for English as a Second Language (ESL) students and English as a Foreign Language (EFL) students (http://professionals.collegeboard.com/higher-ed/placement/accuplacer/english). In the ACCUPLACER Coordinator's Guide, however, the intended test taker population is described as "students who have learned English as a second language or who are native English speakers with limited proficiency" (The College Board, 2007a, p. 5 and p. 15). And yet in that same document, the intended population for the Listening test is described as "non-native English speaking students" (p. 5). It is unclear why the intended test-taker population is not described consistenly, what is meant by native English speakers with limited proficiency, and who the actual intended test taker population is. Perhaps as a result of this confusion, the current ACCUPLACER website does not mention "native English speakers with limited proficiency". Nonetheless, it is important to point out that some colleges do use the

ACCUPLACER ESL (formerly called the LOEP) as a placement test for native-English speakers, raising questions about the validity of the test for this population.

ACCUPLACER ESL offers a total of five tests: Reading Skills, Listening, Essay Writing (WritePlacer), Sentence Meaning, and Language Use yet only the Reading Skills, Sentence Meaning, and Language Use tests are used in colleges for placement. In 2009, ACCUPLACER released a new platform, ACCUPLACER i3, with improved features. With the exception of WritePlacer, the tests have remained the same but "the questions have been reviewed and refreshed where necessary" (The College Board, 2009, p.2).

Table 1 below presents a side-by-side representation of the three ESL placement tests.

Table 1
*Overview of the Three Most Commonly Used ESL Placement Tests*

<table>
<tr><th>CELSA</th><th>COMPASS ESL</th><th>ACCUPLACER ESL</th></tr>
<tr><td rowspan="4">Reading/Grammar<br>-understanding of meaning in a context, as well as grammatical ability. Measures language ability in a holistic manner.</td><td>Reading<br>- reading explicitly stated material<br>- inferential reading</td><td>Reading Skills<br>-comprehension of short passages</td></tr>
<tr><td>Listening<br>-the ability to understand standard American English</td><td>Listening<br>-the ability to listen to and to understand one or more persons speaking in English</td></tr>
<tr><td rowspan="2">Grammar/Usage<br>-sentence elements<br>-sentence structure and syntax</td><td>Sentence Meaning<br>-understanding of word meanings in one-or-two sentence contexts</td></tr>
<tr><td>Language Use<br>-grammar and usage</td></tr>
<tr><td>E-write<br>-direct measure of writing skills</td><td>WritePlacer ESL<br>-direct measure of writing skills</td></tr>
</table>

## Main Findings: ESL Placement Tests

All three tests, ACCUPLACER ESL, COMPASS ESL, and CELSA, are designed to measure English language proficiency for non-native speakers of the language. However, this construct is operationalized differently in each of these tests. Thus by choosing one of these tests, community colleges are also choosing, either consciously or unconsciously, to measure and thus privilege certain aspects of English language proficiency.

CELSA measures very broadly defined aspects of language ability, "reading and grammar in a context," but no information is provided as to specifically what aspects of reading and grammar are assessed. CELSA is composed of 75 multiple-choice questions combining three formats of cloze items (rational, fixed ratio, and multiple choice) into one test (Association of Classroom Teacher Testers, 2000, p. 6). According to the Test Administrators' Manual "its integrated format provides a global measure of language proficiency" (p. 7). Writing, speaking and listening skills are not assessed on the CELSA. According to the ACTT, "each passage [in CELSA] involves common situations experienced by students and encountered in most teaching materials" (http://www.assessment-testing.com/cfaqs.htm). However, the teaching materials probably refer to adult ESL teaching materials. None of the topics are academic in nature. Passages on the test are scaled at the beginner, intermediate, and advanced levels. The test-administration takes approximately 60 minutes in all: 15 minutes are dedicated to the delivery of instructions and 45 minutes for the test itself. Overall, although it is the most commonly used ESL test in California community colleges (Consultation Council Task Force on Assessment, 2008), there is minimum correspondence between the language of the CELSA and the language of the TLU domain that would seem to be relevant for language minority students in community colleges interested in pursuing academic goals (Bunch & Panayotova, 2008).

The COMPASS ESL and ACCUPLACER ESL tests used in California community colleges also measure reading and grammar, albeit in very different ways than CELSA. COMPASS ESL and ACCUPLACER ESL provide detailed descriptions of the constructs assessed in each test in online materials. Unlike CELSA, the tests in COMPASS ESL and ACCUPLACER ESL include mostly academic content. COMPASS ESL and ACCUPLACER ESL also differ in important ways. Overall, ACCUPLACER ESL seems to emphasize discrete aspects of language, whereas COMPASS ESL emphasizes language use in context, as exemplified in the use of longer passages in the reading and grammar tests. Below, the Reading and Grammar tests of the two batteries are described in more detail and compared to each other.

## Detailed Comparison: COMPASS ESL vs. ACCUPLACER ESL

**Reading**. COMPASS Reading "assesses a student's ability to recognize and manipulate Standard American English in two major categories: Referring (reading explicitly stated materials) and Reasoning (inferential reading) . . . Most materials are reading passages, ranging

in length from several sentences to many paragraphs" (http://www.act.org/compass/tests/esl.html).

The ACCUPLACER Reading Skills test measures students' understanding of short passages, with "half of the Reading Skills test contain[ing] straightforward comprehension items (paraphrase, locating information, vocabulary on a phrase level, and pronoun reference, and the other half assess[ing] inference skills (main idea, fact versus opinion, cause/effect logic, identifying irrelevant information, author's point of view, and applying the author's logic to another situation)" (The College Board, 2007a, p. 5).

The COMPASS ESL Reading and the ACCUPLACER ESL Reading differ in a number of ways. COMPASS presents sample questions in four distinct levels of proficiency and provides in its website detailed descriptors of each of these levels. ACCUPLACER does not show sample items along different levels of proficiency but does include proficiency statements at three levels of performance. The main difference between the tests relates to the length of the reading passages, and the number of questions that follows each passage. ACCUPLACER's passages are only one paragraph long and are either 50 words or fewer or between 50-90 words, and each passage is followed by a single multiple-choice question. COMPASS, starting on Level 2, presents short paragraphs each followed by two questions and by Level 4 the passages are three or more paragraphs, followed by approximately five multiple-choice questions.

In terms of topical characteristics, the COMPASS ESL Reading test includes passages that vary widely from academic subjects (such as psychology and urban planning) to everyday subjects. Similarly, in the ACCUPLACER ESL Reading Skills test "there is a variety of passage content such as the arts, human relationships, physical science, history/social sciences, and practical situations" (The College Board, 2007a, p.5).

**Grammar.** In COMPASS ESL, grammar is assessed in one Grammar/Usage test, whereas in ACCUPLACER ESL, grammar is assessed in two distinct tests: Sentence Meaning and Language Use. The COMPASS ESL Grammar/Usage test assesses knowledge of Sentence Elements (including verbs, subjects and objects, modifiers, function words, conventions, and word formation) and Sentence Structure and Syntax (including word order, relationships between and among clauses, and agreement, as well as how grammar relates to communication beyond the sentence level). COMPASS items test grammar through a modified cloze format or through multiple-choice items based on a reading passage (http://www.act.org/compass/tests/esl.html). Grammar items are presented in four distinct levels of proficiency and detailed descriptors are provided at each level.

The ACCUPLACER Language Use test assesses similar constructs as the Sentence Elements section of the COMPASS Grammar/Usage test. It measures knowledge of grammar using two different task types: single-word fill in the blank, and sentence combining. The

Sentence Meaning test assesses the understanding of word meanings in one or two sentence contexts and "the sentences are drawn from the subject areas of natural science, history/social studies, arts/humanities, psychology/human relations, and practical situations" (The College Board, 2007a, p.5). This test is similar to the Sentence Structure and Syntax section of the COMPASS ESL Grammar/Usage test. ACCUPLACER ESL gives proficiency statements for three different proficiency levels.

Overall, the grammatical structures tested on the ACCUPLACER ESL and COMPASS ESL are generally the same for both tests; as most selected-response tests of grammar, they best capture what a test taker knows about grammar and not necessarily their ability to use it productively for communicative purposes. However, COMPASS makes more of an effort to assess grammar in context, using longer paragraphs ranging up to 200 words, compared to ACCUPLACER passages, which are typically a maximum of two sentences.

## English Placement Tests

COMPASS, ACCUPLACER, and CTEP are the three commercially available, most commonly used English placement tests in California's community colleges. COMPASS is a system of tests which "includes placement tests that admissions personnel can use to place students in courses appropriate to their skill levels, diagnostics tests that faculty can use to identify specific subject areas where students may need help, and extensive demographics that advisors can use for a thorough understanding of students' support needs" (http://www.act.org/compass/tests/index.html). The COMPASS placement battery is computer adaptive, and includes three tests, Reading, Writing Skills, and the E-write Essay. However, the E-writer Essay test is not used in California community colleges.

The ACCUPLACER battery is designed "to assist with the determination of course placements that are appropriate for students. ACCUPLACER tests can also be used to monitor student course progress and to suggest whether remediation is still needed or if a change in course assignment is recommended" (The College Board, 2007a, p. 1). Like, COMPASS, the ACCUPLACER battery includes three tests: Reading Comprehension, Sentence Skills, and WritePlacer Plus, but the WritePlacer plus is not used in California community colleges. All tests, except for the WritePlacer Plus, are computer adaptive, though paper and pencil options may be available as well.

The CTEP (College Tests for English Placement) is the third most common placement test used in California community colleges. Primarily used to place students in reading and writing classes, CTEP was developed by community college instructors to "measure the skills necessary for success in community college English classes" (http://www.swccd.edu/Pdfs/CTEPSampleFormforStudents.pdf). The CTEP has three subtests:

Reading Comprehension, Sentence Structure and Grammar, and Sentence and Syntax Skills. There is no essay writing task on the CTEP. CTEP is administered as a paper and pencil test.

Table 2 presents the three most commonly used English placement tests.

Table 2
*Overview of the Three Most Commonly Used English Placement Tests*

| CTEP | COMPASS | ACCUPLACER |
|---|---|---|
| **Reading Comprehension**<br>-Seven passages, 35 MC questions | **Reading**<br>-Referring: reading explicitly stated material<br>-Reasoning: inferential reading | **Reading Comprehension**<br>-Reading passage<br>-Sentence relationship task |
| **Sentence Structure and Grammar**<br>-Error correction, cloze, sentence combining<br><br>**Sentence and Syntax Skills**<br>-Cloze task within one text | **Writing Skills**<br>-Usage and mechanics<br>-Rhetorical skills | **Sentence Skills**<br>-Sentence error correction<br>-Construction shift task |
| | **E-write**<br>-Direct writing sample | **WritePlacer Plus**<br>-Direct writing sample |

### Main Findings: English Placement Tests

All three English placement tests assess reading comprehension and grammar broadly defined. Reading comprehension tasks across the three tests require the test taker to respond to questions based on a passage. The nature of the passage and the number of questions that follow each passage vary significantly from one test to another. For example, in COMPASS, reading passages are multiple paragraphs long and are followed by four to eight questions, whereas in ACCUPLACER passages are one paragraph long and each is followed by only one question. ACCUPLACER also has one task type within the Reading Comprehension test called Sentence Relationships that measures cohesion and rhetorical organization.

Grammar is operationalized differently across the three tests, although all tend to focus less on sentence-level grammatical features and more on advanced grammar involving sentence structure, rhetorical skills and sentence logic. Both CTEP Sentence and Syntax and COMPASS Writing Skills include error correction tasks within the context of an entire essay, but COMPASS appears to require a more sophisticated and higher level of knowledge of mechanics.

## Detailed Comparison: English Placement Tests

This section compares how COMPASS, ACCUPLACER, and CTEP measure reading comprehension and grammar.

**Reading Comprehension.** The three reading comprehension tests have similar tasks to evaluate students' reading ability. ACCUPLACER Reading Comprehension consists of 20 multiple choice questions, organized into two types of tasks: 1) a reading passage followed by a comprehension question intended to test main idea, secondary idea, application or inference, and 2) a Sentence Relationship task in which the relationship between sentences, be it supportive, repetitive, or contradictory must be determined in order to assess test-takers' knowledge of cohesion and rhetorical organization (The College Board, 2007b, p 2). In the first task type, each reading passage is followed by only one multiple choice question, and although the Coordinator's Guide states that both "short and long narratives are provided" (The College Board, 2007a, p. 3), all but one of the sample passages are one paragraph long. The topics of the reading passages represent a variety of academic topics drawn from social sciences, natural and physical sciences, human relations and practical affairs, and the arts.

COMPASS Reading passages consist of multiple paragraphs and each is followed by multiple questions (4 to 8 in the sample items provided). The Reading comprehension test evaluates students based on the categories of Reasoning and Referring: "Referring items pose questions about material explicitly stated . . . and reasoning items assess proficiency at making appropriate inferences, developing a critical understanding of the text, and determining the specific meanings of difficult, unfamiliar, or ambiguous words based on the surrounding context" (ACT, 2004, p. 1). COMPASS reading passages are taken from five sources: practical reading, prose/fiction, humanities, social sciences, and natural sciences.

The CTEP Reading Comprehension section consists of seven passages, with 35 multiple-choice questions in total. The one sample item available includes a two-paragraph long passage followed by four multiple-choice questions. The skills measured are main idea, literal comprehension, inferential comprehension, critical/evaluative comprehension, and vocabulary in context. Test takers are given 30 minutes to complete 35 questions. Due to a lack of specific documentation on the test, it is unclear what the topics for the passages are, though a sample item provided relates to a topic in history.

**Grammar/Writing Skills.** ACCUPLACER Sentence Skills, COMPASS Writing Skills, and CTEP Sentence Structure and Grammar, and Sentence and Syntax Skills all address grammar broadly defined. Beginning with ACCUPLACER, the Sentence Skills section is comprised of 20 items of two different types: 1) sentence error correction and 2) rewriting sentences in a construction shift task. The construction shift task requires the test-taker, given a sentence-starter, to select the appropriate choice of words which would complete the sentence without

corrupting the intended meaning. Test materials state that various skills are tested, including the logic of the sentence, sentence completion, and the relationship between coordination and subordination (The College Board, 2007a, p. 4). Like other sections of the ACCUPLACER, topics are taken from a wide variety of academic texts.

COMPASS Writing Skills is described as being designed to determine whether students are ready for classes with heavy writing demands, or whether students "require developmental writing instruction prior to entry into those courses" (http://www.act.org/compass/sample/writing.html). COMPASS Writing Skills tests students on Usage and Mechanics (Punctuation, Grammar, and Sentence Structure) as well as Rhetorical Skills (Strategy, Organization, and Style) by presenting extended passages with embedded errors which students must correct by selecting the corrected form from five-option multiple choice answers. In this section students can opt to correct sections of the text that may indeed already be correct, creating a task that more authentically approximates the actual editing process of a piece of writing.

The CTEP test contains two grammar sections, 1) Sentence Structure and Grammar and 2) Sentence and Syntax. First, the Sentence Structure and Grammar section evaluates test-takers' knowledge of grammar primarily through error correction. The test consists of thirty multiple choice items which must be completed in 20 minutes. Students demonstrate a command of mechanics through sentence correction (similar to the Usage and Mechanics section in COMPASS Writing Skills), cloze tasks, and sentence combining. Second, the Sentence and Syntax section is structured as a series of cloze tasks within one text, creating greater contextual cues and interdependence between one item and the next. This test is made up of forty items, and test takers have 15 minutes in which to complete all questions. Sample items show that in the Sentence and Syntax section, test takers must fill in the blanks with a single correct word and distracters are occasionally very obviously incorrect and easy to eliminate. This section of CTEP is also similar to the COMPASS Writing Skills section in that the text corrections take place within the context of an entire essay rather than in shorter pieces of text; however, the tasks required in the COMPASS are much more sophisticated and demonstrate a higher level of knowledge of mechanics than those measured on the CTEP.

## ESL vs. English Placement Tests

As mentioned earlier, when language minority students decide to go to a community college, they typically fill out an application, they take the placement test, and then they meet with a counselor. In many colleges, the decision of what tests they should take, ESL or English, is made by the students themselves who self identify as an ESL student or a mainstream student (see Bunch et al., 2011 for more details). Thus language minority students, depending on the college and their decision-making, could end up taking either the ESL or the English placement tests. This section thus directly compares these tests to each other.

## Main Findings: ESL vs. English Placement Tests

Broadly, the ESL and the English placement tests reflect different constructions of language proficiency. Although both the ESL and English batteries assess reading, grammar, and writing, only the ESL batteries assess listening. The ESL tests define language ability more discretely whereas the English tests measure more contextual uses of language. For the skills assessed in both batteries, the main difference lies in the range of language ability levels targeted by the items.

The English placement tests measure language proficiency at higher levels. In general, only the highest proficiency levels in the ESL tests approximate the constructs assessed in the English test. The English tests are also more academic in nature and formal in register whereas the ESL tests include academic and non-academic content and both formal and informal registers. These differences might be because the ESL placement tests are often used to place students along a continuum of both academic and non-academic ESL courses (e.g. courses designed for students seeking English instruction for basic workplace and community-survival purposes). Regular English tests place students into remedial or college-level English courses which, although they are based on very different orientations toward language, literacy, and academic support from college to college (Bunch et al., 2011), are ostensibly designed to prepare students for academic work in English.

## Detailed Comparison: ESL vs. English Placement Tests

**COMPASS ESL vs. COMPASS English**. The COMPASS ESL consists of a battery of 4 tests: Reading, Listening, Grammar/Usage, and E-write, whereas the COMPASS English consists of three tests: Reading, Writing Skills, and E-write. The English battery does not include a test of Listening or Grammar/Usage, although the Writing Skills test is to a large extent a test of advanced grammar. Table 3 presents the tests in the English and ESL batteries of the COMPASS.

Table 3
*COMPASS ESL and COMPASS Test Batteries*

| COMPASS ESL | COMPASS |
|---|---|
| **Reading**<br>-reading explicitly stated material<br>-inferential reading | **Reading**<br>-referring: reading explicitly stated material<br>-reasoning: inferential reading |
| **Grammar/Usage**<br>-sentence elements<br>-sentence structure and syntax | **Writing Skills**<br>-usage and mechanics<br>-rhetorical skills |
| **E-Write**<br>-direct writing sample | **E-write**<br>direct writing sample |
| **Listening** | |

*Reading*. Both the ESL and English Reading tests assess the same reading constructs: referring and reasoning. Referring items pose questions about material explicitly stated in a passage whereas reasoning items assess proficiency at making appropriate inferences. The ESL test, however, evaluates these constructs along a wider range of ability than the English test, from low proficiency to advanced with the highest level (Level 4) of the ESL test being comparable to the difficulty level of the English Reading Comprehension test.

*Grammar*. There is some overlap between the COMPASS ESL Grammar/Usage test and the COMPASS Writing Skills test. As discussed above, the Writing Skills test does not assess basic knowledge of grammatical skills; it taps only into the more advanced levels of usage and mechanics. As would be expected, the ESL test measures a broader range of ability in grammar, from very low to advanced. The Writing Skills test also assesses rhetorical skills such as writing strategy, organization, and style, that are not assessed in the ESL Grammar/Usage test.

**ACCUPLACER ESL vs. ACCUPLACER English.** Table 4 below presents the tests in the ESL and English batteries of ACCUPLACER. ACCUPLACER ESL includes five tests whereas the English version consists of only three tests. Both include a reading test, a writing test, and at least one test of grammar, but there are differences in the ways in which these skills are evaluated in the ESL and the English versions of the test. The ACCUPLACER English battery does not include a test of listening.

Table 4
*ACCUPLACER ESL and ACCUPLACER Test Batteries*

| ACCUPLACER ESL | ACCUPLACER |
|---|---|
| **Reading Skills**<br>-comprehension of short passages | **Reading Comprehension**<br>-reading passage<br>-sentence relationship task |
| **Sentence Meaning**<br>-understanding of word meanings in one-or-two sentence contexts | **Sentence Skills**<br>-sentence error correction<br>-construction shift task |
| **Language Use**<br>-grammar and usage | |
| **WritePlacer**<br>-direct measure of writing skills | **WritePlacer Plus**<br>-direct measure of writing skills |
| **Listening** | |

*Reading*. Comparing the ESL Reading Skills test and the English Reading Comprehension test, we find that though they are based on the same general model, there are some interesting differences. The ACCUPLACER ESL Reading Skills test measures students' understanding of short passages. Items include "straightforward comprehension" and " inference skills (main idea, fact versus opinion, cause/effect logic, identifying irrelevant information, author's point of view, and applying the author's logic to another situation)" (The College Board, 2007a, p. 5). The English test is divided into two parts; the first part, like the ESL Reading Skills test evaluates test-takers' ability to answer questions based on a reading passage. The questions however primarily target higher level skills (main idea, secondary idea, application and inference). The second part of the Reading Comprehension test, Sentence Relationships, which assesses student's ability to identify relationships between sentences (i.e. cohesion and rhetorical organization), is not included in the ESL version.

*Grammar*. Both the ESL and the English batteries assess grammatical knowledge, though the focus is different for the two different populations of test-takers. As mentioned above, the ESL battery includes two tests, Sentence Meaning and Language Use, which place a greater emphasis on grammatical form and mechanics, whereas the English battery includes only one Sentence Skills test. The English Sentence Skills test does not focus on discrete grammatical features but rather on the broader knowledge of sentence structure including logic of the sentence, complete sentences, and the relationship between coordination and subordination. This test also includes two task types: sentence correction and construction shift. The construction shift items are similar to the sentence combining items in the ESL Language Use test.

## IMPLICATIONS FOR THE TESTING AND PLACEMENT OF US-EDUCATED LANGUAGE MINORITY STUDENTS

This report has provided a detailed review and side-by-side comparison of the constructs assessed by the most widely used commercially available placement tests in CA community colleges. Two of the batteries reviewed, ACCUPLACER and COMPASS, are also the two most commonly used for placement in colleges nationwide (Hughes & Scott-Clayton, 2010). The purpose for this content analysis was to better understand the tests that are used for placement in particular as they affect language minority students. Future research should examine what these tests actually measure by examining test taker performance on these tests and conducting construct validation investigations. A better understanding of the constructs assessed by the tests these students face can help community colleges and instructors make decisions about how to better serve the unique needs of this growing population.

The comparison among the three most prominent community college ESL placement tests revealed differences in the ways in which each test has operationally defined the construct of language ability. The CELSA, the most widely used ESL test in California's community colleges, is the one that stands out as most limited in the ways identified by Bunch and Panayotova (2008). Originally designed for an adult school population, the CELSA assesses test takers' reading and grammar knowledge using a cloze format on passages that are not academic in nature. Although COMPASS ESL and ACCUPLACER ESL have the capability to assess listening, reading, grammar and essay writing, only the reading and grammar tests are used for placement. Unlike CELSA, these tests use a variety of tasks that include both non-academic as well as academic content. COMPASS ESL and ACCUPLACER ESL also differ from each other in the ways in which they operationalize these language constructs. Tasks in COMPASS ESL tend to provide more context than tasks in ACCUPLACER ESL as evident in the longer passages in the reading and grammar tests.

The three English placement tests assess reading comprehension and grammar. The COMPASS and ACCUPLACER batteries include a direct assessment of writing, but these essay writing tests are not used in California community colleges. CTEP, COMPASS and ACCUPLACER all assess reading comprehension by targeting both basic and higher order reading skills but the tasks differ significantly across tests. For example, COMPASS includes longer passages followed by multiple questions, whereas ACCUPLACER includes short passages each followed by only one question. Grammar is also operationalized differently across the three tests. All tests tend to focus more on advanced grammar involving sentence structure, rhetorical skills and sentence logic than on discrete grammatical features. Although both CTEP and

COMPASS include error correction tasks within the context of an entire essay, COMPASS appears to require a more sophisticated and higher level of knowledge of mechanics.

When compared to each other, the ESL and the English placement tests used in CA community colleges reflect a similar construction of language proficiency in that they all assess reading and grammar. The ESL tests focus more on the building blocks of proficiency whereas the English tests measure higher levels of proficiency and more contextual uses of language. The English tests are more academic in nature and formal in register whereas the ESL tests include academic and non-academic content and both formal and informal registers.

## Appropriateness of Placement Tests for the US-LM Population

As discussed in the Introduction, language proficiency is a complex construct. Both the English and the ESL placement tests used in California community colleges measure a very narrow portion of students' linguistic abilities, namely their ability to read and understand relatively short passages and their knowledge of grammar. As a result these placement tests provide a very limited picture of what a student can do with language. These tests do not assess students' speaking or listening abilities, or their ability to write for a number of academic and non-academic purposes. This is particularly problematic for US-LM or Generation 1.5 students whose language profiles are different than those of ESL students and monolingual students (Valdes and Figueroa, 1994; DiGennaro, 2009). As stated earlier, US-LM students, due to their experiences with English in naturalistic settings and limited formal instruction of the language, typically have stronger listening and speaking skills but weaker metalinguistic knowledge of grammar than the international students typically served by ESL programs. By not assessing listening and speaking, and focusing instead on assessing knowledge of grammar, the placement testing approach in California community colleges may be producing an incomplete and distorted picture of what Generation 1.5 students can actually do with language.

As this report has shown, however, there are important differences between the ESL and the English placement tests and also within the ESL and English tests in terms of how language ability is operationally defined and assessed. Although they all only measure reading and grammar, thus providing a very limited picture of students' language ability, COMPASS appears to measure these two skills in more contextualized ways than the other tests. This is important given that US-LM students' "skills are stronger with context-embedded language" (Di Gennaro, 2008, p. 71).

A recent study that examined the actual linguistic differences between Generation 1.5 and international students in the context of writing provides evidence that Generation 1.5 students have different linguistic profiles and perform differently on writing assessments (Di Gennaro, 2009). Di Gennaro (2009) found that Generation 1.5 students' writing ability was not uniform across various dimension of writing. Specifically she found that for international

students the various dimensions of writing "may develop concurrently, while for Generation1.5 students, these components are imbalanced, suggesting the Generation 1.5 students may be strong in one area at the same time that they are lacking in another" (p. 551). Di Gennaro warns that "if writing evaluators attend to only one or two of these writing components, they may not have a complete picture of G1.5 students' writing strengths and weaknesses" (p. 543). If US-LM students' language ability is imbalanced across and within skills, then making inferences about their language ability on the basis of their performance on reading and grammar alone may indeed result in an inaccurate assessment of their language ability (see also Valdés & Figueroa, 1994). Further research is needed to examine differences between Generation 1.5 students and international students in areas other than writing, such as their performance on standardized ESL placement tests like COMPASS and ACCUPLACER. Also, given that many US-LM students self identify as non-ESL and take the English placement tests, research is needed that examines differences and similarities between Generation 1.5 students and monolingual English-speaking students at community colleges in terms of their characteristics and their performance on placement tests.

Interestingly, Di Gennaro (2009) also found that the writing performance of Generation 1.5 students was more problematic for raters to judge and she speculates that this may be because "the scales used to score the essays were devised primarily to evaluate the type of writing produced by the [international student] writers, and thus were not sensitive enough to evaluate Generation 1.5 essays more accurately" (p. 550). Di Gennaro concludes that "assessments designed for the [international student] population may not function as expected for Generation 1.5 students" (p. 553). It is unclear who constitutes the intended population of the commercially-available placement tests used in California community colleges. The English placement tests do not specify an intended population, other than "students." The ESL placement tests on the other hand specify "ESL students," and "non-native speakers of English." Given the evidence that the language characteristics of various groups of non-native speakers of English, such as international students and Generation 1.5 students, may vary, it is important that testing companies provide evidence that their tests are appropriate for all the groups of test takers targeted by their tests. Unfortunately, for this study we were not able to obtain technical reports from COMPASS or ACCUPLACER to determine whether US-LM students were included in the norming population. Given the fact that US-LM students are rarely identified as such, it is unlikely that information on this population during the norming process, even if such students were included, is available. In community college programs, courses, and placement tests, the ESL/English dichotomy "disguises" the US-LM student population and allows for their needs to be ignored. Part of the problem is that there is very little research that documents the unique language characteristics of US-LM students in comparison to other groups such as international students and native speakers. This is further complicated by the fact that the US-LM student population itself is also very diverse (Di Gennaro, 2008).

The concerns about the appropriateness of using commercially available tests for placing US-LM students in community colleges are also relevant to the use of writing samples for placement. Writing samples are used by several colleges both for ESL and English placement, sometimes in combination with a commercially developed test. As demonstrated by Di Gennaro's research, the use of writing samples also require a careful examination of how the writing construct is defined (as reflected in the rubric) and how the sample is scored to ensure that the diverse and unique language abilities of US-LM students are accurately and fairly assessed.

## Placement and Alignment Issues

The close examination of the content of commonly used community college placement tests revealed that the construct of language ability is operationalized in different ways. It is important to understand these differences because, as Bachman & Palmer (1996) explain, a test's construct definition and the characteristics of the test tasks determine the kinds of interpretations that can be made on the basis of scores and the domains to which these interpretations can be generalized. In addition to considerations regarding the test-taker population, for effective placement on the basis of scores derived from these tests, there should be a certain level of correspondence between what is assessed by the test and what is taught in the course sequences. Thus, it is important for colleges to consider the following: How well are the constructs assessed by the placement test used at a particular college aligned to the curriculum to which students are placed on the basis of scores? In other words, does the type of language assessed on the placement test reflect the types of language addressed in the course sequence, ESL or English? The issue of alignment is critical in determinations of the construct validity and authenticity of the interpretations about test takers' language ability obtained from these tests.

Both the COMPASS and ACCUPLACER batteries provide resources and guidelines to colleges that they can use for standard-setting and validation. And California community colleges are required to conduct a validation study of the placement tests they adopt. Part of this study involves reviewing the content of test items and mapping the alignment between those items and the course sequence. The assessment and validation process, governed by state and systemwide policy (California Community College Assessment Association [CCCAA], 2005; California Community Colleges, 2001), often includes some faculty involvement, but this involvement varies from college to college, and after a test is adopted, ESL and English faculty may have little familiarity with the placement tests and the types of information they provide about students' language ability (Bunch et al., 2011). The primary purpose of a placement test is to efficiently group students. However, the tighter the alignment between the test and the program, the more likely that the test could not only be used for efficient grouping but also "provide information on which to base diagnostic information or sequencing of instruction materials." (Green and Weir, 2004).

In addition to the alignment between the test and the courses, it is also critical to consider the extent to which the constructs addressed in the ESL and English course sequence into which students are placed are aligned to the types of language use that students need in order to succeed in content courses and college in general. If a US-LM student who wants to pursue a higher education is placed in ESL, will he or she receive the type of English language instruction that will help her succeed not only in the ESL class but also in future English and content courses? For example, even if the current placement tests are adequately aligned to the courses into which students are placed, US-LM students may not be receiving the education they need. It is possible that neither the test nor the program into which they are placed, ESL or English, account for these students' abilities and therefore impede their progress instead of facilitating their academic trajectories. Ultimately, the goal of matriculation, placement, and ESL and English course sequence should be to maximize students' chances of success in pursuing academic pathways. Future research should carefully examine the language demands of community college content courses and compare those demands to the language instruction students typically receive in ESL and English courses.

## Placement as a High-Stakes Decision

Placement tests are typically viewed as low or medium stakes because decisions made can be easily reversed (e.g. moving a student to an easier or more challenging class based on an in-class diagnostic assessment), and because in many contexts the results do not impact students' access toward credit-bearing coursework (for example, university foreign language placement exams place students in varying levels of coursework, but all courses carry credit that can be used toward an undergraduate degree). In the case of California community college placement testing, however, the stakes are high, both because placement determine how far away from college-level English coursework students begin, and because in many colleges it is very difficult to change placements once a semester begins (Bunch et al. 2011). Improper placement can result in students having to take additional semesters of ESL or English which could delay their academic goals, reduce their motivation, and cause them to drop out. Furthermore existing research on the predictive validity of the placement English tests suggest that the evidence "is not as strong as desirable": The COMPASS scores predict performance on the target course with a B or higher with 60 to 70% accuracy, and for ACCUPLACER the placement accuracy rates range from 59 to 66% (Hughes & Scott-Clayton, 2010, p. 17). Thus, because placement is a high stakes decision and placement tests can only provide limited information, the decision should not be made on the basis of only one test score (American Education Research Association, 2000; American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). This is true for all students but especially true for US-LM students for whom the score might be even less representative of what they can actually do using language. Even though officially state and systemwide policy mandate the use of placement should be done on the basis of multiple

measures for placement in California community colleges (CCCAA, 2005), as Bunch et al. (2011) reveal, this is not often the case. Local implementation of multiple measures varies widely and is sometimes minimal to nonexistent for many students (Bunch et al., 2011).

Information from K-12 schools could serve as important measures to inform student placement. First, in an initiative relevant to all students, several colleges are currently piloting an Early Assessment Program (EAP), based on a system used for several years in the California State University system, in which successful scores on an enhanced 11th grade English Standards Test can be used to exempt students from the placement process and to enroll directly into college-level English courses. Second, for US-LM students, given that reclassification is done on the basis of the CELDT and adequate performance on content tests among other things, it seems reasonable that those students who are reclassified before grade 12 should be able to take the regular English placement tests. Research that explores the relationship between language minority students' K-12 English language designations and potential placements in community colleges could prove to be particularly useful. If it is the case that students reclassified as Fluent English Proficient (RFEP) in K-12 schools demonstrate high levels of success in college-level English courses, then exempting such students from English or ESL placement tests would save colleges precious financial resources and time, and it would allow students to enroll in credit-bearing courses sooner so that they can pursue their academic and professional goals more efficiently.

This report has aimed to provide a framework for considering what aspects of language proficiency and academic skills different tests are designed to measure, provide specific information on the most commonly used tests in California community colleges, and raise issues specific to US-LM students. The issue is not simply whether an ESL or an English test is most appropriate for this population, but also whether an ESL or an English course sequence would better serve students' needs. It may also be possible that neither of those two options are adequate, and that other possibilities should be explored. In their instructional and testing practices, colleges might need to think more about students' academic goals and trajectories rather than simply placing students into the existing "pre-collegiate" ESL and English "slots" (see Bunch et al., 2011). Ultimately, it is our hope that discussions about placement tests can foster thoughtful conversations regarding the larger challenges involved in addressing the needs of US-LM students in California community colleges as they seek promising academic and professional futures.

## REFERENCES


ACT (2004) COMPASS/ESL Reading: Sample Test Questions—A guide for students and parents. Retrieved from http://www.act.org/compass/sample/pdf/reading.pdf

ACT (2007). *Compass Guide to Successful ESL Course Placement*. Retrieved from http://www.act.org/compass/pdf/ESLGuide.pdf

ACT (2008). Answers to Frequently Asked Questions about COMPASS e-Write and ESL e-Write. Retrieved from http://www.act.org/compass/pdf/ewritefaq.pdf

Association of Classroom Teacher Testers (2000). CELSA Combined English Language Skills Assessment: Test Administrator's Manual and Technical Guide for Ability to Benefit. Montecito, CA: Association of Classroom Teacher Testers.

American Education Research Association (2000). AERA position statement concerning high stakes testing in preK-12 education. Retrieved from http://www.aera.net/?id=378

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington D.C.: American Psychological Association.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bailey, T. (2009). Rethinking developmental education in community college. Research Brief #40. New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc.tc.columbia.edu

Bailey, T., Jeong, W. D., & Cho, S.-W. (2010). Student progression through developmental sequences in community colleges (Research Brief #45). New York, NY: Community College Research Center, Teachers College, Columbia University. Retrieved from http://ccrc.tc.columbia.edu

Bunch, G. C. (2008). Language minority students and California community colleges: Current issues and future directions. *Community College Policy Research* [Electronic Journal], California Community College Collaborative (C4), University of California, Riverside. Retrieved from http://escholarship.org/uc/item/6x39n5kd

Bunch, G. C. (2009). Immigrant students, English language proficiency, and transitions from high school to community college. In T. G. Wiley, J. S. Lee & R. Rumberger (Eds.), *The education of language minority immigrants in the United States* (pp. 263-294). Clevedon, England: Multilingual Matters.

Bunch, G. C., Endris, K., Panayotova, D., Romero, M., and Llosa, L. (2011). *Mapping the terrain: Language testing and placement for US-Educated language minority students in California's community colleges*. Report prepared for the William and Flora Hewlett Foundation. Available at http://escholarship.org/

Bunch, G. C., & Panayotova, D. (2008). Latinos, language minority students, and the

construction of ESL: Language testing and placement from high school to community college. *Journal of Hispanic Higher Education, 7*(1), 6-30.

California Community College Assessment Association [CCCAA]. (2005). *Assessment Q&A: Questions and answers on assessment for use in the California Community Colleges*. Author. Retrieved from http://www.cccco.edu/LinkClick.aspx?fileticket=qLFSNlpn5Ls=&tabid=628&mid=1718

California Community Colleges (2001). Standards, policies, and procedures for the evaluation of assessment instruments used in the California community colleges. 4th Ed. (Revised March, 2001). Retrieved from http://www.cccco.edu/LinkClick.aspx?fileticket=ubGoW34y0Mo=&tabid=628&mid=1719

California Department of Education. (2009a). *California English Language Development Test (CELDT): Understanding and Using 2009–10 Individual Results*. Retrieved from http://www.cde.ca.gov/ta/tg/el/documents/celdt09astpkt1.pdf

California Department of Education. (2009b). Standardized Testing and Reporting (STAR) Program: Summary of 2009 Results. Retrieved from http://www.cde.ca.gov/ta/tg/sr/

California Department of Education. (2008). CELDT Released Test Questions. Retrieved from http://www.cde.ca.gov/ta/tg/el/documents/celdtrtqs0408.pdf

Consultation Council Task Force on Assessment (2008). Report of the Consultation Council Task Force on Assessment to the Board of Governors of the California Community Colleges. Retrieved from http://www.cccco.edu/Portals/4/Executive/Board/2008_agendas/january/3-5_Assessment%20TF%20Report%2001-08.pdf

di Gennaro, K. (2009). Investigating differences in the writing performance of international and Generation 1.5 students. *Language Testing, 26*(4) 533–559

di Gennaro, K. (2008). Assessment of Generation 1.5 learners for placement into college writing courses. *Journal of Basic Writing, 27*(1), p. 61-79.

Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). *Listening to teachers of English language learners: A survey of California teachers' challenges, experiences, and professional development needs*. Santa Cruz, CA: The Center for the Future of Teaching and Learning.

Green, A. B. & Weir, C. J. (2004). Can placement tests inform instructional decisions? *Language Testing, 21*(4), 467.494.

Harklau, L., Losey, K. M., & Siegal, M. (Eds.). (1999). *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hughes, K. L. & Scott-Clayton, J. (2010). Assessing developmental assessment in community colleges: A review of the literature. Community College Research Center Working Paper #19. New York: Teachers College, Columbia University

Roberge, M., Siegal, M., & Harklau, L. (Eds.). (2009). *Generation 1.5 in college composition: Teaching academic writing to U.S.-educated learners of ESL*. Mahwah, NJ: Lawrence Erlbaum Associates.

Roberge, M. M. (2002). California's generation 1.5 immigrants: What experiences, characteristics, and needs do they bring to our English classes? *CATESOL Journal, 14*(1), 107-129.

The College Board (2009). ACCUPLACER i3. Retrieved from http://professionals.collegeboard.com/profdownload/accuplacer-whats-new.pdf

The College Board (2008). WritePlacer Guide with Sample Essays. Retrieved from http://professionals.collegeboard.com/profdownload/accuplacer-writeplacer-sample-essays.pdf

The College Board (2007a). ACCUPLACER Coordinator's Guide. Retrieved from http://www.matyconn.org/MathIssues_StatewideInitiatives/Accuplacer-coordinatorsGuide.pdf

The College Board (2007b). ACCUPLACER Sample Questions for Students. Retrieved from http://professionals.collegeboard.com/profdownload/accuplacer-sample-questions-for-students.pdf

Valdés, G. & Figueroa, R. A. (1994). *Bilingualism and testing: A special case of bias*. Westport, CT: Ablex Publishing.

Venezia, A., Bracco, K. R., & Nodine, T. (2010). One-shot deal?: Students' perceptions of assessment and course placement in California's community colleges. San Francisco, CA: WestEd.

Woodlief, B., Thomas, C., & Orozco, G. (2003). *California's gold: Claiming the promise of diversity in our community colleges*. Oakland, CA: California Tomorrow.

## APPENDIX A: Community College ESL Placement Tests vs. CELDT

This appendix compares the ESL community college placement tests and the ESL test used at the K-12 level: the California English Language Development Test or CELDT. The CELDT is a standardized test designed to assess the English proficiency of English learners in California. All students whose home language is not English must take the test within 30 calendar days after they are enrolled in a California public school for the first time to determine if they are English learners. The CELDT must also be given once a year in the fall to English learners until they meet the criteria for reclassification as Fluent English Proficient.

Following the organization of the California ELD Standards, the CELDT covers four skill areas: Listening, Speaking, Reading, and Writing and four different grade categories: kindergarten-grade 2, grades 3-5, grades 6-8, and grades 9-12. Results are reported according to five proficiency levels—beginning, early intermediate, intermediate, early advanced, and advanced—for each skill area and overall.

A comparison of the characteristics and constructs assessed by the CELDT Grades 9-12 and the community college ESL placement tests sheds important insights into the level of alignment in terms of language expectations across these two educational contexts. The next section provides a detailed comparison of the three commercially available, most commonly used community college ESL placement tests to the CELDT Grades 9-12 in terms of their characteristics and language constructs assessed. We begin by comparing each CELDT Subtest to the corresponding COMPASS ESL and ACCUPLACER ESL test. Next we compare CELDT as a whole to CELSA. CELDT is described using information from the "Released Test Questions" document prepared by the California Department of Education released in April 2008 (California Department of Education, 2008). Table 5 provides a visual representation of the subtests/tests in the CELDT and the ESL community college placement tests.

Table 5
*Comparison of CELDT and the ESL Community College Placement Tests*

| CELDT (K-12) | COMPASS ESL | ACCUPLACER ESL | CELSA |
|---|---|---|---|
| **Reading**<br>-word analysis<br>-fluency& vocabulary<br>-reading comprehension | **Reading**<br>- reading explicitly stated material<br>- inferential reading | **Reading Skills**<br>-comprehension of short passages | **Reading/Grammar**<br>-understanding of meaning in a context, as well as grammatical ability. Measures language ability in a holistic manner. |
| **Listening**<br>-following oral directions<br>-listening comprehension<br>-extended LC | **Listening**<br>-the ability to understand standard American English | **Listening**<br>-the ability to listen to and to understand one or more persons speaking in English | |
| **Writing**<br>-grammar& structure<br>-sentences<br>-short Composition | **Grammar**<br>-sentence elements<br>-sentence structure and syntax | **Sentence Meaning**<br>-understanding of word meanings in one-or-two sentence contexts | |
| | | **Language Use**<br>-grammar and usage | |
| | **E-write**<br>-direct measure of writing skills | **WritePlacer ESL**<br>-direct measure of writing skills | |
| **Speaking**<br>-oral vocabulary<br>-speech functions<br>-choose/give reasons<br>-4-picture narrative | | | |

## Main Findings: Community College ESL Placement Tests and CELDT

The most noticeable difference between CELDT and the community college placement tests is that the CELDT assesses all four skills with grammar assessed within each of the skills subtests. The community college tests do not test Speaking, Listening or Writing and have separate grammar tests. Mode of delivery and scoring are also a major difference. Some sections of CELDT are teacher administered and scored whereas the community college tests (except for CELSA) are computer administered and scored. The CELDT, overall, is significantly less challenging than the community college placement tests other than the CELSA. For example, the Writing section of the CELDT requires students to write an explanation using a paragraph, whereas two of the community college writing tests require students to write a fully developed,

argumentative essay. The topics on the CELDT are typically school-based but the tasks are generally less academic in what they require students to do. Most of the released CELDT items were comparable to the lower level items in the community college tests.

These differences between the K-12 and the community college placement tests reflect the different purposes of the tests themselves. The CELDT is a language proficiency test but also an achievement test that demonstrates the extent to which students have mastered the ELD standards they developed through instruction in K-12 (except for when the CELDT is used for initial identification and placement). Also, federal law requires that all four skills be assessed. The community college placement tests are designed primarily for initial placement purposes and as a result they prioritize practicality (e.g. multiple-choice items that are computer administered and scored).

### Detailed Comparison: Community college ESL placement tests and CELDT

**Reading.** The CELDT Reading Section includes three task types: Word Analysis, Fluency and Vocabulary, and Reading Comprehension. In Word Analysis, students are asked to recognize and select correct English forms. For example, in the item presented in the Sample, students are expected to select an appropriate plural form with regard to correct morphemic and spelling use. Items in the Fluency and Vocabulary section require students to use context to determine meanings of unknown vocabulary. In the sample provided, students are required to select an appropriate synonym as the underlined word from the multiple-choice options. In the Reading Comprehension section, test-takers read passages of varying lengths and answer questions based on their general comprehension of the text. Students may be required to answer more than one question on each passage.

Comparing Reading on the CELDT and the community college placement tests proves a difficult task since the composition of reading sections differs significantly from test to test. For practical purposes, we have focused on the Reading Comprehension section exclusively, since the Word Analysis and Fluency and Vocabulary sections of the CELDT overlap more with the grammar sections on the college tests. The released test item for the Comprehension section of the CELDT Grades 9-12, consists of a job announcement followed by 3 multiple-choice questions and resembles the types of items on the lower levels of COMPASS ESL Reading (levels 1 and 2) Both CELDT and the COMPASS ESL lower level reading sections utilize visual aids to assist the test taker with comprehension. There is no pictorial support on the ACCUPLACER Reading comprehension section. Overall, CELDT measures reading at a lower level than the community college placement tests.

**Grammar.** Grammar is tested in the COMPASS ESL Grammar/Usage test and ACCUPLACER's Language Use and Sentence Meaning tests. In the CELDT, grammar is assessed primarily in the Grammar and Structure section of the Writing subtest. In the CELDT, the

Grammar and Structure tasks within the Writing section seem typical of multiple choice grammar tasks. For example, the sample item provides a modified cloze prompt, in which test takers must select the appropriate comparative form from three multiple choice options. One format of the ACCUPLACER Language Use section tests very much the same concepts as the CELDT, but goes beyond the CELDT specifications to also assess sentence combining. The COMPASS ESL Grammar/Usage section also tests a similar construct, but branches further into more sophisticated grammar or sentence structure and clause relationships. It should be noted that some of these concepts are evaluated on the CELDT test, but in the Writing tasks (Sentences and Short Compositions), rather than through multiple choice selection. The Sentence test on the CELDT is unique in its nature compared to the community college placement tests. In this task, students respond to a picture by producing a sentence to describe it. Scoring is based largely on word order and correctness of syntax. Interestingly, the rubrics to evaluate the Sentences section and the Short Compositions section remain constant from 2nd through 12th grade. Overall, the community college tests assess a broader range of grammar ability at higher levels than the CELDT.

**Listening.** The CELDT Listening subtest includes 20 questions divided into three different task types of increasing difficulty: Following Oral Directions, Listening Comprehension (Teacher Talk), and Extended Listening Comprehension. In Following Oral Directions students are typically given a one-sentence oral directive prompt which they need to complete. In Listening Comprehension students are given a three-sentence oral prompt describing a school situation and must answer basic listening comprehension questions. In Extended Listening Comprehension, students listen to a longer oral passage and answer a number of comprehension questions based on passage content.

The primary differences between the listening sections on the CELDT and the listening tests available in the college batteries lie in the mode of delivery and the nature of the listening passages. The CELDT listening test is read aloud by a test administrator, most likely a teacher, whereas the college tests are delivered in a pre-recorded format. As for the listening passages, at the higher levels the community college placement tests include academic lectures whereas the CELDT the passages are less academic in that they focus on school situations broadly. ACCUPLACER and CELDT test takers are given opportunities to hear sections or directions more than once as well as strategies to support listening comprehension. In ACCUPLACER, re-listening to sections of the test is permitted twice after the initial reading, and likewise, the CELDT prompts may be read twice to students, though questions can be read only once. COMPASS does not specify if re-listening is admissible, but does encourage test takers to take notes in an effort to test listening skills rather than short term memory skills. In all tests, pictures are used as listening support in the lower levels, with clusters of questions relating to the same listening comprehension passages as the items increase in difficulty.

**Writing.** The CELDT Writing subtest includes three components: Grammar and Structure, Sentences, and Short Composition. The Grammar and Structure section is presented in a multiple choice format, whereas the Sentences and Short Composition sections are presented in constructed response formats. Grammar and Structure requires the students to make MC selections based on correct grammar and spelling, appropriate lexical choices, and transitions. The Sentences section requires students to write a single sentence to describe a picture, thereby demonstrating their writing ability at the sentence level. Topics for this task are not necessarily academic in nature. The sentences are scored on a rubric according to the following 6 categories: Content, Response, Subject and predicate, Grammar and syntax, Vocabulary, Spelling and mechanics (California Department of Education, 2008, p. 73), and scores are reported between 0-3 according to levels of proficiency (0-No Communication, 1-Emerging Communication, 2-Basic Communication, and 3-Fully Competent Communication). The Short Composition task is composed of a simple prompt, to which students must craft a response to a topic or situation. For example, the sample prompt requires students to write a paragraph about something they would like to learn to do and why. The compositions are scored on a rubric according to the same 6 categories stated above, and the scores are reported between 0-4 according to levels of proficiency (0-No Communication, 1-Emerging Communication, 2-Developing Communication, 3-Competent Communication, and 4-Expressive Communication). Sentences and Short Compositions are typically scored on site by teachers.

The constructs assessed in the Writing section of the CELDT overlap with both the writing and grammar tests available in the community college ESL placement batteries. The Short Composition task on CELDT most closely resembles COMPASS ESL e-Write and the ACCUPLACER WritePlacer-ESL, in that they all require the student to produce an original piece of writing in response to a prompt. However, the CELDT writing test differs from the two community college writing tests in terms of 1) mode of administration and assessment (human administered and assessed vs. computer administered and assessed), 2) genre of writing elicited by the prompts (explanation vs. argumentative), 3) length of expected response (a paragraph of 3 or more sentences vs. a fully developed essay), and 4) the operational definition of writing ability (more local vs. more global). Both the e-Write and the WritePlacer-ESL are computer administered and computer scored by a system called IntelliMetric, by Vantage Technologies, which is calibrated to simulate a human rating process. In the case of the CELDT, the students handwrite the paragraph, and their submissions are teacher-scored, with no information provided as to how this is carried out. Important differences between the CELDT and the e-Write and WritePlacer ESL relate to the genre and the expected level and length of the writing sample. The community college tests elicit argumentative writing whereas the CELDT elicits an explanation. As observed in student samples and in rubric specifications, essays written on the CELDT are much less sophisticated than those on the college tests given the nature of the prompt. Though CELDT documentation does not have explicit length requirements; in the assessed domain of Subject-predicate, a length of at least three sentences is necessary in order to receive the highest rating of 4-Expressive Communication. Likewise, the community college

writing tasks do not specify length requirements; yet, the complexity of the tasks requires a student to write more in order to fulfill the task. As suggested by the sample WritePlacer prompts, students are asked to compare and contrast options, and make arguments for their rationales based on examples and support. This task certainly could not be completed well in three sentences, and thus, the sophistication of writing expected is greater in the college tests than the CELDT. The operational definitions of the construct of writing ability in these tests, as reflected in their respective rubrics, are also differently. CELDT defines the construct of writing ability more narrowly (e.g. Subject-Predicate) than the two community college writing tests.

**Speaking.** The community college placement batteries do not include a test of Speaking. The CELDT however has a Speaking section that is administered one-on-one and includes four components: Oral Vocabulary, Speech Functions, Choose and Give Reasons, and 4-Picture Narrative. In Oral Vocabulary, the student must identify an object, and then speak about it, identifying its name, use, or other distinguishing characteristics. Students are also expected to identify opposites of words in prompts and say them aloud. For Speech Functions, students are provided a context and then must make a request or a statement to express a need. In Choose and Give Reasons, test takers are given a choice between two things, activities or options and must make a statement about them, providing rationales for their choices. In the 4-Picture Narrative section, students look at four sequenced pictures and tell the story that they depict. In all CELDT Speaking sections, prompts and questions can be read only once, and though it is untimed, the test takes approximately 10 minutes for students to complete.

### CELDT vs. CELSA

Both CELDT and CELSA require test-takers to make grammatical judgments to fill in the blanks correctly rendering CELSA most similar to the CELDT Grammar and Structure section within its Writing subtest. Unlike CELDT's four skills testing, CELSA does not test speaking, listening, or writing.

# APPENDIX B: Review of Commercially-Available Tests Not Currently Used for Placement in California's Community Colleges

## Description of COMPASS ESL and ACCUPLACER ESL Listening and Essay Writing Tests

**Listening.** It is difficult to compare the COMPASS ESL Listening and ACCUPLACER ESL Listening tests because ACCUPLACER sample items are not publicly available. As with Reading, COMPASS Listening includes questions along four levels of proficiency that range from basic comprehension of words or phrases to the ability to listen to academic lectures. COMPASS Listening items increase in difficulty depending on the proficiency levels; pictorial support is included at Level 1 and, at higher levels, listening sections include two or three paragraphs with clusters of two or three items based on each passage. By level 4 the listening passages consist of academic lectures followed by several multiple-choice questions and note-taking is allowed. The descriptors for level 4 describe students as having near-native levels of proficiency:

> "Students at Level 4 are able to understand linguistically complex discussions, including academic lectures and factual reports. Though there may be occasional trouble with colloquialisms, idiomatic language, or rapid native speech, they are able to use context clues to aid comprehension and have acquired an understanding of most discourse markers. They have acquired the ability to comprehend implications, inferences, emotional overtones, differences in style, and shifts in register. Level 4 students understand almost all reductions, elisions, and blends in the spoken language" (http://www.act.org/compass/esl/desc/listen.html).

The ACCUPLACER ESL Listening test includes 20 questions which measure the ability "to listen and understand one or more people speaking in English" (The College Board, 2007a p. 5). ACCUPLACER's description of the Listening test suggests that the input consists of conversations in both everyday and academic contexts, but it is not clear whether lectures are included. Students listen to recorded conversations which take place "in academic environments such as lecture halls, study sessions, a computer lab, the library, the gymnasium, and so forth; and in everyday environments such as at home, at a store, at a restaurant, at a dentist's office, listening to the radio, reading the newspaper, and performing tasks at work" (The College Board, 2007a, p.6). Test takers look at pictures of the speakers while listening.

Both COMPASS ESL and ACCUPLACER ESL describe test takers at the highest level of performance as being able to comprehend spoken material in both everyday and academic situations.

**Writing.** The writing tests of both COMPASS ESL (e-Write) and ACCUPLACER ESL (WritePlacer-ESL) are computer administered and scored based on proficiency in several different domains. The COMPASS ESL e-Write claims to be a direct measure of students' writing in general and is scored "using analytic scoring rubrics, with an overall score derived from a weighting and summing of the assigned analytic scores" (ACT, 2008, p.20). The analytic scoring focuses on five domains: Development (35%), Organization (15%), Focus (10%), Language Use (35%), and Mechanics (5%), which are weighted according to perceived importance to the task of writing. (ACT, 2008, p. 16).

ACCUPLACER WritePlacer ESL, prior to the new version introduced in 2009, also included five domains: Organization, Focus, Vocabulary, Sentence Structure, and Mechanical Conventions. Three of these domains overlap entirely between the writing tests (Organization, Focus, and Mechanics), however the non-overlapping dimensions, e-Write's Development and Language Use and ACCUPLACER's Vocabulary and Sentence Structure differ widely in their scope. ACCUPLACER documentation states that "because Vocabulary and Sentence Structure are skill areas that nonnative speakers need to develop, those dimensions of writing are included in addition to three rhetorical dimensions of writing that apply to both native and nonnative writers" (The College Board, 2007a, p. 10). COMPASS e-Write does not place importance on Vocabulary as a separate skill area, but incorporates it into the larger, heavily weighted Language Use domain. In fact the Language Use and the Development domain together account for 70% of the total writing ability construct in e-Write, while the three overlapping domains only account for 30%. So, although it might seem that the two writing tests are very similar in the sense that they share three domains, in reality they differ greatly as 70% of what is valued in COMPASS e-write is not considered at all in WritePlacer ESL. The two tests therefore define the construct of writing ability differently, meaning that what colleges will learn about what students can do with writing correspondingly varies.

The recently updated ACCUPLACER WriterPlacer ESL scoring rubric measures four dimensions—Word Use, Sentence Use, Grammar, and Organization & Development. The first three dimensions correspond, to some extent, to the dimensions of Vocabulary, Sentence Structure, and Mechanical Conventions in the original rubric. The last dimension in this new WriterPlacer ESL rubric, however, integrates the original domains of Focus and Organization as well as a new Development dimension. Although the revised rubric seems to assess a wider range of writing skills, the structure and presentation of the new rubric, however, indicate that the rubric puts much more emphasis on the basic language skills than on the more higher-order rhetorical skills, when compared to the COMPASS ESL e-write. This means that different aspects of students' writing skills are emphasized quite differently in the two tests. For example, ACCUPLACER's WriterPlacer ESL's "Organization and Development" dimension in the new rubric integrates elements of "Focus," "Organization" and "Development" into one of the four equally weighted analytic dimensions. If these three domains of writing together accounted for 60% of the total writing ability construct in the COMPASS ESL e-write (Focus, 10%,

Organization, 15%, Development 35%), they together only account for 25% of the overall writing ability as defined by the new WritePlacer ESL rubric. Likewise, whereas the domains of "Language use" and "Mechanics" together account for 40% of the total score on writing ability in the COMPASS ESL e-write, the three domains of "Word Use," "Sentence Use" and "Grammar", (which more or less represent the same constructs as "language use and mechanics" in COMPASS), account for a total of 75% of the overall writing score in the new WriterPlacer ESL test. COMPASS ESL e-Write and ACCUPLACER WritePlacer ESL reflect very different representations of the writing construct.

### Description of the COMPASS and ACCUPLACER English Essay Writing Tests

This section describes the new ACCUPLACER WritePlacer Plus, and the COMPASS E-write. There is no essay writing section on the CTEP. The new WritePlacer Plus and E-write both require students to produce a writing sample in response to a prompt. Although WritePlacer Plus topics are not specifically stated, a sample prompt requires students to write an essay on whether "obstacles or disadvantages [can] be turned into something good." WritePlacer Plus prompts are claimed to be "carefully designed so that the student can respond quickly in a variety of ways." In addition these prompts are designed to "stimulate critical thinking and are relevant to any number of fields and interests" (The College Board, 2008, p. 1). WritePlacer scores are based upon the writer's ability to "clearly and effectively express [his or her] position" in writing (The College Board, 2008, p. 1). And scores are reported on a scale of 0-8 using a holistic rubric that evaluates the following writing features: Purpose and Focus, Organization and Structure, Development and Support, Sentence Variety and Style, Mechanical Conventions, and Critical Thinking. Compared with the old version of the rubric used prior to 2009, which included the dimensions of Focus, Organization, Development and Support, Sentence Structure and Usage, and Mechanical Conventions, the new rubric not only includes more analytic dimensions but it also attaches more importance to the discourse level writing skills than the original rubric did. The new Critical Thinking dimension, for instance, focuses on assessing "the extent to which [the student] communicate[s] a point of view and demonstrate[s] reasoned relationships among ideas" (The College Board, 2008, p. 1). Likewise, the change of "Sentence Structure and Usage" in the old rubric to "Sentence Variety and Style" in the new version indicates a shift of focus from the assessment of linguistic features of written language to the evaluation of rhetorical features of written discourse.

Compass E-write is very similar in nature to the ACCUPLACER WritePlacer Plus. In both tasks, students need to take a position about a particular issue and support their opinion with examples and evidence. E-Write topics are "framed within a familiar context. This might be a community or a school setting where a problem or issue related to that setting is presented" (ACT, 2008, p.3). The E-write is evaluated holistically according to how well students formulate a clear and focused position on the issue defined in the prompt, support that position with reasons and evidence appropriate to the position taken and the specified concerns of the

audience, as well as how well they develop the argument in a coherent and logical manner, and express ideas using clear, effective language. Secondary scores are provided which evaluate the test takers' writing ability in more specific areas, including: Focus, Content, Organization, Style, and Conventions. E-write scoring is somewhat complex based primarily on the fact that there are two scoring systems in use simultaneously. Writing prompts designed before 2006 are scored on a 4-point scale and then doubled in order to align scores with the pre-computer rating system. Final scores are reported from 2-8. Prompts designed after 2006 are scored on a 6 point scale, and then doubled to "align with the original scoring model of two rater scores" (ACT, 2008, p. 6). These scores are reported from 2-12. Neither the WritePlacer nor the Compass E-Write essay task is timed.

### Comparison of the COMPASS and ACCUPLACER Essay Writing tests: ESL vs English

**COMPASS Essay Writing: ESL vs. English.** The COMPASS ESL E-write and the English E-write both claim to be direct measures of students' writing ability in general. What distinguishes these two direct writing assessments is the nature of the prompts as well as the scoring rubrics used to evaluate the written product of the test takers. Scrutiny of both kinds of prompts indicates that the ESL prompts are shorter in length and use simpler language and sentence structure. More importantly, the English essay prompt contains more information than the ESL essay writing prompt. Specifically, the English essay prompt provides a concrete situation and a clear purpose for test takers to address in their writing. The audience is also specified in the prompt, and therefore test takers are expected to address this particular audience effectively in their essay. The COMPASS ESL E-write prompt provides no context, no specified audience, and no particular purpose. The topic of the writing is also much simpler and general on the ESL essay test than it is on the regular English essay test.

The ESL E-write is scored analytically with a major focus placed on Development (35%), Organization (15%), Focus (10%), Language Use (35%), and Mechanics (5%). The English E-write is scored holistically and emphasizes advanced writing skills such as audience awareness and content appropriateness in relation to audience and purpose, as well as effective writing styles, which are not evaluated on the ESL writing test.

**ACCUPLACER Essay Writing: ESL vs. English.** The WritePlacer essay test is designed for native speakers of English and "measures writing skill at the level expected of an entry-level college student" (The College Board, 2007a, p. 10). Writing is evaluated according to 6 domains: Purpose and Focus, Organization and Structure, Development and Support, Sentence Variety and Style, Mechanical Conventions, and Critical Thinking. However, in the new WritePlacer ESL, writing is evaluated based on 4 domains: Word Use, Sentence Use, Grammar, and Organization and Development. Although both tests are scored holistically, the ESL test is scored on a 6-point scale whereas the English test is scored on an 8-point scale. A comparison of the domains of writing that are evaluated on these two writing tests reveals that these two

scoring systems are quite different. The former provides a more comprehensive assessment of writing with a focus on the communicative function of a written text, whereas the latter assesses writing more locally and puts an emphasis on the correctness of language rather than the effective use of language.