Listening. Learning. Leading.®

Research Report
ETS RR–11-22

# Examining American Post-Secondary Education

Ou Lydia Liu

May 2011

# Examining American Post-Secondary Education

Ou Lydia Liu
ETS, Princeton, New Jersey

May 2011

As part of its nonprofit mission, ETS conducts and disseminates the results of research to advance quality and equity in education and assessment for the benefit of ETS's constituents and the field.

To obtain a PDF or a print copy of a report, please visit:

http://www.ets.org/research/contact.html

**Technical Review Editor:** Donald Powers

**Technical Reviewers:** Thomas Van Essen and John W. Young



## Table of Contents

## List of Figures

**Abstract**

The purpose of this report is to identify the most prominent issues in U.S. higher education and to develop strategic research plans to address the issues that are most relevant to ETS's capabilities in measurement and assessment through the ETS's higher education research initiative. In the United States, issues related to higher education such as improved performance and effective accountability have received unprecedented attention from stakeholders at many levels. At the national level, President Obama has set forth an ambitious agenda for American postsecondary education such that by 2020, the United States should once again have the largest concentration of citizens with a postsecondary degree. At the corporate level, ETS, as the world's largest educational research and testing organization, is ready to move beyond testing program-based research in higher education and has the capability to deal with some of the most thorny issues in higher education. By strategically expanding post-secondary research, ETS will establish itself as a pioneer and thought leader in the field of higher education.

The first part of the research report identifies four key issues existing in American higher education: *enrollment and performance, retention and degree attainment, student learning and experience,* and *learning outcomes and accountability*. The second part of the research report develops an ETS research agenda with short-term and long-term plans to address these issues. The agenda specifies short-term and long-term research goals that are specific, attainable, and measurable. Research findings from the studies proposed in this agenda have a potential for advancing understanding of the current situation and future needs of American higher education and also contributing to enhanced student learning at postsecondary institutions. Reaffirming and strengthening American higher education is critical to this country's success in the 21st century.

Key words: higher education, retention rates, graduation rates, learning outcomes, accountability

**Acknowledgments**

I want to sincerely thank Cathy Wendler for her continuous mentoring and support during the development of the ETS higher education research agenda. Cathy, Brent Bridgeman, David Payne, Kate Kazin, and John Young also provided valuable comments on an earlier draft of this research report. I also want to thank David and Kate for sharing their insights on the selection of research topics, and providing access to external resources in higher education. I am grateful to Zhitong (Lin) Yang for his assistance with producing the graphs in this research report. I thank Rosalea Courtney for her editorial assistance. I am also grateful to the Validity Research Initiative led by Brent Bridgeman for funding this work.

## Part I: An Overview of Postsecondary Education in the United States

Higher education is considered a springboard to many opportunities in a knowledge society. A college degree is usually associated with more prestigious jobs and higher salaries. In 1990, college graduates on average were paid more than high school graduates, by $11,145 (in 2008 dollars), and the difference increased to $22,720 in 2008 (U.S. Department of Education, National Center for Education Statistics, 2008). A college degree also has the potential for broadening one's experience and cultivating critical thinking and analytical reasoning skills which bring citizens lifelong benefits. Besides personal benefits, enhanced academic achievement also benefits the society at large. A highly educated workforce will contribute effectively to the economic development of the nation. A citizen's college degree attainment will likely translate into a higher probability of degree attainment in the next generation. Society also benefits from high concentrations of college graduates with respect to stability of families, lower crime rates, and more active civic participation (Hill, Hoffman, & Rex, 2005).

In the United States, higher education has become the prerequisite for a highly-skilled and globally-competitive workforce. High-paying jobs once available to high school graduates are no longer available to them due to the shift in the economy from a manufacturing society to a knowledge society (Wendler, Bridgeman, Cline, Millett, Rock, Bell, & McAllister, 2010). President Obama has emphasized the importance of higher education and urged more Americans to obtain a college degree. Higher education has become a gateway to many options and opportunities, including increased financial success and social benefits, and these benefits are likely to compound over one's lifetime (Hill, Hoffman, & Rex, 2005).

Educational Testing Service (ETS) has been an active contributor to U.S. higher education for the past sixty years. Some of the most well-known testing programs such as the SAT® and GRE® are either owned by or operated through ETS. These programs have benefited thousands of institutions in their admission decisions for both undergraduate and graduate programs. ETS offers a wide range of products and services that serve an array of stakeholders in higher education. For example, the ETS Test of English as a Foreign Language (TOEFL iBT®) provides opportunities for international students whose native language is not English to pursue all levels of education in English-speaking countries; the Major Field Tests capture comprehensive learning outcomes gained by undergraduate students in major fields of study; and the ETS Proficiency Profile provides measures for key college-level skills such as critical

thinking, reading, writing, and mathematics. In sum, ETS's assessment portfolio in higher education significantly expands access to college and graduate education by serving students of all backgrounds.

Building on the important roles it plays in U.S. higher education, ETS seeks to undertake more responsibilities to advance this country's post-secondary education. Despite its leading role in the world, higher education in the United States has been thwarted by a variety of long-standing issues such as low retention rates, low degree attainment rates, and ambiguous accountability mechanisms. ETS, as the world's largest educational testing and research organization, is well positioned to undertake research that will address some of the thorniest issues that are most relevant to ETS's capabilities in higher education. These broad-based issues concern many important aspects of college and graduate education and transcend any specific testing programs at ETS.

The purpose of this research report is to identify the most prominent issues in U.S. higher education and to develop strategic research plans to address these (or a subset of these) issues. Findings from this research will advance ETS's understanding of the current situation and future needs of American postsecondary education. The findings will also have the potential for positioning ETS as a thought leader in higher education research, particularly in areas of assessment and measurement where ETS has the strongest capabilities.

## Urgent Issues in American Higher Education

America has been a world leader in college education and graduate studies for the past century (U.S. Department of Education, 2006). The quality and variety of U.S. higher education institutions have made them the "gold standard" worldwide. However, the leading position of the United States is now being seriously challenged. Part of this challenge arises from the continuing high attrition rates, low degree attainment rates (ACT, 2008) and ambiguous criteria for evaluating learning outcomes (Liu, 2009a; U.S. Department of Education, 2006). The situation of U.S. higher education is further complicated by the rapidly changing demographics of prospective students (Kirsch, Braun, Yamamoto, & Sum, 2007), unequal access, ever-increasing tuition and fees, and dim prospects for job placement upon graduation. As America struggles to advance its higher education system, many other countries have experienced booming success with their college education systems. According to the official statistics released by the Chinese Ministry of Education (*www.moe.gov.cn*), the number of Chinese college graduates increased

from 829,070 to 1,594,130 from 1997 to 2007, almost doubling in the last ten years. For the U.S. to maintain its global competitiveness, dramatic changes are needed to improve American postsecondary education.

Through a broad-based review of American higher education, we identified four areas of research that deserve the most urgent attention: enrollment and performance, retention and degree attainment, quality instruction, and learning outcomes. These broad areas are some of the most important aspects of higher education and will continue to be determining factors of the prospect of U.S. higher education. Gaining a deep understanding of these issues will help shape ETS's short- and long-term research agenda in postsecondary education.

## Enrollment and Performance

As of 2007, about 18.2 million students enrolled in degree-granting institutions in the United States, consisting of 7.8 million males and 10.4 females (U.S. Department of Education, National Center for Education Statistics, 2009). Students pursue various levels of degrees (i.e., associate, bachelor's, master's, doctoral, and first-professional) at two-year and four-year colleges and universities.

### Enrollment by Ethnicity

The percentage of high school graduates who enrolled in a degree-granting institution in the same year was 56% for Black students, 62% for Hispanic students, and 72% for White students[1] in 2008 (Figure 1; Aud, Fox, &, 2010).

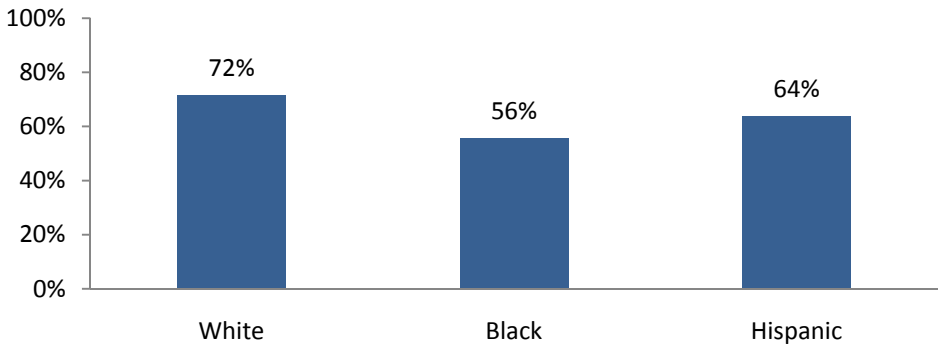


***Figure 1.*** **Percentage of high school completers who were enrolled in 2- or 4-year colleges the October immediately following high school completion, by race/ethnicity: 2008. Data are from Department of Commerce, Census Bureau, *Current Population Survey* (CPS), October 2008.**

The ethnic and racial gaps in higher education enrollment have long been a source of concern. In 2008, only 32% Black and 26% Hispanic 18- to 24-year-olds were enrolled in degree-granting institutions, compared to 44% White and 58% Asian students (Figure 2; Aud, Fox, & KewalRamani, 2010). Among the students who started at a community college, Black and Hispanic students (9.1% and 7.5%, respectively) are also less likely to transfer from community colleges to 4-year colleges than their White and Asian peers (12.8% and 13.8%, respectively).

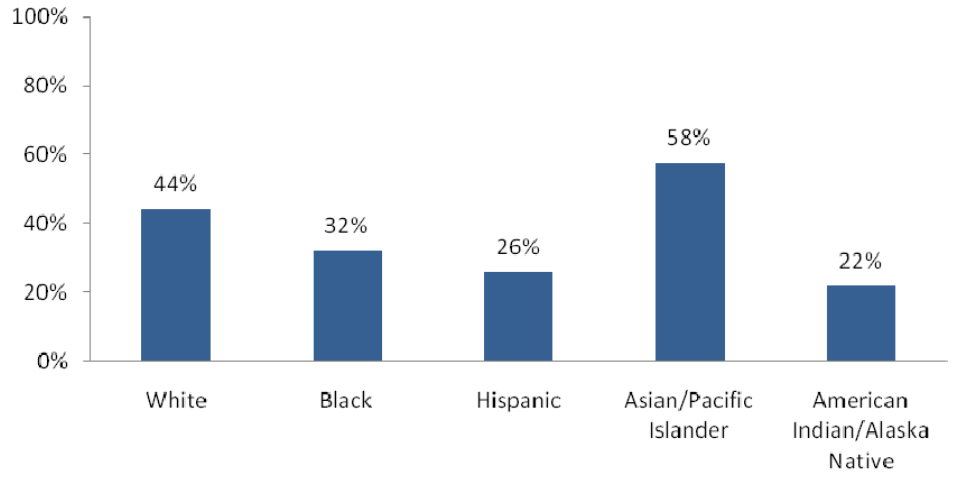


***Figure 2.*** **Percentage of 18- to 24-year-olds enrolled in colleges and universities, by race/ethnicity: 2008. Data are from U.S. Department of Commerce, Census Bureau, *Current Population Survey* (CPS), October 2008.**

The low enrollment rates of Black and Hispanic students may be explained by their performance in high school and also their educational expectations, as students' academic preparation in high school significantly predicts their transition to college. Data from the *Education Longitudinal Study: 2002* suggest that Black and Hispanic students on average earned fewer high school course credits (17.8 and 17.6 respectively) than their White and Asian counterparts (19.0 and 19.5 respectively) (NCES, 2010*a*). When students are less well prepared in high school, they are less likely to go to college. And even when they do enroll in a college, they are more likely to drop out than better prepared students due to limited college readiness (Wendler et al., 2010).

Racial/ethnic groups also show differential course performance in undergraduate education. According to the data from the National Center for Education Statistics (Figure 3), a

considerably larger proportion of Black and Hispanic students (35.58% and 29.77%, respectively) obtain GPAs lower than 2.5 compared to White and Asian students (21.53% and 21.76%, respectively). The performance gap also occurs among the highest performing students. Only 18.92% of Black and 22.96% of Hispanic students achieved GPA at 3.5 or higher as compared to 32.39% of White and 31.45% of Asian students.

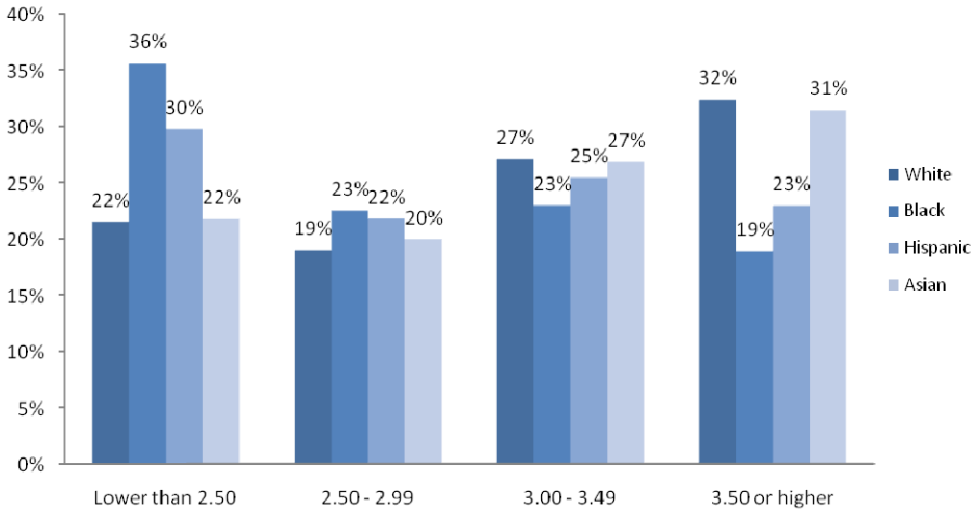


***Figure 3*. GPA by ethnicity. Data are from U.S. Department of Education, National Center for Education Statistics, 2007-08 *National Postsecondary Student Aid Study* (NPSAS: 08).**

Differential performance by race/ethnicity also exists among community college students. Greene et al. (2008) examined the academic performance of 3,143 students enrolled from 2004 to 2008 in 36 Florida community colleges. Based on a hierarchical linear analysis, the authors found that Black and Hispanic students achieved significantly lower grades (-.34 in SD, $p<.001$ and -.09 in SD, $p<.05$, respectively) than their White peers.

Based on the above synthesis, there is an urgent need to increase college enrollment for minority students. To that end, it is important to improve high school graduation rates among minority students and help high school completers experience a smooth transition to college. Research is needed to examine the factors that influence high school students' commitment to going to college, such as high school academic preparation, socioeconomic status, knowledge about college life, aspirations for college, etc..

## Enrollment by Gender

In 2008, the percentage of 18- to 24-year-olds enrolling in colleges and universities was 37% for males and 42% for females (Aud, Fox, & KewalRamani, 2010). Black students showed the largest gender differences in undergraduate enrollment, with females accounting for 64% of the enrollment (Aud et al., 2010). Figure 4 shows the percentage of male and female enrollment at degree-granting institutions. Overall, females comprise about 57% of the entire college population. Although the percentage of females is larger than the percentage of males at all types of institutions, the percentages of females are particularly high at private institutions (66% at two-year and 59% at four-year private institutions). In fact, the percentage of female students has been increasing steadily for about three decades now.

Of the students pursuing postbaccalaureate degrees in 2008, about 59% females and 41% were males. Again, the difference in enrollment was largest for Black students, with females comprising 71% of the Black student enrollment (Aud, Fox, & KewalRamani, 2010).

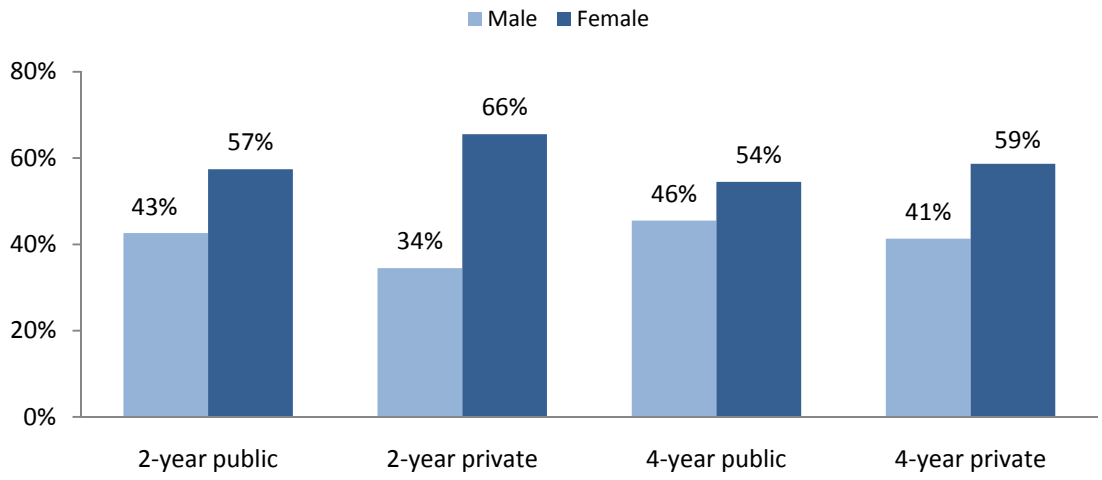


***Figure 4.*** **Percentage of students attending 2- and 4-year institutions by gender. Data are from U.S. Department of Education, National Center for Education Statistics, *Integrated Postsecondary Education Data System* (IPEDS), Spring 2009, Enrollment component.**

## Retention and Degree Attainment

**Retention Rates.** Low retention rates, especially from the first year to second year in college, have been a major setback for American higher education at all types of institutions. Among research/doctoral institutions, about 81% of freshmen return after their first year at

private universities and 74% of freshmen return at public universities. The problem of losing students is particularly serious for two-year colleges (ACT, 2008). Almost half of the students drop out after the first year of study at two-year colleges (Figure 5).

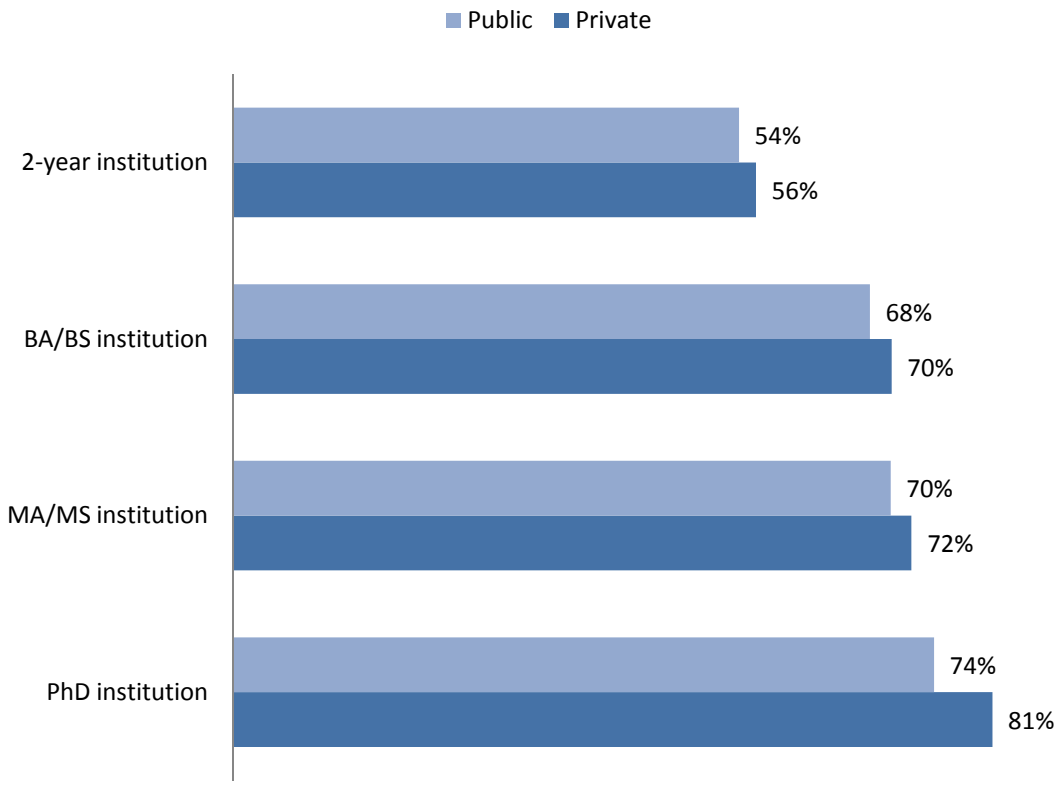


***Figure 5.*** **Retention rate: Freshman to sophomore year Note: The retention rate here refers to the percentage of students returning to the same institutions. It does not include students who transfer to another institution (http://www.act.org/news/releases/2009/1-22-09.html). Data are from ACT (2009). *Retention/Completion Summary Tables.* Iowa City, IA: ACT.**

## Graduation Rates

Completion rates are generally low for all types of institutions. Only 65% and 49% students complete within five years of enrollment at research/doctoral private and public universities, respectively (Figure 6). In general, the best graduation rate is seen at four-year private institutions, including PhD, master's, and baccalaureate institutions, where 55%–65% students obtain a bachelor's degree within five years of enrollment. The worst completion rate is

seen at two-year public colleges, where only 28% obtain a degree within three years of enrollment (Figure 6). Among the students who did not receive a degree from the community college in which they initially enrolled, some of them may have transferred to another community college or transferred to a four-year college or university.

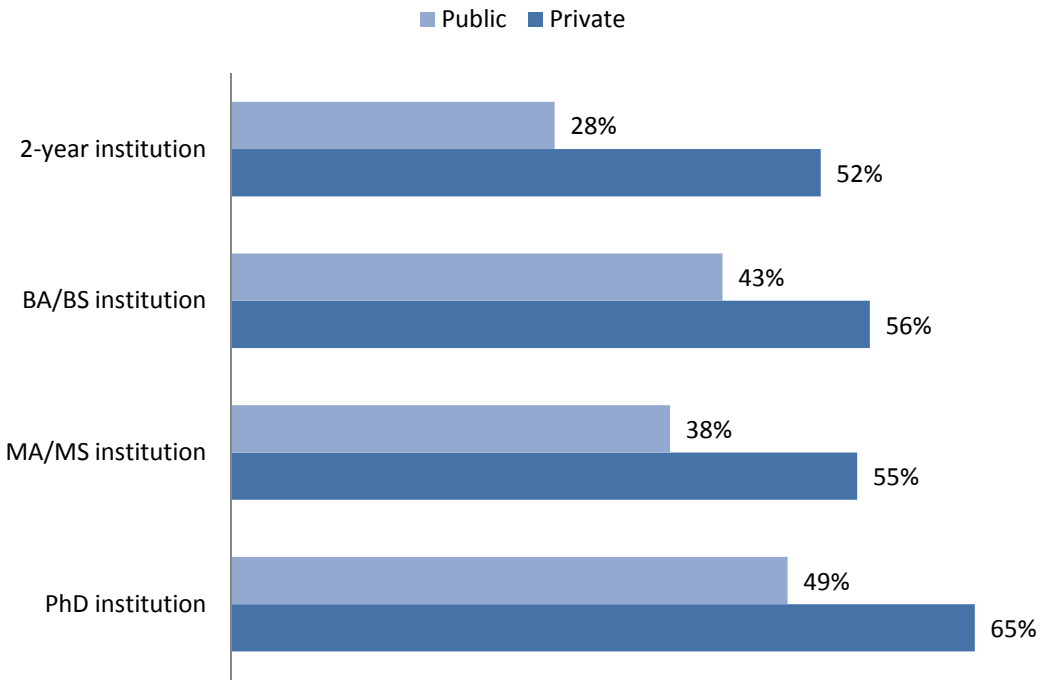


***Figure 6.*** **Completion rate: Two-year and four-year colleges. Data are from ACT (2009).** ***Retention/completion summary tables.*** **Iowa City, IA: ACT.**

**Time to completion.** When examining the graduation rates by the number of enrollment years at two-and four-year institutions (Figures 7 and 8), the general trend is that many students take a considerably longer amount of time to obtain a degree than two or four years. For example, at two-year college, only 19% of the students graduated within two years, 31% of the students graduated within three years, and 37% graduated within four years. Similarly, only 36% of the students obtain a degree within four years at four-year universities, 52% obtained the degree within five years and 57% in six years.

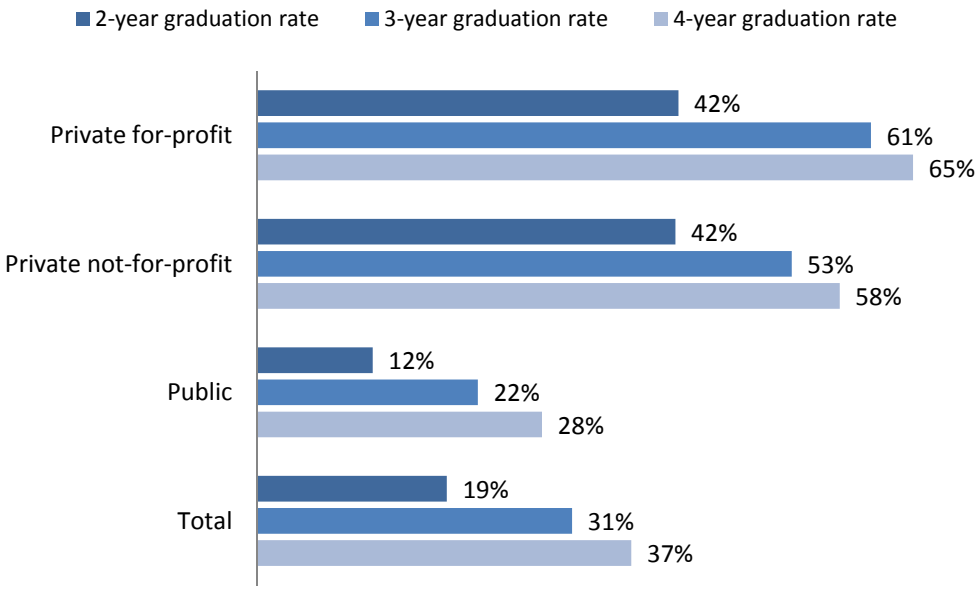


***Figure 7.*** **Two-year college graduation rate. Data are from U.S. Department of Education, National Center for Education Statistics, *Integrated Postsecondary Education Data System* (IPEDS), spring 2009, Graduation Rates component.**

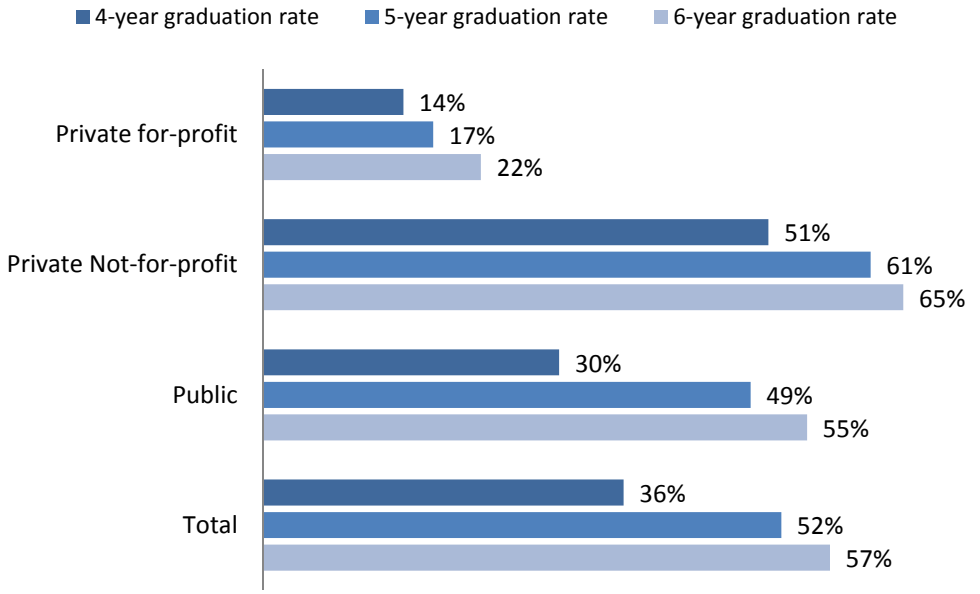


***Figure 8.*** **Four-year college graduation rate. Data are from U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), spring 2009, Graduation Rates component.**

The low retention rate, low graduation rates, and long time to complete present a serious problem for the effective use of institutional resources. They drain financial resources for the students and their families. Without higher retention and graduation rates, America will be hard-pressed to reach the goal of reclaiming itself as a world leader in the concentration of citizens with postsecondary degrees (Kelly, Schneider, & Carey, 2010).

**Degree attainment by race/ethnicity.** Besides the overall low degree attainment rate, the degree attainment rates are particularly low for Black and Hispanic students. Among the students who enroll at a degree-granting institution, Black and Hispanic students are less likely than their White and Asian peers to obtain a degree within six years (NCES, 2005). For students starting at a four-year institution, only 43.4% of Black and 44.0% of Hispanic students completed a bachelor's degree within five years of enrollment as compared to White (61.9%) and Asian (69.1%) students (U.S. Department of Education, National Center for Education Statistics, 1996).

For students who transfer from community colleges to four-year colleges, Black and Hispanic students are also less likely (32.6% and 30.0%, respectively) to earn a bachelor's degree than their White and Asian counterparts (46.7% and 48.8%, respectively; NCES, 2005).

In 2008, about 29% of U.S. adults over the age of 25 had at least a bachelor's degree. Thirteen percent of Hispanics and 20% of Blacks obtained at least a bachelor's degree, as compared to 33% of Whites and 52% of Asians. Thus, it is critical to improve the degree attainment rate for Hispanics and Blacks.

To improve graduation rates, it is important to understand the reasons why students do not complete their studies. Financial factors could be a significant reason, but other factors, such as motivation, academic engagement and career aspirations, could also profoundly influence students' persistence to degree attainment (Bradley & Blanco, 2010; Mullin, 2010). In 2002, the NCES designed the Education Longitudinal Study of 2002 (ELS, 2002) which followed a national sample of tenth grade students as they progressed through high school and on to higher education or the workforce. The study first collected data from a nationally representative sample of students in 2002, and gathered additional information from the same group of students in 2004 and 2006. Results from this study revealed significant differences in educational expectations of students by race/ethnicity (Table 1: NCES, 2010*b*). Only 57% of the Hispanic high school

students expect to receive a bachelor's degree or higher, as compared to 67% Black, 71% White, and 80% Asian students.

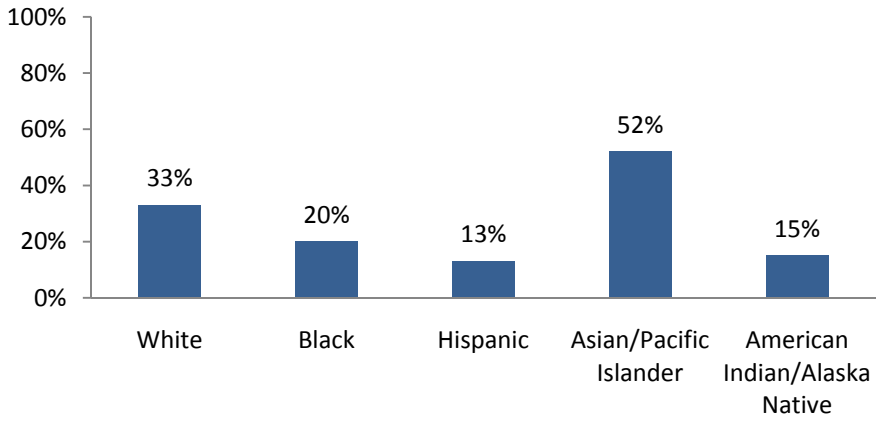


***Figure 9.*** **Degree attainment by race/ethnicity. Data are from U.S. Department of Commerce, Census Bureau, *Current Population Survey* (CPS), *Annual Demographic Supplement,* and *Annual Social and Economic Su*pplement, 2008.**

**Table 1**

***Educational Expectations of High School Seniors by Ethnicity***

| | Don't know (%) | High school or less (%) | Some college (%) | Bachelors (%) | Graduate/advanced degree (%) |
|---|---|---|---|---|---|
| White | 7.0 | 4.7 | 17.3 | 35.0 | 36.0 |
| Black | 8.6 | 5.0 | 18.8 | 32.1 | 35.3 |
| Hispanic | 13.5 | 6.4 | 23.1 | 28.2 | 28.8 |
| Asian | 6.9 | 2.5 | 10.4 | 32.6 | 47.6 |

*Note.* Data from U.S. Department of Education, National Center for Education Statistics, 2002/2006 Educational Longitudinal Survey, 2010.

Going forward, it is of central importance to understand the factors that affect students' persistence in pursuing a college degree from both cognitive and noncognitive perspectives. Contributing factors such as socioeconomic status, change in family status, commitment to earning a degree, change of interest, engagement, career aspirations, and high school academic achievement should all be considered (Museus & Quaye, 2009;

Oseguera & Rhee, 2009). It is also important to explore the possibilities of intervention programs aiming to identify at-risk students and provide effective strategies to help students stay in college. Factors such as motivation, interests, and career aspiration are malleable traits and the enhancement of these traits among at-risk students should be incorporated into the intervention programs. At the institutional level, an institution's size, control, selectivity, expenditure, and financial aid can also contribute to students' degree completion (Oseguera & Rhee, 2009; Oseguera, 2005–2006).

**Degree attainment by gender.** Comparison by gender shows that females are more likely to graduate from four-year universities within five years of enrollment than males, except at private for-profit institutions (Figure 10). Overall, 55% of females graduate within five years at four-year institutions as compared to 49% males. Within each ethnicity, more females than males received degrees in 2008. The gender difference in degree recipients is again most prominent for Black students, with females receiving 68% of the degrees awarded in 2008 (Aud et al., 2010).

Examination of educational expectations by gender unveils some interesting findings. Although similar percentages of male and female students expect to attain a bachelor's degree, females (41.3%) have much stronger aspirations for graduate and advanced degrees than males (28.9%) (Table 2).

**Table 2**

***Educational Expectations of High School Seniors by Gender***

| | Don't know (%) | High school or less (%) | Some college (%) | Bachelors (%) | Graduate/advanced degree (%) |
|---|---|---|---|---|---|
| Males | 9.4 | 6.9 | 20.5 | 34.4 | 28.9 |
| Females | 7.4 | 3.1 | 15.6 | 32.6 | 41.3 |

*Note.* Data from U.S. Department of Education, National Center for Education Statistics, 2002/2006 Educational Longitudinal Survey, 2010.

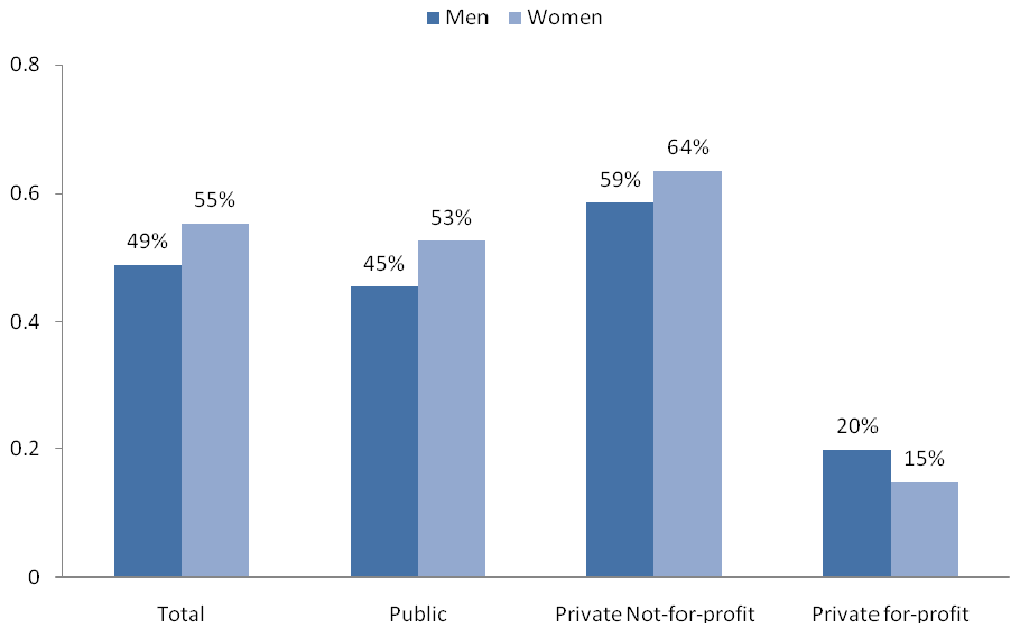


***Figure 10.*** **Degree attainment rate by gender. Data are from U.S. Department of Education, National Center for Education Statistics, I*ntegrated Postsecondary Education Data System* (IPEDS), spring 2009, Graduation Rates component.**

## Degree Major Composition

Among the declared bachelor's degrees granted in 2006–2007, business and health sciences were the most popular majors (Figure 11). It is noteworthy that a very low percentage of students receive bachelor's degrees in mathematics (1%) and physical sciences (1%). For master's degrees, education (29%) and business (25%) had the largest number of graduates (Figure 12). Some of the most popular doctoral specializations (Figure 13) include health sciences (14%), education (14%), engineering & technology (13%), and biological and biomedical sciences (11%). Figure 14 shows the first-professional degrees.

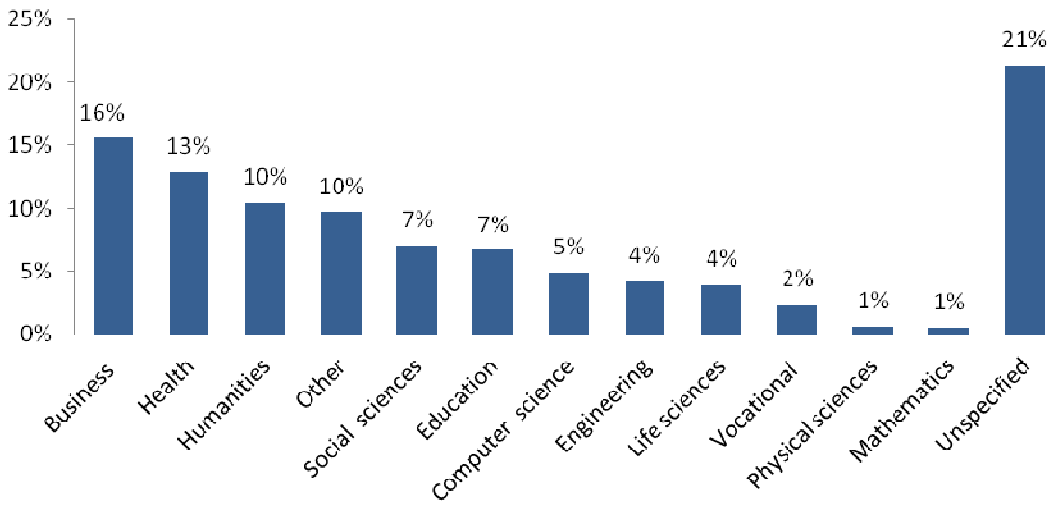


***Figure 11.*** **Bachelor's degrees: Major composition. Data are from U.S. Department of Education, National Center for Education Statistics,** ***2003–04 National Postsecondary Student Aid Study*** **(NPSAS:04).**

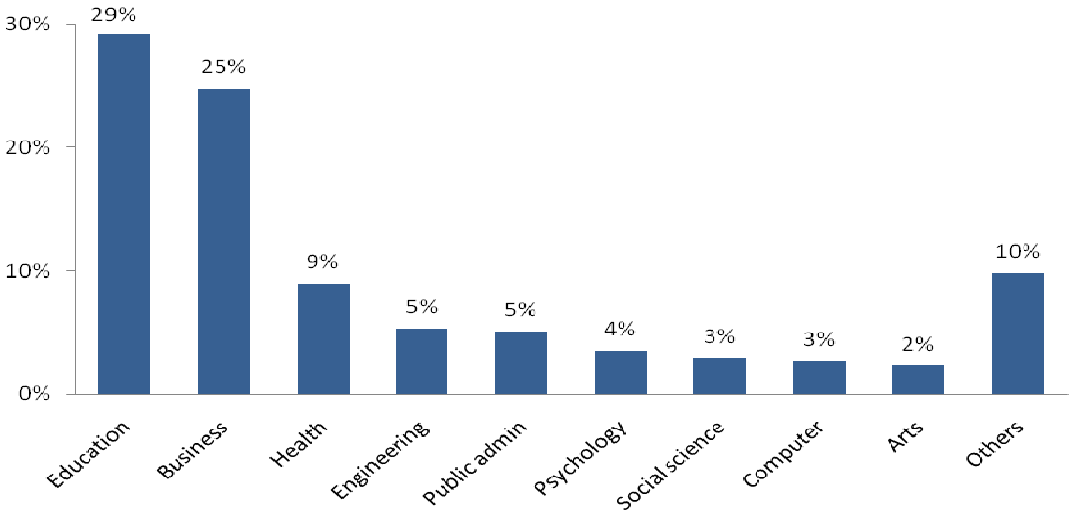


***Figure 12.*** **Master's degrees: Major composition. Data are from U.S. Department of Education, National Center for Education Statistics (NCES), 1996–97 and 2006–07** ***Integrated Postsecondary Education Data System, Completions Survey*** **(IPEDS-C:97) and fall 2007.**

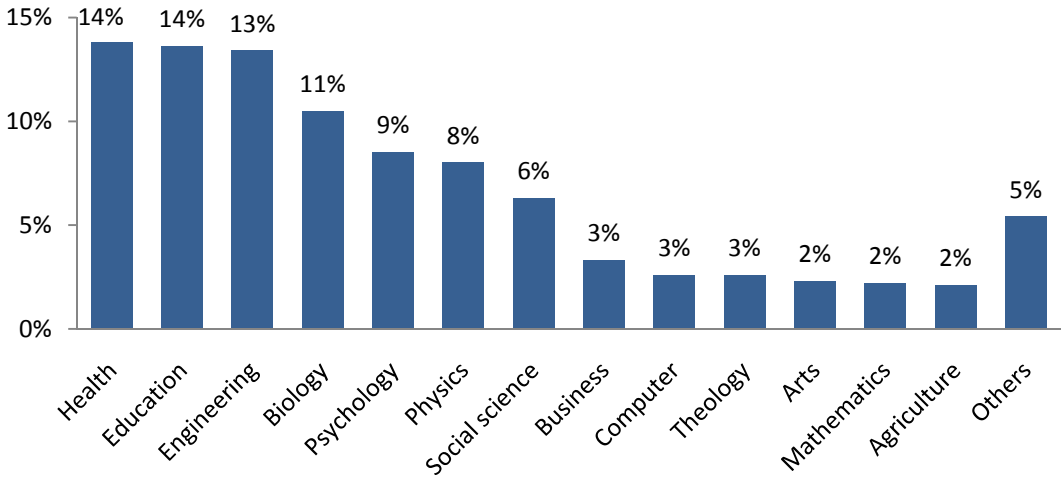


***Figure 13.*** **Doctoral degrees: Major composition. Data are from U.S. Department of Education, National Center for Education Statistics (NCES), 1996–97 and 2006–07 *Integrated Postsecondary Education Data System, Completions Survey* (IPEDS-C:97) and fall 2007.**

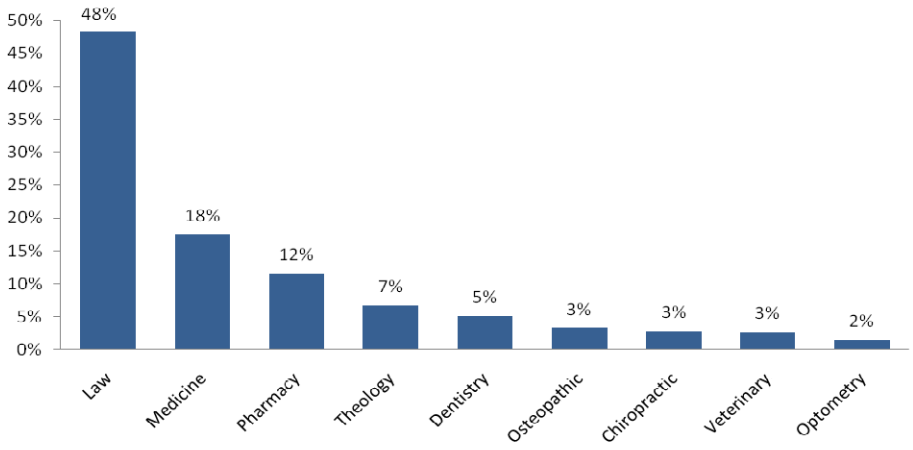


***Figure 14.*** **First-professional major composition. Data are from U.S. Department of Education, National Center for Education Statistics (NCES), 1996–97 and 2006–07 *Integrated Postsecondary Education Data System, Completions Survey*, (IPEDS-C:97) and fall 2007.**

The racial/ethnic differences in degree attainments continue to be amplified in engineering, math, and physical science majors. There are very few African American and

Hispanic students receiving a bachelor's degree in these fields (Figure 15; Snyder, Dillow, & Hoffman, 2009).

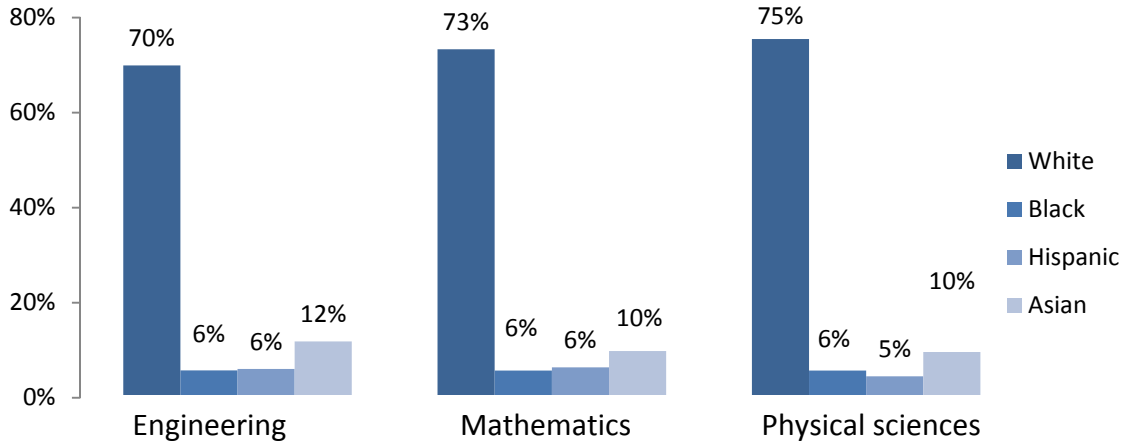


***Figure 15.*** **Ethnic composition of degree holders by major field. Data are from U.S. Department of Education, National Center for Education Statistics, *Higher Education General Information Survey* (HEGIS), *Degrees and other formal awards conferred surveys,* 1976-77 and 1980-81; and 1989-90 through 2007-08 *Integrated Postsecondary Education Data System, Completions survey* (IPEDS-C:90-99), and fall 2000 through fall 2008.**

## Student Learning and Experience

This literature search has identified two issues involving student learning and experiences that need further research: (a) student evaluation of instruction and learning, and (b) the evaluation of the effectiveness of distance education.

## Student Evaluation of Instruction and Learning

Many institutions rely on student self-report surveys for evaluating instruction and student learning and experiences in college. The National Survey of Student Engagement (NSSE, *http://nsse.iub.edu/*) is one of the most widely used surveys of student engagement and learning at four-year colleges and universities. The Community College Survey of Student Engagement (Center for Community College Student Engagement, 2009) is another widely used survey designed for students at community colleges. Institutions use results from

the surveys to understand how students spend time in college and identify aspects of the college experiences that can be possibly improved.

Besides using standardized instruments to gather information on student engagement and experiences, many institutions also use standardized instruments for student evaluation of instruction. For example, the SIR II™ and eSIR II programs offered by ETS are designed to measure student perception of eight dimensions of instruction: Course organization and planning; faculty communication; faculty/student interaction; assignments, exams and grading; instructional methods and materials; course outcomes; student effort and involvement; and course difficulty, workload and pace. Effective use of the information will help institutions evaluate and possibly improve the quality of learning and experiences provided to their students.

Student evaluation of instruction is an important means to improve teaching quality and help instructors adjust their teaching practices to suit students' needs. Student ratings of a course could be influenced by a number of factors including the course type, the delivery mode (i.e., traditional or online), gender of both the student and the instructor, student performance in the course, student perceived difficulty of the course, and the evaluation instruments per se (Davidovitch & Soen, 2006).

Research shows that instructors' gender is related to student ratings of instructors. Female instructors consistently receive lower ratings than male instructors do (Centra & Gaubatz, 2000; Basow, 1995). Instructor gender also interacts with the major field of the course. Basow (1995) found that it was mainly in natural sciences courses that male professors received some of the highest ratings and it was mainly in humanities that female professors were rated highest. Student gender may also impact the ratings they give to their instructors. Both male and female students tend to give higher ratings on enthusiasm to their same gender professors (Centra & Gaubatz, 2000; Basow, 1995).

The evaluation instruments could also have an impact on student rating. Based on data from 5,616 students at a large university, Landrum and Braitman (2008) found that when the

instrument was changed from a 10-point response scale to a 5-point one, student mean ratings of the instructor decreased significantly ($p<.05$).

As the evaluation results are used by institutional administrators to make various decisions related to the instructor and the course, it is necessary to further the research on evaluation instruments to examine their reliability, validity, and usage by students of different gender, language, and major field of study. It is also important to examine the relationship between students' self-reported learning and experiences—and their actual course performance.

**Experience and Learning in Distance Education**

The issue of student learning experiences applies to distance education even more, as students enrolled in distance education often lack the opportunities for face-to-face interactions with the instructor and classmates. The past 20 years have witnessed a steady increase in opportunities related to distance learning in the United States. A prominent feature of distance education is that it is an educational situation where the instructor and the students are separated by time, location, or both.

According to Khan (1997, p. 6), distance education can be defined as "hypermedia-based instructional program[s], which utilizes the attributes and resources of the World Wide Web to create a meaningful learning environment where learning is fostered and supported." Given that computers and use of the internet are the integral features of modern distance education, distance education and online education are used interchangeably hereafter in this paper.

In 2002, about 1.6 million students enrolled in at least one online course at a degree-granting institution, constituting about 9.6% of the total enrollment. In 2007, the number of online students rapidly increased to 3.9 million and comprised 21.9% of the total enrollment (Allen & Seaman, 2008).

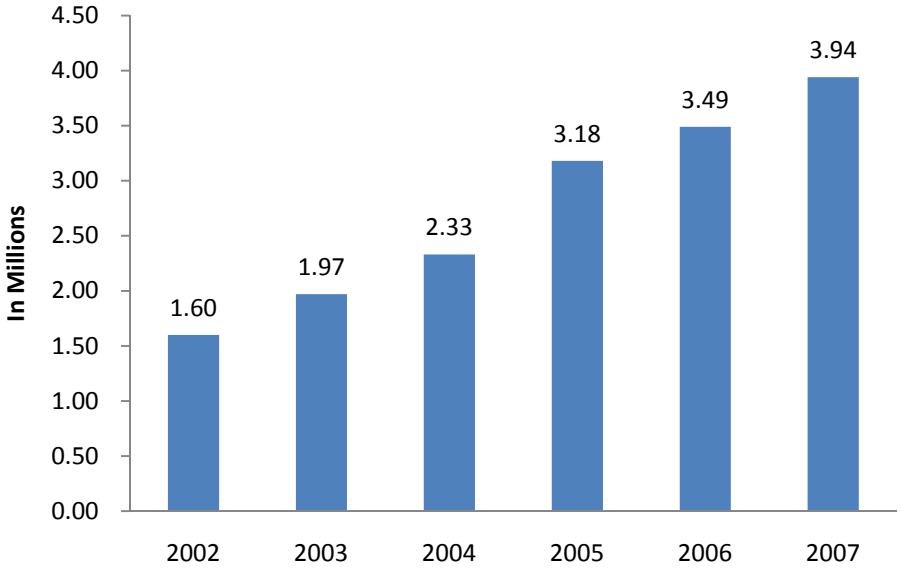


***Figure 16.*** **Number of students taking at least one online course. Data are from Kuh, G., & Ikenberry, S. (2009).** ***More than you think, less than we need: Learning outcomes assessment in American higher education.*** **Champaign, IL: National Institute for Learning Outcomes Assessment (NILOA).**

Compared to traditional in-class education, online education relaxes the constraints of time and location. Students who enroll in online courses no longer need to physically appear in a classroom at a fixed time, but have the flexibility to learn the materials at a time and location that is convenient to them. Online education has benefited many working adults who are motivated to pursue further education for career advancement but cannot afford the time required to attend a traditional class. For example, Wal-Mart, the largest retailer in the United States, has chosen American Public University, an online university, to provide training and credits for its employees in subjects like ethics and retail inventory management (Parry, 2010). Online education also breaks down national boundaries by being available to students in foreign countries. The current economic downturn has fueled the development of online education. A recent study (Allen & Seaman, 2008) indicated that rising unemployment has propelled many working professionals to register for online courses. The rising fuel cost is another reason that more students are now choosing online education over traditional classroom-based education.

However, online education is not expanding without any controversy. The value of online education compared to classroom-based education has been constantly questioned by academic

officers, faculty members and the public. For example, when asked about whether online education is critical to an institution's long-term strategy in 2007, 15% of the academic officers at 2,577 U. S. postsecondary institutions disagreed, increasing from 13.1% in 2002 (Allen & Seaman, 2008). Among all types of institutions, baccalaureate institutions are the ones that reported the lowest rating of the importance of online education, at 35.4%, as compared to 65.8% at masters' institutions and 54.8% at doctoral/research institutions (Allen & Seaman, 2008).

The public's reservation towards online education stems from a number of factors. Among these, the foremost concerns are student experience and the quality of online education (Zhao, Lei, Yan, Lai, & Tan, 2005). In an online educational context, instructors tend to monitor student learning less, due to the lack of in-class interactions and no face-to-face meetings. In addition, the evaluation methods that can be used to test students' mastery of knowledge are limited since students will not be required to be present in a classroom to take a quiz or an examination, which may be particularly important for certain disciplines. For example, a mathematics course may require a proctored math exam while other courses that use essays as an evaluation method may not have this requirement.

Given the potential limitations of online education, it becomes critically important for the online programs to demonstrate their validity, particularly in the areas of student experience and learning outcomes. Many studies have examined the effectiveness of distance education, but most use nonexperimental design and therefore the generalizability of the conclusions from these studies is limited (Lee, Driscoll, & Nelson, 2004). Meta-analysis has become an effective way to synthesize the findings on learning outcomes of distance education. Zhao et al. (2005) conducted a meta-analysis investigating the effectiveness of distance education. They reviewed 8,840 papers relevant to distance education, published 1966–2002, and selected 51 studies which met their five selection criteria (i.e., journal article, complete reference information, inclusion of evaluation, comparison of distance education and face-to-face instruction, and inclusion of empirical data). The meta-analysis was focused on the effect size (*d*) of the differences in outcomes (e.g., course grades, student evaluation of learning, student satisfaction, student participation, etc.) between the distance and face-to-face instruction. Zhao et al. found no significant difference between the two instructional methods. However, when factoring in the year of publication, the authors discovered that studies published before 1998 reported smaller effect sizes than studies published later. Their explanation was that technologies developed

rapidly after 1998, thus enabling more effective and powerful delivery of distance education than before.

In addition to distance education, there is *blended education*, which is a hybrid of traditional, face-to-face education and online education (Rovai & Jordan, 2004). The online component can be a natural extension of classroom instruction. Blended education has become a popular instructional mode in higher education given its flexibility and convenience and also the opportunity for face-to-face instruction. Blended education is designed to maximize the use of institutional infrastructure and ensures a certain amount of student-instructor interaction. If designed appropriately, blended education can contribute to reduced student withdrawal and improved student success (Dziuban & Moskal, 2001).

A decade into the 21st century, distance education has demonstrated a strong momentum for continuing growth, both in the United States and around the world. Many issues await answers in reexamining the role of distance education in advancing knowledge and skills. For example, what features of distance education are associated with more student success? What kinds of students are best suited for distance education? Also, do students have the same experiences and learning outcomes in distance education as in traditional classrooms? Further research is needed to provide insight into these critical questions about distance education.

**Learning Outcomes and Accountability**

Accountability in higher education has recently received unprecedented attention from many levels of stakeholders, including the Federal Government, accrediting organizations, higher education organizations, and the public. One of the critical issues on accountability is how institutions can provide evidence of student learning that is comparable across institutions.

Many American colleges and universities use standardized outcomes assessment to evaluate student learning. For the past five years, at least 1,387 institutions (41% are four-year public universities and 25% are two-year public colleges; Figure 17) have used at least one form of standardized outcomes assessment (ETS, 2010). A recent survey of 2,809 regionally accredited, undergraduate-degree-granting, two- and four-year public and private institutions revealed that 39% of the institutions use one form of standardized measure of learning outcomes (National Institute for Learning Outcomes Assessment, 2009). Compared to subject-specific standardized tests commonly used in K–12, the standardized assessment in higher education tends to measure

general knowledge and skills deemed important for college students, such as critical thinking, reading, writing, and quantitative reasoning. Institutions use outcomes assessment data for a variety of purposes: fulfilling accreditation requirements, responding to accountability calls, informing strategic planning, and modifying general education curriculum. Accountability-related uses have gained important momentum, given the attention accountability has received in higher education. The U.S. Department of Education established a Commission on the Future of Higher Education in 2005, with the aim to improve higher education in the United States. Accountability is one of the four areas identified by the Commission that need urgent reform. The Commission's first report released in September, 2006 (U.S. Department of Education, 2006) comments on the nation's "remarkable absence of accountability mechanisms to ensure that colleges succeed in educating students" (p. *x*) and points out that accountability is vital to ensuring the success of reforms in the other three key areas (i.e., access, quality of instruction, affordability). This report also calls for solid evidence of how much students have learned in college and emphasizes that their evidence should be comparable across institutions.

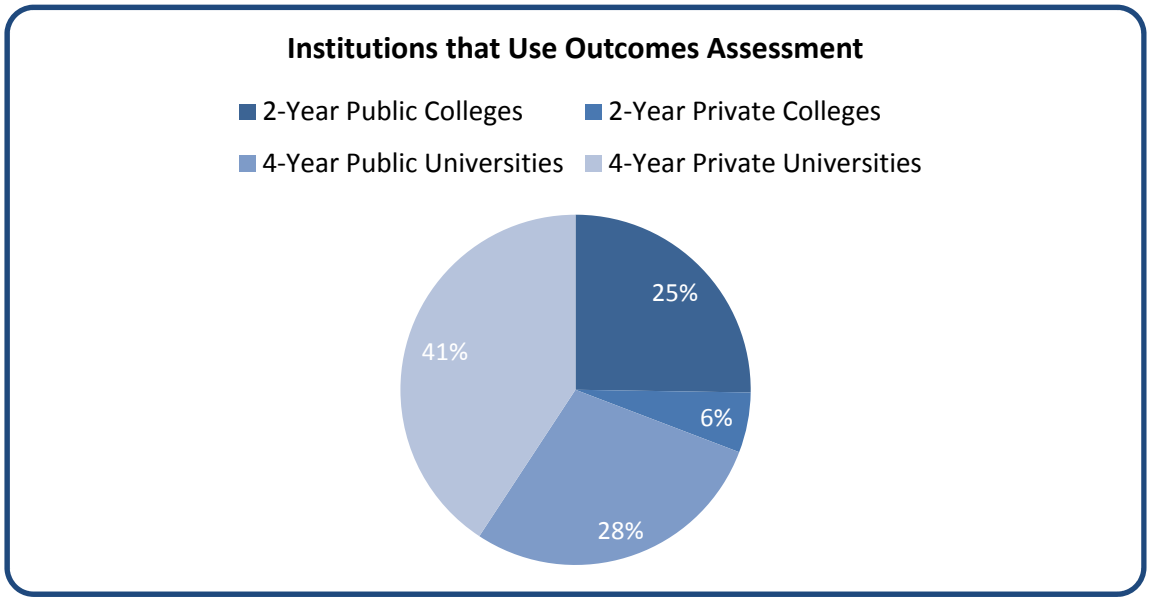


***Figure 17.*** **Institution composition for use of outcomes assessment From *Market research of institutions that use outcomes assessment* by ETS, 2010. Princeton, NJ: Author. Copyright 2010 by Educational Testing Service.**

The development of accountability initiatives serving different purposes and representing different segments of higher education has expanded rapidly over the last several years. For instance,

the Voluntary System of Accountability (VSA; *http://www.voluntarysystem.org*), developed by the American Association of State Colleges and Universities (AASCU) and the Association of Public and Land-Grant Universities (APLU), provides an opportunity for public colleges and universities to demonstrate evidence of student learning, using standardized outcomes assessment. The VSA initiative has attracted participation from 331 institutions as of 2010, representing 63% of AASCU and APLU members. All VSA participating institutions volunteer to use one form of standardized outcomes assessment to report their students' learning outcomes. The Transparency by Design initiative, initiated by WCET (*http://www.wcet.info/2.0/index.php?q=TransparencyByDesign*), a division of the Western Interstate Commission for Higher Education, aims to provide quality assurance on the learning outcomes of distance education institutions serving adult learners. Similarly, the Voluntary Framework of Accountability (VFA; *http://www.aacc.nche.edu/resources/aaccprograms/vfa/Pages/default.aspx*), sponsored by the American Association of Community Colleges, the Association of Community College Trustees (ACCT) and the College Board, tends to measure outcomes and processes at the community college level.

Despite the wide use of outcomes assessment, there is a serious lack of research evidence showing how the assessments should be implemented and how scores should be interpreted. Many institutions administer outcomes assessment to their freshmen and seniors and use the score difference to determine an institution's "value-added". However, problems in the sampling technique and value-added methodology potentially threaten the validity of the findings. The following section provides an overview of value-added research in higher education and summarizes four areas of research that need the most urgent attention.

**Value-Added Research in Higher Education Using Outcomes Assessment**

More space is devoted to this topic than other topics since value-added research in higher education is of particular interest and has generated controversial debates. Value-added research has been conducted widely in K–12. However, in the field of higher education, value-added research is new and relatively rare. Council for Aid to Education (CAE) first included a value-added component in its Collegiate Learning Assessment (CLA) score report (CAE, 2007). This value-added method was later adopted by VSA.

In VSA, the value-added of an institution is defined as the performance difference on a standardized test between freshmen and seniors at that particular institution, after controlling for admission test scores (e.g., SAT or ACT scores). In current VSA practice, the freshmen and seniors represent different groups of students. Therefore, it is a cross-sectional design, a longitudinal design, in which the *same* group of students is tested twice, once when they are in their freshman year and again when they are in their senior year. The longitudinal design is the preferred design, but the cross-sectional design is currently used because of its practical advantages (Voluntary System, 2008). It is easier and also costs less to test two groups of students at the same time than to track the same group of students over a number of years. Given the relative recency of value-added research in higher education, no studies have been conducted in this context to investigate the comparability in results between the cross-sectional and longitudinal designs.

The VSA currently uses Ordinary Least Squares (OLS) regression models to compute value-added scores for institutions. The analysis is conducted at the institution level, with standardized test scores as the dependent variable and mean SAT or ACT scores as the independent variables (VSA, 2008) (The ACT scores are converted to the SAT scores so that they are on a common metric). Using the ETS Proficiency Profile as an example, the value-added computation follows three steps:

1. Institution mean SAT score is used to predict mean Proficiency Profile score in an OLS regression model. This analysis is conducted separately for freshmen and seniors;
2. a residual score is calculated based on the OLS model from step (a) separately for freshmen and seniors, respectively. The residual is the difference between the observed and expected mean ETS Proficiency Profile score, given the mean SAT score; and
3. the difference in the residual scores between the seniors and freshmen determines the final value-added score for that institution.

The final value-added score, essentially, is a *difference in differences*. Each institution is then ranked to ten decile groups on the basis of the value-added score. For example, if an institution is

ranked in group 9, it suggests that this institution has yielded more value-added than 80% of the institutions included in the analysis.

**Current Challenges in Outcomes Assessment in Higher Education**

Although the Voluntary System of Accountability holds great potential for providing direct assessment of student learning and making results comparable across institutions, several challenges exist in its design, value-added methodology, and interpretation of the value-added scores. In this section I summarize the most prominent challenges in the implementation and interpretation of outcomes assessment.

Challenge One: Insufficient evidence of student motivation in taking low-stakes tests. Although outcomes assessments have important implications for institutions, they bear little or no direct relationship to individual students. Students' low motivation in test-taking poses a threat to both the test results and the validity of the score interpretations (Haladyna & Downing, 2004; Wise & DeMars, 2005, 2010; Wise, Wise, & Bhola, 2006). A useful theoretical basis for evaluating student test- taking motivation is the expectancy-value model, with *expectancy* referring to students' beliefs that they can successfully complete a task and *value* referring to the beliefs that it is important to complete a task (Pintrich & Schunk, 2002). Based on this theoretical model, researchers developed self-report surveys to measure student motivation in taking low-stakes tests. The Student Opinion Survey (SOS; Sundre, 1997, 1999; Sundre & Wise, 2003) is one of the widely used surveys capturing students' reported effort and perceived importance of the test. A general conclusion from studies investigating the relationship between student motivation and test performance is that highly motivated students tend to perform better than less motivated students (Cole & Osterlind, 2008; O'Neil, Sugrue, & Baker, 1995/1996; Sundre, 1999; Wise & DeMars, 2005; Wise, Wise, & Bhola, 2006). A meta-analysis of 12 studies consisting of 25 effect sizes (Cohen, 1988) showed that the mean performance difference between motivated and unmotivated students could be as large as .59 standard deviation (Wise & DeMars, 2005). Besides relying on student self-report, researchers also examine response time effort (RTE) for computer-based, unspeeded tests to determine student motivation (Wise & Kong, 2005; Wise & DeMars, 2006). Results show that RTE is significantly correlated with student self-report motivation and is also a significant predictor of their test performance.

To eliminate the impact of student low test-taking motivation on test results, researchers explored ways to filter responses from unmotivated students identified through either their self-report or their response item effort (Wise & Kong, 2005; Wise & DeMars, 2005, 2006; Wise, Wise & Bhola, 2006). Findings are consistent that after controlling for student general ability (e.g., SAT scores), motivation filtering helps improve student performance and improve the validity of the test results (Wise & DeMars, 2005, 2010; Wise, Wise, & Bhola, 2006; Wolf & Smith, 1995).

Realizing the important impact of student motivation on test results, researchers explored ways to boost student motivation in taking low-stakes tests. Common practices include increasing the stakes of the tests (Sundre, 1999), providing extra monetary compensation for performance (Baumert & Demmerich, 2001; Braun, Kirsch, & Yamamoto, 2010; O'Neil, Abedi, Miyoshi, & Mastergeorge, 2005; O'Neil, Sugrue, & Baker, 1995/1996), and providing feedback (Baumert & Demmerich, 2001; Wise, 2004). Increased stakes and extra payment for performance are proven to be effective ways to motivate students (Braun, Kirsch, & Yamamoto, 2010; O'Neil, Sugrue, & Baker, 1995/1996; Sundre, 1999). Despite the intuitive appeal of providing feedback, it has not been demonstrated to have an impact on either student motivation or their performance (Baumert & Demmerich, 2001; Wise, 2004).

Challenge Two: Unclear evidence of what outcomes assessment relates to. Besides showing student progress on core educational outcomes required by VSA, it is critical for the outcomes measures to demonstrate how they relate to other important success indicators, such as cumulative college GPA, retention, degree-completion, and graduate school application. Very few existing studies answer these questions. Among the few, Hendel (1991) examined the relationship between student performance on the *Academic Profile*[2] (Proficiency Profile's predecessor) and college GPA and reported a significant correlation between these two. Marr (1995) also reported that students with higher GPA tended to obtain higher Proficiency Profile scores. Before the inception of the CLA test, Klein, Kuh, Chun, Mamilton, and Shavelson (2005) investigated the relationship between prototype CLA test batteries and student GPA. The test batteries included combinations of GRE-type writing prompts, a critical thinking test (Ewell, 1994) and performance tasks (Klein, 1996). The reported correlation between the test batteries and cumulative GPA ranged from .51 to .64 (Klein et al., 2005). Except for results from these

few studies, very little is known about how student scores on outcomes assessment are related to other academic achievements.

VSA is also interested in providing information on student learning experiences and engagement in college. For example, VSA uses the National Survey of Student Engagement (NSSE, 2008) scale, among other measures, to document student experiences in undergraduate education. However, there is limited evidence on how student self-report experience relates to their performance on outcomes assessment. It is important to find out whether positive experiences are associated with enhanced learning in college. Even more, institutional administrators, business leaders, and other stakeholders need to understand whether scores on outcomes tests can predict student success beyond undergraduate education. For example, are high performing students on outcomes assessment more likely to apply for graduate programs or to be more successful in the work force? Such information is needed to allow a better interpretation of the results from outcomes assessment.

**Challenge Three: Methodological issues with the current value-added method**. As described earlier, the current method used to determine value-added scores in VSA employs OLS regression models for freshmen and seniors, with institution mean standardized test score as the dependent variable and mean admission test score as the independent variable. There are a few potential problems with this method. First, the institution is the unit of analysis in the current model and researchers are concerned that all of the student-level information is ignored (Klein, Freedman, Shavelson, & Bolus, 2008; Liu, 2009b, in press). Analysis results could be more reliable using student-level information given that the number of students far exceeds the number of institutions in the equation. Second, the current method uses admission scores (e.g., SAT or ACT) as a sole predictor in calculating value-added scores. There are many other factors that could influence student learning growth in college at both the student and institution level. Student variables such as gender, ethnicity, and language status, and institutional characteristics such as selectivity, type of institutions (private or public) and resources could all have a significant impact on student performance (Liu, 2008, 2009a, in press; Borden & Young, 2008). Finally, the current method uses OLS regression models to analyze student performance on outcomes tests. Given the hierarchical structure of the data with students nested within an institution, multi-level modeling (Rabe-Hesketh & Skrondal, 2008; Raudenbush & Bryk, 2002) may be more appropriate than OLS models. One of the assumptions of the OLS models is that all

the observations are independent (Stone, 1995). In the case of the value-added calculation for VSA purposes, student learning is likely to be affected by the unique characteristics of the institution they attend. Therefore, test scores of students attending the same institution cannot be considered independent. As a solution to this issue, multi-level modeling relaxes this constraint by differentiating the variance in student performance due to within-institution factors and between-institution factors (DiPrete & Forristal, 1994; Rabe-Hesketh & Skrondal, 2008; Raudenbush & Bryk, 2002; Singer, 1998). Empirical studies show that by switching from OLS method to multi-level method, the value-added results for higher education institutions become more reliable and more stable across years (Liu, in press; Steedle, 2010)..

**Challenge Four: No evidence of the comparability of results between the preferred longitudinal design and the current cross-sectional design.** VSA currently adopts a cross-sectional design in value-added calculation. Institutions often test a group of freshmen and a group of seniors to obtain value-added results. Since the freshmen and seniors are not the same group of students, the value-added results are confounded by the differences between these two groups of students to an unknown degree. The cross-sectional design assumes that the freshman and senior classes have comparable student characteristics including entering academic ability. However, the senior class is often more selective than the first year class due to reasons such as attrition. Liu (2009b) examined the difference in SAT scores between 4,373 freshmen and 1,823 seniors from 23 institutions. Results showed that the senior students had a significantly higher mean SAT score than freshmen and the score difference could be as large as .57 standard deviation on a 1,600-point score scale.

Given the lack of comparability between freshmen and seniors, a longitudinal study is needed to track the same group of students from enrollment to graduation. This way, results between the longitudinal and the cross-sectional designs can be compared. As the VSA overview (VSA, 2008) points out, the cross-sectional design was selected because it is “quicker, simpler, and less costly to implement.” However, the choice of the cross-sectional design was not supported by any research evidence. Such evidence is needed for the cross-sectional design to be implemented at a larger scale in determining institutional value-added. If it becomes apparent that the longitudinal design is superior to the cross-sectional design, then the practice of using a cross-sectional design should be revisited.

Besides these above challenges, we also need to consider the representativeness of the student sample recruited to take the test. Given the rapidly increasing use of outcomes assessment, there is an urgent need to address these critical issues for the assessment results to be used and interpreted in a valid way for accountability purposes.

It's noteworthy that most of the current outcomes assessments are designed for four-year colleges and universities, and very few assessments are designed for community colleges. With the rapid development of the accountability initiative (i.e., Voluntary Framework of Accountability) for community colleges, there will be an urgent need for outcomes assessment specially tailored for community college students.

## Part II: An ETS Research Agenda for Higher Education

Four broad areas of research have been identified in the above review: enrollment and performance, retention and degree attainment, student learning and experiences, and learning outcomes and accountability. These four research areas capture some of the thorniest issues in American higher education. The following section is devoted to an ETS research agenda that outlines short-term and long-term research goals in addressing some of the issues raised in the review. In developing the research agenda, considerations were given to a number of factors including the urgency of the issue, its relevance to ETS capabilities and vision, and the resources needed to undertake the research activities. Considerations were also given to the alignment between the research studies proposed in the agenda and ETS existing research initiative projects in terms of depth, scope, and timeline.

Although graduate education is an important segment of higher education, this research report tends to focus on community colleges and undergraduate education. Many of the graduate-level issues are summarized in the Wendler et al. (2010) report.

### What Are the Short-Term Goals of the Research Agenda?

The short-term goal of the research agenda is to move ETS beyond testing program-based research in higher education. A large proportion of the current higher education research ETS is undertaking is focused on specific testing programs. Addressing issues in the higher education realm will allow ETS to contribute to solving some of the higher-level and broader issues in postsecondary education. Given ETS's reputation in educational testing and its existing research in outcomes assessment in higher education, the short-term research goal will be focused on two

major issues: retention and graduation rates, and outcomes and accountability. These two issues were selected as they represent the most concerned issues facing U.S. higher education and also they are closely related to ETS capabilities. Some of the other issues reviewed in this report (e.g., access to college) may align with the research taking place at the ETS Policy Evaluation and Research Center (PERC).

**What Are the Key Questions That Must Be Addressed for Achieving the Short-Term Goals?**

- To what extent can existing measures, including (but not limited to), SAT/ACT scores, high school courses taken, and grades in the those courses, reference letters, record of extracurricular activities, college performance, college experience, be combined to predict undergraduate degree completion? We will start with a literature review in year one to identify the factors that have an impact on degree completion [2011-2014].

- How does student motivation in taking low-stakes outcomes assessment affect their test performance? Does item format (e.g., multiple-choice vs. constructed-response) have an impact on student motivation? What are the strategies that institutions can use to monitor and improve student motivation? [2011-2012]

- What measures are appropriate to assess learning outcomes of community college students? Given the rapid development of the Voluntary Framework of Accountability, an outcomes assessment for community college students is needed. Results from this study will either inform the design of a new assessment or demonstrate the viability of current outcomes assessment for community college students. [2011-2013]

- Can a cross-sectional design adequately approximate a longitudinal design in measuring value-added in higher education? What statistical model should be used to compute value-added? This study is part of a pending study proposed to IES. [2011-2015]

- What is the validity evidence for using outcomes assessment for international institutions? [2012-2013]

- Is there any difference in outcomes (e.g., course grades, time to degree completion) between students taking online education and traditional education? [2012-2013]

- How should outcomes assessments be designed for adult learners? This question targets the new accountability initiative Transparency by Design, serving adult learners.

The above research questions focus on general issues in higher education. The following is a list of short-term questions that are program-specific.

- How should the Proficiency Profile test be revised to effectively measure critical thinking skills of college students? Is it possible to replace the current optional essay section with constructed-response critical thinking prompts? Can automated scoring methods (like *e-rater®)* be used to score these prompts?

- Is there any gender bias in student evaluation of teaching using the eSIR II program? This study was requested by the eSIR II program and funded for 2011

- How do students with an ELL background perform on the ETS Proficiency Profile?

- Can discipline-specific outcomes assessments (e.g., MFT) be used to evaluate program effectiveness?

**What Are the Long-Term Goals of the Research Agenda?**

The long-term research goal for the ETS higher education research agenda is to establish ETS as a thought leader in postsecondary education, both in the United States and worldwide. ETS will conduct cutting-edge research at all levels of higher education (e.g. community colleges, four-year colleges and universities, and graduate programs) and aim to expand access and improve learning for all student populations (e.g., traditional students, minorities, and international students) in higher education. Particularly, given its unparalleled capabilities in measurement and assessment, ETS will undertake pioneering research on developing and validating assessments that serve a range of purposes including diagnosis, formative and summative evaluation, learning outcomes, accreditation, and accountability. Through the research activities, ETS will forge partnerships with influential higher education organizations and significantly expand ETS brand recognitions worldwide.

**What Are the Key Questions that Must Be Addressed for Achieving the Long-Term Goals?**

- What are the strategies to increase high school graduation rates to expand the pool of college-bound students?

- What are the strategies to increase the percentage of college enrollment for Black and Hispanic students?
- What are the strategies to improve retention and graduation rates, especially at community colleges?
- What are the learning experiences of students in distance or blended education? What are the learning outcomes of these students? How do student learning experiences affect their learning outcomes?
- What knowledge and skills should college graduates master to become an effective contributor in a global workforce?
- How can outcomes assessment data be used to help Historically Black Colleges improve retention and graduation rates?
- How can assessments for accountability purposes also be used for curriculum and instruction improvement?
- In 1955, there were about 50,000 international students in U.S. higher education. The number increased to 690,923 in 2010 (Institute of International Education, 2010). It has been assumed that the increase of international students is changing the landscape of U.S. higher education. But there is little if any formal study showing what exactly these changes are. We ask the questions: How do international students help change the landscape of American graduate education? What can we do to retain the talents after the international students graduate?

## Resources Needed and Partnerships

### Research Expertise

At least two full-time research scientists are needed to carry out the short-term research studies in higher education. The research scientists should be familiar with relevant literature in higher education, on topics such as retention, completion rates, and outcomes issues, and have adequate training in research design and quantitative methods. It is also important that the research scientists collaborate with other scientists or psychometricians with statistical expertise

on projects that require advanced statistical skills. Adequate research support from research assistants is also critical to achieving the research goals.

### Partnerships

To execute the research plans, ETS needs to collaborate with higher-education institutions and organizations. While the institutions and organizations can benefit from ETS's capabilities in measurement and testing, ETS can benefit from their student and faculty resources. External funding will be a major funding source for these long-term research activities, given the depth and scope of these studies.

An effective information-sharing mechanism, both within and outside of ETS, is critical, to ensure a successful undertaking of both the short-term and long-term research studies. Specifically, researchers need to reach out to the higher-education community, understand the issues that institutional administrators and leaders are most concerned with, readily prioritize the research tasks, and make informed decisions. Researchers will use international, national, and regional conferences, forums, and advisory meetings as platforms to explore the future of higher education, and carefully examine the specific roles ETS can play in advancing higher education in the United States. Specifically, the higher education organizations ETS could reach out to include the Association for Institutional Research (AIR), the Association of Public and Land-grant Universities (APLU), the American Association of State Colleges and Universities (AASCU), the State Higher Education Executive Officers (SHEEO), American Association of Community Colleges (AACC), National Association of College Admissions Counseling (NACAC) and the American Council of Education (ACE). Inside ETS, Research and the Program Division need to work closely in identifying potential partnerships with members of various ETS advisory groups (e.g., the Higher Education Advisory Council, the National Community College Advisory Council, regional College and University Advisory Councils). Researchers will also work closely with the ETS external grants office in securing external funding.

## Conclusions

This research effort will contribute to improved understanding of the current situation and future needs of American higher education. For the United States to reaffirm its position as a world leader in postsecondary education, research on issues such as access, quality, and

accountability is urgently needed. Major setbacks, such as low retention and graduate rates, and unclear criteria for effective learning outcomes also need to be rectified in order to achieve President Obama's goal that by 2020, the United States should have the largest proportion of college graduates. ETS is in a unique position to tackle some of the issues, given its unparalleled research and testing capabilities and international recognition.

Addressing the research goals will also bring organizational benefits to ETS. It will help ETS establish leadership in the field of higher education, both domestically and internationally. Through undertaking the research activities, ETS will have expanded access to external resources. ETS will be able to engage important stakeholders in higher education when executing the research plans including researchers, university administrators, faculty, higher education organizations, policy makers, and business leaders. ETS will have the opportunities to secure partnerships with institutions and organizations that share the same goals. The dissemination of the research results will also help increase brand recognitions of ETS higher education services and products. The proposed research is firmly aligned with ETS mission of promoting learning and performance and supporting education for students of all backgrounds.

## References


ACT. (2008). 2008 Retention/Completion Summary Tables. Iowa City, IA: ACT.

Allen, I. E., & Seaman, J. (2008). *Staying the Course: Online Education in the United States, 2008.* Needham MA: Sloan consortium. Retrieved from http://www.sloan-c.org/publications/survey/pdf/staying_the_course.pdf

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology, 87,* 656–665.

Baumert, J., & Demmrich, A. (2001). Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16, 441–462.

Borden, V. M. H., and Young, J. W. (in press, 2008). Measurement validity and accountability for student learning. In V. M. H. Borden and G. R. Pike (Eds.), *Assessing and accounting for student learning: Finding a constructive path forward.* San Francisco: Jossey-Bass.

Bradley, A. P., & Blanco, C. D. (2010). *Promoting a culture of student success: How colleges and universities are improving degree completion*. Atlanta, GA: Southern Regional Education Board.

Braun, H., Kirsch, I., & Yamamoto, K. (2010). *An experimental study of the effects of monetary incentives on performance on the 12th grade NAEP reading assessment*. Princeton, NJ: ETS.

Center for Community College Student Engagement. (2009). *Benchmarking & Benchmarks: Effective Practice with Entering Students.* Austin, TX: The University of Texas at Austin, Community College Leadership Program.

Centra, J. A & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Educational Psychology, 70*(1), 17–33.

Chen, X., Wu, J., & Tasoff, S. (2010a). Academic Preparation for College in the High School Senior Class of 2003−04. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010169.

Chen, X., Wu, J., & Tasoff, S. (2010b). Postsecondary Expectations and Plans for the High School Senior Class of 2003−04. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010170.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.

Cole, J. S., & Osterlind, S. J. (2008). Investigating differences between low- and high-stakes test performance on a general education exam. The *Journal of General Education, 57*(2), 119–130.

Council for Aid to Education (CAE). (2007*). Collegiate Learning Assessment (CLA).* New York, NY: CAE.

Davidovitch, N., & Soen, D. (2006). Class attendance and students' evaluation of their college instructors. *College Student Journal, 40,* 691–703.

Davis, W. J., & Bauman, K. J. (2008). *School enrollment in the United States: 2006*. Washington, DC: U.S. Census Bureau.

DiPrete, T. A., & Forristal, J. D. (1994). Multilevel models: Method and substance. *Annual Review of Sociology, 20,* 331–357.

Dziuban, C., & Moskal, P. (2001). Evaluating distributed learning in metropolitan universities. *Metropolitan Universities, 12*(1), 41–49.

ETS (2010). *Market research of institutions that use outcomes assessment*. Princeton, NJ: ETS.

Ewell, P. T. (1994). *A policy guide for assessment: Making good use of the tasks in critical thinking*. Princeton, NJ: ETS.

Greene, T. G., Marti, C. N., & McClenney, K. (2008). The effort-outcome gap: Differences for African American and Hispanic community college students in student engagement and academic achievement. *The Journal of Higher Education, Vol. 79*(5), 513–539.

Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17–27.

Hendel, D. D. (1991). Evidence of convergent and discriminant validity in three measures of college outcomes. *Educational and Psychological Measurement, 51*(2), 351–358.

Hill, K., Hoffman, D., & Rex, T. R. (2005). *The value of higher education: Individual and societal benefits.* Tempe, AR: ASU School of Business.

Institute of International Education (2010). *Open Doors 2010 Fast Facts*. Retrieved from http://www.iie.org/en/Research-and-Publications/Open-Doors/Data.

Kelderman, E. (2010). *Duncan challenges black colleges to improve teacher training and graduation rates***.** Retrieved from http://chronicle.com/article/Duncan-Challenges-Black/65774/**.**

Kelly, A. P., Schneider, M. & Carey, K. (2010). *Rising to the Challenge: Hispanic College Graduation Rates as a National Priority.* Washington, DC: American Enterprise Institute.

Khan, B. H. (1997). Web-based instruction: What is it and why is it? In B. H. Khan (Ed.), *Web-based instruction* (pp. 5–18). Englewood Cliffs, NJ: Educational Technology Publications.

Kirsch, I., Braun, H. Yamamoto, K & Sum, A. (2007) *America's perfect storm*: *three forces changing our nation's future*. Princeton, NJ: ETS.

Klein, S. (1996). The costs and benefits of performance testing on the bar examination. *The Bar Examiner, 65*(3), 13–20.

Klein, S., Freedman, D., Shavelson, R. & Bolus, R. (2008). Assessing school effectiveness. *Evaluation Review, 32*(6), 511–525.

Klein, S., Kuh, G., Chun, M., Hamilton, L. & Shavelson, R. *(*2005). An approach to measuring cognitive outcomes across higher-education institutions. *Journal of Research on Higher Education, 46*(3), 251–276.

Kuh, G., & Ikenberry, S. (2009). *More than you think, less than we need: Learning outcomes assessment in American higher education*. Champaign, IL: National Institute for Learning Outcomes Assessment (NILOA).

Landrum, P. E., & Braitman, K. A. (2008). The effect of decreasing response options on students' evaluation of instruction. *College Teaching, Vol. 56*(4), 215–217.

Lee, Y., Driscoll, M. P., & Nelson, D. W. (2004). The past, present, and future of research in distance education: Results of a content analysis. *The American Journal of Distance Education. 18*(4), 225–241.

Liu, O. L. (2008). *Measuring learning outcomes in higher education using the Measure of Academic Proficiency and Progress (MAPP™).* (ETS Research Report No. RR-08-047). Princeton, NJ: ETS.

Liu, O. L. (2009a). Measuring value-added in higher education: Conditions and caveats. Results from using the Measure of Academic Proficiency and Progress (MAPP™). *Assessment and Evaluation in Higher Education, 34*(6), 1–14.

Liu, O. L. (2009b). *R&D Connections: Measuring learning outcomes in higher education* (Report No. RDC-10). Princeton, NJ: ETS.

Liu, O. L. (in press). Value-added assessment in higher education: A comparison of two methods. *Higher Education*.

Marr, D. (1995). *Validity of the academic profile*. Princeton, NJ: ETS.

Mullin, C. M. (2010, June). *Rebalancing the mission: The community college completion challenge* (Policy Brief 2010-02PBL). Washington, DC: American Association of Community Colleges.

Museus, S. D., & Quaye, S. J. (2009). Toward an intercultural perspective of racial and ethnic minority college student persistence**.** *The Review of Higher Education, 33*(1), 67–94.

National Center for Education Statistics (NCES) (2005). *The condition of education 2005*. Washington, D.C.: U.S. Government Printing Office. .

National Survey of Student Engagement (2008). *Promoting engagement for all students: The imperative to look within*. Bloomington, IN: Indiana University.

O'Neil, H. F., Abedi, J., Miyoshi, J., & Mastergeorge, A. (2005). Monetary incentives for low-stakes tests. *Educational Assessment, 10*(3), 185–208.

O'Neil, H. F., Jr., Sugrue, B., & Baker, E. L. (1995/1996). Effects of motivational interventions on the National Assessment of Educational Progress mathematics performance. *Educational Assessment*, 3, 135–157.

Oseguera, L. (2005-2006). Four and six-year baccalaureate degree completion by institutional characteristics and racial/ethnic groups. *The Journal of College Student Retention, 7*(1–2), 19–59.

Oseguera, L. & Rhee, B. S. (2009). The influence of institutional retention climates on student persistence to degree completion: a multilevel approach. *Research in Higher Education, 50*(6), 546–569.

Parry, M. (2010). *Wal-Mart employees get new college program—Online.* Retrieved from http://chronicle.com/blogPost/Wal-Mart-Employees-Get-New/24504/

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Planty, M., Hussar, W., Snyder, T., Provasnik, S., Kena, G., Dinkes, R., KewalRamani, A., & Kemp, J. (2008). *The condition of education 2008* (NCES 2008–031). Washington, DC: U. S. Printing Office.

Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRamani, A., Kemp, J., Bianco, K., & Dinkes, R. (2009). *The condition of education 2009* (NCES 2009–081).

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata.* (2nd ed.). College Station, TX: Stata Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rovai, A. P., & Jordan, H. M. (2004). Blended learning and sense of community: A comparative analysis with traditional and fully online graduate courses. *International Review of Research in Open and Distance Learning, 5*(2), 2–13.

Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Education and Behavioral Statistics, 24* (4), 323–355.

Snyder, T. D., Dillow, S. A., & Hoffman, C. M. (2009). *Digest of Education Statistics 2008* (NCES 2009-020). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

Steedle, J. (2010, May). Improving the reliability and interpretability of value-added scores for post-secondary institutional assessment programs. Paper presented at the annual conference of the American Educational Research Association, Denver, CO.

Stone, C. (1995). *A course in probability and statistics*. Belmont, CA: Duxbury Press.

Sundre, D. L. (1997, April). *Differential examinee motivation and validity: A dangerous combination.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Sundre, D. L. (1999, April). *Does examinee motivation moderate the relationship between test consequences and test performance*? Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Sundre, D. L., & Wise, S. L. (2003, April). Motivation filtering: An exploration of the impact of low examinee motivation on the psychometric quality of tests. Paper present at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

U.S. Census Bureau. (2008). *2008 American community survey 1 -year estimates*. Retrieved from: http://factfinder.census.gov/servlet/DTTable?_bm=y&-context=dt&-ds_name=ACS_2008_1YR_G00_&-CONTEXT=dt&-mt_name=ACS_2008_1YR_G2000_B02001&-tree_id=306&-redoLog=false&-all_geo_types=N&-currentselections=ACS_2006_EST_G2000_B02001&-geo_id=01000US&-search_results=01000US&-format=&-_lang=en

U.S. Department of Education. (2006). *A test of leadership: charting the future of U.S. higher education.* Washington, D.C.: Author.

U.S. Department of Education, National Center for Education Statistics. (1996). *Beginning postsecondary students longitudinal study* (BPS:96/01). Retrieved from http://nces.ed.gov/search/?as_sitesearch=nces.ed.gov/pubsearch/pubsinfo.asp&output=xml_no_dtd&site=nces&client=nces&q=96/01&ie=UTF-8&ip=192.168.132.61&access=p&sort=date:D:L:d1&entqr=3&entsp=a

Voluntary System of Accountability. (2008). *Background on learning outcomes measures.* Retrieved from www.voluntarysystem.org/ index.cfm?page=about_cp

Wendler, C., Bridgeman, B., Cline, F., Millett, C., Rock, J., Bell, N., & McAllister, P. (2010). *The path forward: The future of graduate education in the United States*. Princeton, NJ: ETS.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*(1), 1–17.

Wise, S.L., & DeMars, C. E. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurement, 43*(1), 19–38.

Wise, S. L., & DeMars, C. E. (2010). Examinee noneffort and the validity of program assessment results. *Educational Assessment, 15,* 27–41.

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education, 18*(2), 163–183.

Wise, V.L. (2004). *The effects of the promise of test feedback on examinee performance and motivation under low-stakes testing conditions.* Unpublished doctoral dissertation, University of Nebraska-Lincoln.

Wise, V. L., Wise, S. L., & Bhola, D. S. (2006). The generalizability of motivation filtering in improving test score validity. *Educational Assessment, 11*(1), 65–83.

Wolf, L. F., & Smith, J. K. (1995). The consequence of consequence: Motivation, anxiety, and test performance. *Applied Measurement in Education, 8*, 227–242.

Zhao, Y., Lei, J., Lai, C., Yan, B., & Tan, H. S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record, 107*, 1836–1884.

## Notes

[1] Data were not available for other ethnicity groups.

[2] Since there is little difference between Academic Profile and Proficiency Profile, Academic Profile will be referred to as Proficiency Profile in the following sections of the report.