

the twentieth century to de-
to measure undergraduate
d universities that granted
ke the Collegiate Learning
005) and at the end of their
o consented to participate
multiple points in their col-
their social and educational
collected course transcript
ools and colleges that the
based on longitudinal data
e of campuses. Colleges in
ctivity, and missions. The
earch institutions, as well
universities (HBCUs) and
s are dispersed nationally
this multifaceted data as
aset.

ur reliance on participat-
n sampling and retention
to which students in our
rom these institutions as
his book's methodologi-
data from the Integrated
eginning Postsecondary
es, students in the DCL
ditional-age undergrad-
and universities they at-
ide. The DCL students'
their English-language
well with national statis-
d college-educated par-
mple of traditional-age
s in both the DCL and
Moreover, the four-year
proportion of white stu-
those of four-year insti-
percentiles of entering
tutions nationwide are
participation required

in our study, however, our sample did have fewer men, as well as fewer students of lower scholastic ability as measured by standardized tests—for example, students' combined scores at the 25th percentile of the SAT were lower in our sample than at DCL institutions or four-year institutions nationwide. Consequently, we believe that any biases introduced into our analysis by the sampling procedures used are likely to be in the direction of leading us toward overestimating students' *positive* educational experiences and institutional success.

The Collegiate Learning Assessment

The Collegiate Learning Assessment (CLA) consists of three open-ended, as opposed to multiple-choice, assessment components: a performance task and two analytical writing tasks (i.e., to make an argument and to break an argument). According to its developers, the CLA was designed to assess "core outcomes espoused by all of higher education—critical thinking, analytical reasoning, problem solving and writing."¹⁰ These *general skills* are "the broad competencies that are mentioned in college and university mission statements."¹¹ Rather than testing for *specific content knowledge* gained in particular courses or majors, the intent was to assess "the collective and cumulative result of what takes place or does not take place over the four to six years of undergraduate education in and out of the classroom."¹² The developers of the CLA argue that it assesses abilities distinct from those measured in general education tests such as the Scholastic Aptitude Test (SAT) and the American College Testing (ACT) program. "Consequently, an SAT prep course would not help a student on the CLA and instruction aimed at improving CLA scores is unlikely to have much impact on SAT or ACT scores."¹³

While the CLA as a whole is considered by some as state-of-the-art, the performance task component is its most well-developed and sophisticated part. Our analysis, which follows in this book, will focus on that component. The performance task allows students ninety minutes to respond to a writing prompt that is associated with a set of background documents. The testing materials, including the documents, are accessed through a computer. The Council for Aid to Education has published several examples of representative performance tasks that are worth describing here in detail.

The "DynaTech" performance task asks students to generate a memo advising an employer about the desirability of purchasing a type of airplane that has recently crashed. Students are informed: "You are the assistant to Pat Williams, the president of DynaTech, a company that makes precision

electronic instruments and navigational equipment. Sally Evans, a member of DynaTech's sales force, recommended that DynaTech buy a small private plane (a SwiftAir 235) that she and other members of the sales force could use to visit customers. Pat was about to approve the purchase when there was an accident involving a SwiftAir 235." Students are provided with the following set of documents for this activity: newspaper articles about the accident, a federal accident report on in-flight breakups in single engine planes, Pat Williams's e-mail to her assistant and Sally Evans's e-mail to Pat Williams, charts on SwiftAir's performance characteristics, an article from *Amateur Pilot* magazine comparing SwiftAir 235 to similar planes, and pictures and descriptions of SwiftAir models 180 and 235. Students are then instructed to "prepare a memo that addresses several questions, including what data support or refute the claim that the type of wing on the SwiftAir 235 leads to more in-flight breakups, what other factors might have contributed to the accident and should be taken into account, and your overall recommendation about whether or not DynaTech should purchase the plane."¹⁰

A second performance task that the Council for Aid to Education has circulated is related to crime reduction. The test instructs students that "Jamie Eager is a candidate who is opposing Pat Stone for reelection. Eager critiques the mayor's solution to reducing crime by increasing the number of police officers. Eager proposes the city support a drug education program for addicts because, according to Eager, addicts are the major source of the city's crime problem." Students again are provided with a set of documents including newspaper articles, crime and drug statistics, research briefs, and internal administrative memos. The CLA requires that students should specifically address the following: "Mayor Pat Stone asks you to do two things: (1) evaluate the validity of Eager's proposal and (2) assess the validity of Eager's criticism of the mayor's plan to increase the number of officers."¹¹

The Council for Aid to Education has also published a detailed scoring rubric on the criteria that it defines as critical thinking, analytical reasoning, and problem solving—including how well the student assesses the quality and relevance of evidence, analyzes and synthesizes data and information, draws conclusions from his or her analysis, and considers alternative perspectives. In addition, the scoring rubric with respect to written communication requires that the presentation is clear and concise, the structure of the argument is well-developed and effective, the work is persuasive, the written mechanics are proper and correct, and reader interest is maintained.¹²

The design of low "a criterion the whole is greater than the parts" require an integrative and measured approach is to "saturate" the test to observe her performance. Thus, she attempts to observe her performance in real-world life situations." G

holistic, real-world to the test" will be analytically solve

The CLA has been Future of Higher Education based assessment comprehensive national different campuses the Carnegie Corporation "rose from the field" and psychometric higher education, that the "C.L.A. is based on standardized tests."

Nevertheless, the falls into several broad categories: increased encroachment and higher education of student learning and deep amongst e Center for Public Policy has deflected the idea things we want und able." ¹³ Resistance has which are not responsible instrument . . . will be and the country," the Colleges and Universities meaningless outcome

These critics of increased such efforts are also u

The design of the prompts and the criteria applied for evaluation follow "a criterion sampling approach to measurement" that "assumes that the whole is greater than the sum of its parts and that complex tasks require an integration of abilities that cannot be captured when divided into and measured as individual components."²⁷ The philosophy behind the approach is to "sample tasks from the domain in which that person is to act, observe her performance, and infer competence and learning."²⁸ The CLA thus attempts to identify "real-world tasks that are holistic and drawn from life situations." Given that the performance tasks involve solving "complex, holistic, real-world problems," college institutions that attempt to "teach to the test" will be schools that teach students "to think critically, reason analytically, solve problems, and communicate clearly."²⁹

The CLA has been lauded by many. For example, the Commission on the Future of Higher Education noted that it "promotes a culture of evidence-based assessment in higher education" and is "among the most comprehensive national efforts to measure how much students actually learn at different campuses."³⁰ The former program director of higher education for the Carnegie Corporation of New York, Daniel Fallon, noted that the CLA "rose from the field" as "the best creative thinking of the academic research and psychometric community" focused on measuring student learning in higher education.³¹ Even testing skeptics, such as James Traub, have noted that the "C.L.A. is light years ahead of the fill-in-the-blanks format of most standardized tests."³²

Nevertheless, the CLA also has its fair share of critics. The criticism falls into several broad categories. First, there are those who resist any increased encroachment of testing and assessment in education in general and higher education in particular. Resistance to standardized assessment of student learning in U.S. higher education has been historically broad and deep amongst educators. As Patrick Callan, president of the National Center for Public Policy and Higher Education, notes: "Higher education has deflected the idea for the past quarter-century by arguing the kinds of things we want undergraduate education to teach are not really measurable."³³ Resistance has been particularly pronounced at private colleges, which are not responsive to public officials. "Trying to create an uber-instrument . . . will be a grave disservice to the individuals, institutions, and the country," the president of the National Association of Independent Colleges and Universities, David Warren, has commented. "We will get a meaningless outcome at a great cost."³⁴

These critics of increased standardized learning assessment argue that such efforts are also unnecessary given the successes of a U.S. higher edu-

cation system that already inherently ensures accountability through market forces. As Princeton professor and former president of the American Council of Learned Societies, Stanley Katz, has noted: "the public is quite satisfied with what higher education is doing on the whole. This is a market system, and the customers are buying. We have by a considerable measure the finest system of higher education in the world. And if that's the case, this is an 'ain't broke, don't fix it' situation."⁵⁰ While we share Katz's sentimental attachment to a U.S. higher education system that has generously provided us with both training and employment, we are skeptical of most of the assumptions inherent in this argument. The "market" system for higher education in the U.S. is characterized by a limited number of selective institutions that share many features in common, that control access to scarce goods (i.e., prestigious credentials) and that are heavily subsidized by public sources of support such as college grant provisions, loan guarantees, tax exemptions, and research grants.

In recent decades, the U.S. higher education system has fallen behind many other countries in terms of the percentage of individuals it graduates.⁵¹ Moreover, whether college students are more effectively educated in the U.S. than abroad is today an open empirical question, but will perhaps not remain so for much longer. The Organisation of Economic Cooperation and Development (OECD) is currently launching a feasibility study for the international Assessment of Higher Education Learning Outcomes (AHELO) that will parallel its earlier efforts that have successfully assessed academic performance of fifteen-year-olds from a comparative international perspective since 2000 with the Programme for International Student Assessment (PISA). The OECD efforts are designed to develop a "direct assessment of learning outcomes in higher education" that "could provide member governments with a powerful instrument to judge the effectiveness and international competitiveness of their higher education institutions, systems and policies in the light of other countries' performance, in ways that better reflect the multiple aims and contributions of tertiary education to society."⁵² It is worth noting here that AHELO decided to embrace and adapt the CLA "to an international context with a view to provide a proof of concept" for its assessment of generic skills that "can be measured across diverse institutions, languages and cultures." In particular, students in multiple countries in 2016 "will complete an online assessment, using their critical skills along with data provided for each task. The questions are not specialized so that they can be answered by most undergraduates, whatever their field of study."⁵³

A second line of criticism is not necessarily opposed to testing itself, but

questions the validity of generalizing from the specific knowledge of the sciences, mathematics, philosophy, and economics to the general. Hoffman Breyer at the University of California, Berkeley, has argued that "a standardized test, which measures only a narrow range of skills and knowledge, could not accurately measure what each student had worked to learn." The American Council on Education has asserted that "the variability of student experiences across majors and institutions makes an overall measure of performance for all students 'will necessarily be flawed.'" The criticisms are unclear, however, as to whether curricular composition itself is a problem for student learning or why or how it would affect when modeling results.

Third, skeptics of the CLA question the instrumental validity of the CLA. This question has now been addressed by the Organisation for Economic Co-operation and Development Fund for the Improvement of Living and Working Conditions, which brought together researchers from the Educational Testing Service (ETS) and the ACT to evaluate the Educational Testing Service (ACT) program. It examined the validity of the ACT's Collegiate Assessment of Academic Proficiency (CAAP) and ETS's Measure of Academic Proficiency (MAP) by administering all three tests in a number of countries participating. While CAAP score reliability with the ACT at the school level (correlations were high) and at the individual level, correlations were high for multiple choice tests of critical thinking skills. However, tests of different constructs should not be used to assess students as individuals (e.g., in smaller samples) or for group-level differences.

Fourth, some higher-education critics question the CLA itself, but the modeling of institutional and individual college-level effects with this

ability through mar-
ent of the American
"the public is quite
ole. This is a market
considerable mea-
ld. And if that's the
hile we share Katz's
tem that has gener-
we are skeptical of
ne "market" system
limited number of
mon, that control
nd that are heavily
e grant provisions,

has fallen behind
dividuals it gradu-
fectively educated
stion, but will per-
of Economic Co-
chng a feasibility
ion Learning Out-
have successfully
a comparative in-
for International
gned to develop a
ation" that "could
ment to judge the
higher education
countries' perfor-
l contributions of
t AHELO decided
ext with a view to
skills that "can be
es." In particular,
an online assess-
or each task. The
d by most under-

testing itself, but

questions the validity of general, broad-based assessments that do not focus on the specific knowledge taught in particular courses and majors (e.g., life sciences, mathematics, physical sciences, and social sciences). Catherine Hoffman Breyer at the University of Washington, for example, has argued that "a standardized test, such as the CLA, with its focus on generic skills and knowledge, could not detect the specialized information and skills each student had worked to master."⁸⁴ In a similar fashion, Steve Chatman at the University of California at Berkeley's Center for Studies in Higher Education has asserted that "because of the differences in undergraduate experiences across majors within an institution, any attempt to capture an overall measure of performance across all of a college or university's students 'will necessarily be biased' by the makeup of its programs."⁸⁵ These criticisms are unclear, however, on why one should not consider a college's curricular composition itself to be an institutional policy associated with student learning or why one could not easily control for these differences when modeling results.

Third, skeptics of the CLA in the past have raised questions about the instrumental validity of the indicator. Some of these concerns, however, have now been addressed by a recent test validity study organized by the Fund for the Improvement of Postsecondary Education (FIPSE). This study brought together researchers from the Council for Aid to Education (CAE), the Educational Testing Service (ETS), and the American College Testing (ACT) program. It examined the instrumental construct validity of the CLA, the ACT's Collegiate Assessment of Academic Proficiency (CAAP) and the ETS's Measure of Academic Proficiency and Progress (MAPP) by administering all three tests in thirteen schools with more than 1,100 students participating. While CAAP and MAPP rely on a multiple choice format, score reliability with the CLA was high when considered at the aggregate school level (correlations of 0.75 to 0.84). In addition, at the individual level, correlations were higher across CLA open-ended and CAAP/MAPP multiple choice tests of critical thinking ($r = 0.53$) than CLA-CAAP/MAPP tests of different constructs ($r = 0.45$). While the results indicate that these tests should not be used as a basis to make institutional decisions about students as individuals (e.g., promotion or course placement), when aggregated in larger samples they can provide reliable estimates of institutional or group-level differences in performance on these tasks.⁸⁶

Fourth, some higher-education practitioners have questioned not the CLA itself, but the modeling approach that the Council for Aid to Education and individual colleges and universities have used to identify institutional effects with this assessment instrument. CLA has generally been

used in a value-added framework, which entails comparing test scores of enrolled freshmen and seniors at an institution in a given year, after controlling for student performance on a prior test such as the SAT or ACT. These comparisons have not typically tracked specific students through college, nor have they accounted for other non-school factors that might be associated with differential rates of learning. Higher-education practitioners, such as Chancellor Howard Cohen of Purdue University Calumet, has questioned whether one “can measure the ‘value added’ in college generally, when so much of the experience of students is beyond the control of colleges.”⁸⁷ If one longitudinally tracked students over time, however, and adequately accounted for a full set of non-school factors—as we will do in this project—even CLA critics such as Wheaton College Dean Gary N. Larson concede that the measurement approach would approximate a “gold standard” for assessing student outcomes.⁸⁸

Although there are significant methodological challenges to our project (including issues of sampling, attrition, and selection that are discussed at length in a methodological appendix), the study generates significant new knowledge to guide future research, policy, and practice. While well short of an experimental research “gold standard,” descriptive findings based on tracking many students enrolled in diverse institutions, with careful longitudinal measurement of a wide range of factors and outcomes over time, yields quite illuminating results on the nature and character of collegiate experiences and variation in student learning that can significantly increase our understanding of the phenomenon.

Other Studies of Learning and Student Trajectories through College

In spite of the increasing attention of policy makers on measuring student learning in higher education, and an extensive tradition of research on academic performance in elementary and secondary education, efforts to directly measure development of general cognitive skills in college have been limited. Over the past decade the most widely used assessment of student learning and personal development in higher education has been the National Survey of Student Engagement (NSSE), which presents students with a questionnaire in multiple-choice format that gauges students’ self-assessment of their learning during college. Since the inception of the NSSE in 2000, more than 1,300 colleges and universities in the United States and Canada have used it to survey students about their learning.

It is unclear, however, whether students can accurately self-report an

assessment of the degree to which they have learned. As young adults, are they aware of their own learning? Can they not identify or define learning? Do they not know whether they have obtained the skills they claim to have? It is well known to be susceptible to bias. For example, as the editor noted, “some 80 percent of students reported that they had learned more and “more than 90 percent of students reported that they had learned more than their average colleagues.”⁸⁹ Others have used NSSE results to assess student learning and self-reporting. A study that systematically demonstrated that self-reported learning was consistently associated with growth in learning outcomes.

Instead of relying on student self-reports, a number of large-scale national projects have been developed by directly relying on different methods. The ACT program developed by the ACT program, which includes the ACT Learning (NSSL) followed approximately 100 institutions through their first year of college from 1992 to 1999. While this project is not a longitudinal study, it provides insights about the relationship between student learning and their improvement in general cognitive skills. In 2006, Charles Bluhm, director of the Center for Arts at Wabash College launched the Center for Arts Education. Starting with a study of liberal arts education, it has been expanded to include a diverse range of liberal arts colleges, regional colleges, and community colleges. Students are tested at their entry into high school and at the end of their senior year. The study measures outcomes, from academic motivation to leadership, moral reasoning, and critical thinking (evaluation of evidence). Although the multiple-choice format has been criticized for its reductionist nature, the study is among the few large-scale studies that have well as nonacademic experiences. The study also provides information on students’ dem

comparing test scores of
a given year, after con-
such as the SAT or ACT.
specific students through
school factors that might
higher-education practi-
due University Calumet,
e added' in college gen-
is beyond the control of
ver time, however, and
ctors—as we will do in
lege Dean Gary N. Lar-
ld approximate a “gold

allenges to our project
on that are discussed at
erates significant new
ctice. While well short
riptive findings based
stitutions, with careful
ers and outcomes over
e and character of col-
g that can significantly

ories

on measuring student
dition of research on
education, efforts to
skills in college have
y used assessment of
r education has been
which presents stu-
that gauges students'
e the inception of the
sities in the United
out their learning.
rately self-report an

assessment of the degree to which they have actually learned general skills. As young adults, are they aware of what they do not know? If students cannot identify or define learning and critical thinking skills, how will they know whether they have obtained them? Self-reported assessments are also well known to be susceptible to inflated perceptions of one's own performance. For example, as the economists Robert Frank and Philip Cook have noted, “some 80 percent of us think we are better-than-average drivers” and “more than 90 percent of workers consider themselves more productive than their average colleague.”⁸⁹ In addition, while George Kuh and others have used NSSE results to identify associations between self-reported student learning and self-reported college engagement, it has not yet been systematically demonstrated that all forms of college engagement are consistently associated with growth on objective measures of learning.

Instead of relying on students' self-reports of their cognitive gains, two large-scale national projects have aimed to measure student learning directly by relying on different modules of the CAAP, an assessment tool developed by the ACT program to measure general college skills including critical thinking, reading, and writing. The National Study of Student Learning (NSSL) followed approximately 4,000 students at twenty-three institutions through their first three years in college, beginning in the fall of 1992. While this project is no longer ongoing, it has provided important insights about the relationship between students' college experiences and their improvement in general skills such as reading, writing, and critical thinking. In 2006, Charles Blaich at the Center of Inquiry in the Liberal Arts at Wabash College launched the Wabash National Study of Liberal Arts Education. Starting with nineteen institutions, the study has since been expanded to include a diverse set of forty-nine institutions including liberal arts colleges, regional universities, research universities, and community colleges. Students participating in the study are surveyed and tested at their entry into higher education, at the end of their first year, and at the end of their senior year. This study assesses a range of college outcomes, from academic motivation and attitudes toward reading and writing to leadership, moral reasoning, and attitudes about diversity, as well as critical thinking (evaluated using the CAAP critical thinking test). Although the multiple-choice framework to assessing college learning can be criticized for its reductionist character, the Wabash and earlier NSSL studies are among the few large-scale efforts to assess how academic as well as nonacademic experiences are associated with student learning, and how those experiences are shaped by student backgrounds. By collecting information on students' demographic characteristics, pre-college attri-

butes, and college experiences, as well as by conducting in-depth interviews with a subsample of students, the Wabash study in particular promises to provide crucial insights into factors shaping student development over four years of college.²⁰

In addition to these studies, which directly measure students' experiences and performance during college, some studies have used standardized test scores, such as SAT and ACT pre-college measures and GRE post-college measures, to approximate a repeated indicator longitudinal assessment design.²¹ Moreover, recent reports from the *Measuring Up* initiative have used professional exams and licensures as a proxy for learning. While these endeavors, which aim to approximate but not directly measure students' progress through college, present important steps in the measurement of student outcomes, they are limited to students who take specific tests, and thus miss a large proportion of students who do not pursue specific educational or occupational paths affected by graduate school or licensure exams immediately after college.

Although scant attention has been dedicated to measuring student learning with objective performance assessment across institutions and over time, several large projects have recently focused on tracking students through college and into the labor force. While ignoring the measurement and modeling of student learning, these endeavors provide useful models for thinking about student experiences and outcomes in higher education. William Bowen and Derek Bok in *The Shape of the River* examined outcomes of minority students admitted to selective colleges under race-sensitive policies relative to the outcomes of their white peers in the 1979 and 1989 entering freshmen cohorts. Non-white students at twenty-eight academically selective and predominantly private colleges "have, overall, performed very well" on a wide range of indications—including graduation rate, fields of study, advanced degree attainment, earnings, and civic engagement.²² The one major exception to this pattern was observed in student academic outcomes measured by college grade point averages. Specifically, Bowen and Bok demonstrated that "black students with the same SAT scores as whites tend to earn lower grades."²³ James Shulman and William Bowen found in subsequent work that while college athletes graduate at relatively high rates from these selective college settings, their grades in college are lower than expected after controlling for prior preparation, and have been deteriorating over time.²⁴

In more recent work, Douglas Massey and his colleagues have tracked a large number of students entering college in the fall of 1999 at a similar set of twenty-eight selective colleges and universities "essentially following the

cohort of freshmen became sophomore. In a series of articles, Massey et al. (2001) report on their findings in particular, similar to Bowen and Bok (1993) that students were high-achieving and academically prepared to the extent to which they had taken measures while at college. They also reported that students who were engaged in fraternities or sororities

These endeavors provide important experiences during the college years, but student learning outcomes and other outcomes, such as college grades, the extent of learning in higher education, student learning value-added, scaled grading systems, and alternative measures across institutions and across teachers within institutions, are grade definition based. Although these data are worth collecting, they do not on their own they provide a complete picture of student learning.

Moreover, past research on student outcomes have often found that non-white students tend to garner more success than the privilege of attending selective institutions differ from the characteristics of individual characteristics of institutions participating in the study (used by Massey et al., 2001) that parents who had graduate degrees that "by any criteria, they are a sample."²⁵ Similarly, studies by Bowen and Bok (1993) on college grade average and, not surpris-

cohort of freshmen entering Bowen and Bok's sample of schools as they became sophomores, juniors, and ultimately for most, graduating seniors.²⁰ In a series of articles and books, Massey and his colleagues focused attention in particular on racial differences in student outcomes. In results similar to Bowen and Bok's earlier work, the lower grades of African-American students were highlighted (net of extensive controls for social background and academic preparation). Massey and his colleagues also identified the extent to which African-American students faced greater economic pressures while at college, and the extent to which students regardless of race who were engaged in many campus activities (other than membership in a fraternity or sorority) earned higher grades.²¹

These endeavors provide invaluable information about students' experiences during their college years. However, they have failed to measure student learning or link student experiences to growth in learning. Among other outcomes, Bowen and Bok as well as Massey et al. report analyses of college grades, the traditional and long-relied-upon method of measuring learning in higher education. Grades are an effective way of measuring student learning within a particular class, since most institutions have a scaled grading system already in place. They are an unreliable comparative measure across classes or schools, however, since inconsistencies exist across teachers within schools and there are discrepancies in scale and grade definition between schools and over time as grade inflation has occurred. Although grades serve a valuable purpose within classrooms and are worth collecting as a component of a larger evaluation strategy, on their own they provide only a very limited and inadequate assessment of student learning.

Moreover, past endeavors examining college students' experiences and outcomes have often focused on selective colleges and the experience of non-white students attending these schools. While selective institutions tend to garner much scholarly attention, most students do not have the privilege of attending such schools. Students attending selective institutions differ from those attending the rest of higher education on a number of individual characteristics as well as outcomes. The median SAT score for institutions participating in the National Longitudinal Survey of Freshmen (used by Massey et al.) was 1,243 and the majority of those students had parents who had graduated from college, leading the authors to conclude that "by any criteria, the twenty-eight institutions constitute an elite sample."²² Similarly, students in the College and Beyond (C & B) dataset studied by Bowen and Bok were more academically prepared than the national average and, not surprisingly, had much higher graduation rates: 85 percent

of C&B students graduated from the *same institution* within six years, compared to the national average of just over 50 percent.¹⁸ Thus, knowing the patterns and consequences of specific activities at elite institutions does not necessarily extend to the majority of students who are attending non-selective colleges and universities. Questions about the growth in student learning over time and the patterns and consequences of different collegiate experiences on average U.S. campuses still remain to be answered.

Outline of our Presentation

In this book we will highlight four core “important lessons” from our research. First, in terms of undergraduate learning, four-year colleges and universities and students attending them are too often “academically adrift.” While U.S. higher education is expected to accomplish many tasks, we draw on students’ reports of their collegiate experiences to demonstrate that undergraduate learning is rarely adequately prioritized. Second, gains in student performance are disturbingly low; a pattern of limited learning is prevalent on contemporary college campuses. Third, individual learning in higher education is characterized by persistent and/or growing inequality. Fourth, while the overall level of learning is low, there is notable variation both within and across institutions that is associated with measurable differences in students’ educational experiences.

In chapter 2 we continue to describe the 2,322 students in our study as they begin their college careers. We focus in particular on the extent to which they are improving their skills in critical thinking, complex reasoning, and writing as measured by the CLA during the first two years in college. Moreover, while inequalities in access persist, higher-education institutions today enroll an increasingly diverse set of students from a variety of backgrounds. We thus examine whether CLA performance at entry into higher education as well as gains over time vary across students from different social backgrounds, focusing in particular on different racial/ethnic groups and students from more or less educated families. This chapter reveals that American higher education is characterized by limited or no learning for a large proportion of students, and persistent or growing inequalities over time.

Chapter 3 examines how students navigate and experience contemporary college cultures. How distinctive are these cultures? Do students’ academic attitudes, behaviors, and values simply reflect their divergent social backgrounds and academic abilities? Or do colleges differ in the extent to which they successfully promote student academic orientations and prac-

tices? We find that academic programs are students rarely seem academically adrift in that colleges vary in entered student behavior.

How are students’ critical thinking, complex reasoning, and social integration to improve student retention of rigorous courses spent studying is highly employment and extracurricular learning, as well as how associated with improvement of learning are low, we contexts that are associated with complex reasoning, and writing.

In our concluding chapter we highlight the extent of higher education. Specifically, we call attention to the fact that academically adrift on contemporary improved measurement associated with student learning in the economy, the consequences of economic competitiveness upon educating a generation deeply, and commitment of such competencies to be successfully understood by college graduates.

stitution within six years, com-
percent." Thus, knowing the
ities at elite institutions does
students who are attending non-
s about the growth in student
sequences of different colle-
still remain to be answered.

important lessons" from our re-
ning, four-year colleges and
are too often "academically
ed to accomplish many tasks,
e experiences to demonstrate
ly prioritized. Second, gains
a pattern of limited learning
s. Third, individual learning
ent and/or growing inequal-
s low, there is notable varia-
associated with measurable
es.

2,322 students in our study
in particular on the extent
ical thinking, complex rea-
uring the first two years in
s persist, higher-education
e set of students from a va-
CLA performance at entry
e vary across students from
lar on different racial/eth-
ated families. This chapter
acterized by limited or no
persistent or growing in-

nd experience contempo-
cultures? Do students' aca-
fect their divergent social
ges differ in the extent to
nic orientations and prac-

ities? We find disturbing evidence that many contemporary college aca-
ademic programs are not particularly rigorous or demanding. Moreover,
students rarely seem to focus on academic pursuits; many appear to be
academically adrift in today's colleges and universities. We show, however,
that colleges vary in the extent to which they support academically ori-
ented student behaviors.

How are students' experiences in college related to their development of
critical thinking, complex reasoning, and writing skills as measured by the
CLA? We address this question in chapter 4, by exploring how academic
and social integration—with the latter being promoted by many colleges
to improve student retention—are related to student learning. The impor-
tance of rigorous coursework requirements, faculty expectations, and time
spent studying is highlighted. In addition, we discuss whether student em-
ployment and extracurricular activities can become a distraction to student
learning, as well as how various college majors and types of coursework are
associated with improvement in CLA performance. While overall levels
of learning are low, we identify specific experiences and higher-education
contexts that are associated with improvement in critical thinking, com-
plex reasoning, and writing skills during the first two years of college.

In our concluding chapter, we argue that the patterns identified in our
study highlight the extent to which institutional reform is required in U.S.
higher education. Specifically, while others have applied the metaphor
of a river to the journey through college of today's students, our findings
call attention to the fact that many undergraduate students are academi-
cally adrift on contemporary campuses. Educational reform requires im-
proved measurement and understanding of the processes and factors as-
sociated with student learning. In an increasingly globalized competitive
economy, the consequences of policy inattention are profound. Regardless
of economic competitiveness, the future of a democratic society depends
upon educating a generation of young adults who can think critically, rea-
son deeply, and communicate effectively. Only with the individual mas-
tery of such competencies can today's complex and competitive world
be successfully understood and navigated by the next generation of col-
lege graduates.