Quality in the Undergraduate Experience: What Is It? How Is It Measured? Who Decides? Summary of a Workshop

DETAILS

98 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-44301-2 | DOI: 10.17226/23514

AUTHORS

Karin Matchett, Maria Lund Dahlberg, and Thomas Rudin, Rapporteurs; Board on Higher Education and Workforce; Policy and Global Affairs; National Academies of Sciences, Engineering, and Medicine

Quality in the Undergraduate Experience

What Is It? How Is It Measured? Who Decides?

SUMMARY OF A WORKSHOP

Karin Matchett, Maria Lund Dahlberg, and Thomas Rudin, Rapporteurs

Board on Higher Education and Workforce

Policy and Global Affairs

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by Grant No. OPVT/Lumina Foundation 9408 from The Lumina Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13:978-0-309-44301-2
International Standard Book Number-10:0-309-44301-6
Digital Object Identifier: 10.17226/23514

Additional copies of this report are available for sale from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2016. *Quality in the Undergraduate Experience: What Is It? How Is It Measured? Who Decides? Summary of a Workshop*. Washington, DC: The National Academies Press. doi: 10.17226/23514.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, non-governmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org.**

PLANNING COMMITTEE FOR QUALITY HIGHER EDUCATION: WHAT DOES IT MEAN, HOW IS IT MEASURED, AND WHO DECIDES? A WORKSHOP

PAUL N. COURANT (*Chair*) Arthur F. Thurnau Professor, Harold T. Shapiro Collegiate Professor of Public Policy, Professor of Economics, Professor of Information, and Faculty Associate in the Institute for Social Research, University of Michigan
ELLEN HAZELKORN, Policy Advisor to the Higher Education Authority, and Director, Higher Education Policy Research Unit (HEPRU), Dublin Institute of Technology, Ireland
PAUL J. LeBLANC, President, Southern New Hampshire University
ALEXANDER McCORMICK, Associate Professor of Educational Leadership and Policy Studies, Indiana University Center for Postsecondary Research, and Director, National Survey of Student Engagement
MARCO MOLINARO, Assistant Vice Provost for Educational Effectiveness, and Director, Center for Educational Effectiveness, University of California, Davis
INDIRA NAIR, Former Vice Provost of Education and Professor, Engineering and Public Policy, Carnegie Mellon University
ROY A. SWIFT, Executive Director, Workcred

National Academies of Sciences, Engineering, and Medicine Staff
THOMAS RUDIN, Director, Board on Higher Education and Workforce
AQILA COULTHURST, Associate Program Officer, Science, Technology, and Economic Policy
MARIA LUND DAHLBERG, Associate Program Officer, Office of Communications
IRENE NGUN, Research Assistant, Board on Higher Education and Workforce
ADRIANA COUREMBIS, Financial Associate, Board on Higher Education and Workforce

Consultant
KARIN MATCHETT, Rapporteur

Acknowledgments

The committee thanks the Lumina Foundation for its support of this workshop.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Academies of Sciences, Engineering and Medicine's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for quality and objectivity. The review comments and draft manuscript remain confidential to protect the integrity of the process.

We wish to thank the following individuals for their review of this report: Karen Elzey, Business-Higher Education Forum; Jordan Matsudaira, Cornell University; Josh Trapani, Association of American Universities; Scott Watson, Education Policy Institute; and Joshua Wyner, The Aspen Institute.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the content of the report, nor did they see the final draft before its release. The review of this report was overseen by Norman Bradburn, University of Chicago. Appointed by the Academies, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the rapporteurs and the institution.

Contents

Chapter 1 Introduction **1**
Chapter 2 Defining Quality **9**
Chapter 3 Improving Quality **19**
Chapter 4 Measuring and Communicating Quality **29**
Chapter 5 Suggested Actions **33**

Appendixes **41**
Appendix A Workshop Agenda **43**
Appendix B Quality in the Undergraduate Experience—A Discussion Document **47**
Appendix C Defining and Measuring Institutional Quality in Higher Education **57**
Appendix D Planning Committee and Staff Biographies **81**

Chapter 1
Introduction

BACKGROUND

Much of the focus on "quality" in undergraduate education in recent years has been on a combination of input factors and outcome measures. Reputation, entrance examination scores and admissions selectivity, financial resources, graduation rates, graduates' employment and earnings, and other attributes are imperfect measures of the overall quality of a college or university, but they do provide some metrics to help consumers assess the value of their investment in postsecondary education. Yet, educators, policymakers, employers, and other interested stakeholders continue to strive for more comprehensive indicators of a "quality undergraduate experience," including those that measure student learning outcomes and graduates' readiness for success in the workforce.

Students, parents, and government agencies—all of which invest heavily in postsecondary education in the country—need as much information as possible about the outcomes of the higher education experience and the extent to which they can expect a fair return on their investment in higher education. Parents and students want some assurance that their investments will result in, among other things, the capacity of the students to secure well-paying jobs and have a fulfilling career. Governments—especially the U.S. federal government, which invests $75 billion annually in higher education,[1] much of it through student support programs such as Pell Grants—also want assurances that their investments will benefit students as well as the larger society. The College Scorecard,[2] released by the U.S. Department of Education in summer 2015, is an

[1] See Schroeder, Ingrid, et al. (2015). *Federal and State Funding of Higher Education: a changing landscape.* The Pew Charitable Trusts. Figure 2. http://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2015/06/federal-and-state-funding-of-higher-education.

[2] The College Scorecard is an online interactive tool developed by the U.S. Department of Education to provide students and families with information to help inform a college search process—including location, size, campus setting, and degree and major programs. Each Scorecard

example of a tool that focuses on a few quantitative indicators of the value of institutions.

A major remaining challenge, then, is to better understand the concept of quality in terms of the full range of student experiences at an undergraduate institution. This can be defined broadly as enabling students to acquire knowledge in a variety of disciplines and deep knowledge in at least one discipline, as well as to develop a range of skills and habits of mind that prepare them for career success, engaged citizenship, intercultural competence, social responsibility, and continued intellectual growth.[3] Although these outcomes are difficult to measure in a standard way that allows for easy comparison across programs and institutions, they are educational outcomes that students, parents, and employers value.

OBJECTIVES FOR THE WORKSHOP

In response to this challenge, an ad hoc planning committee of the National Academies of Sciences, Engineering, and Medicine (the Academies) Board on Higher Education and Workforce (BHEW), with funding from the Lumina Foundation, organized a workshop in Washington, D.C., on December 14-15, 2015. As outlined in the Statement of Task, the workshop goals were

1. To engage scholars and researchers—as well as leaders from higher education, business, civic organizations, and government—in focused discussions about quality in the undergraduate educational experience.
2. To begin to understand how to define and measure those factors that contribute to a quality educational experience that are difficult to quantify but represent the core elements of a successful undergraduate experience for most students.
3. To identify key questions and research themes for possible further study on the definition, measurement, and determination of a quality education.
4. To stimulate further research and dialogue among education leaders and policymakers on the topic of quality, which could in turn influence both *institutional* policy and practice and *public* policies at the federal and state levels.

The planning committee intended for the discussions among college and university faculty and administrators; state and federal agency officials, legislators, and staff; accreditors; policy organizations; business leaders and

---

also includes five pieces of data about a college: costs, graduation rate, loan default rate, average amount borrowed, and employment. More information is available at http://www.ed.gov/news/press-releases/education-department-releases-college-scorecard-help-students-choose-best-college-them.

[3] Workshop participants were asked to provide their own definitions of quality as it pertains to undergraduate education. Those definitions are presented in Chapter 2.

industry associations; students; and other stakeholders to focus on improving our understanding, definition, and measurement of educational quality across the range of undergraduate institutions in the United States.

GUIDANCE AND MATERIALS GIVEN TO WORKSHOP PARTICIPANTS PRIOR TO THE EVENT

Prior to the workshop, each participant received two background papers that set the stage for the presentations and panel discussions: "Quality in the Undergraduate Experience: a discussion document" (see Appendix B) and a commissioned paper authored by Jordan Matsudaira, assistant professor in the Department of Policy Analysis and Management at Cornell University, "Defining and Measuring Institutional quality in Higher Education" (see Appendix C).

The planning committee's discussion document focused on five themes: the measurement of student learning; qualitative factors often cited as important outcomes of undergraduate education; the importance and challenges of assessment; federal policy implications of assessing quality; and the importance of context with regard to institution type, learning environments, and student goals. It concluded with a set of questions intended to guide the workshop discussions:

1. What actions are required in the next 2 years to move us from current models of measuring student learning (e.g., VALUE Rubrics,[4] PULSE,[5] and DQP[6]) that are implemented on an ad hoc basis to a system of quality measurement whereby a group of like institutions adopts a standard set of indicators and reports their results, keeping in mind the work of the Voluntary System of Accountability (VSA)[7] and the related community college effort, the Voluntary Framework of Accountability?[8] What are the next steps in the process of implementing such a system, even on a pilot basis?
2. Now that the College Scorecard[9] has been released, what further steps should the federal government (and, possibly, state governments) take to improve public information about the quality of undergraduate institutions? Are there improvements to the College Scorecard that are feasible and desirable in the near term? If so, who should be responsible

[4] Valid Assessment of Learning in Undergraduate Education, see https://www.aacu.org/value/rubrics.

[5] Partnership of Undergraduate Life Science Education, see http://www.pulsecommunity.org/.

[6] Degree Qualifications Profile, see http://degreeprofile.org/.

[7] See http://www.voluntarysystem.org/.

[8] See http://vfa.aacc.nche.edu/Pages/default.aspx.

[9] See https://collegescorecard.ed.gov/.

for implementing them? What structures should be put in place to assure that the College Scorecard is well-curated and can improve over time?

3. Can—and should—a group be assembled to create a core set of principles to guide the development of a general framework for measuring quality in undergraduate education—one that can be adopted by nearly any type of institution (e.g., 4-year university, 2-year college, online institution, "boot camp")? If so, who should be involved in that process, who should lead it, and who should fund it? How could such an entity build on many of the existing rubrics and tools that have been recently developed?
4. What might be the most appropriate role, if any, for the Academies? Could they, for example, serve an integration and synthesis role, bringing together and leveraging the good work that is under way (including DQP, VALUE, VSA, and perhaps other emerging programs)? Might they also seek to broaden the emphasis from defining competencies and outcomes to working out the quite thorny assessment and consumer information components?

The Matsudaira paper provided background on the topic and a substantial overview of the research that has already been conducted on defining and measuring institutional quality. Among the key points made in the Matsudaira paper were as follows:

> The goal of developing quality indicators for higher education is to enable better decision-making on the part of prospective students, higher education officials, and policymakers to improve the quality of education offered by institutions and to guide students to institutions offering better quality. Institutional quality is multidimensional, and the various users of quality information might place different weight on each dimension of quality.
>
> Quality should be viewed as the extent to which an institution increases the likelihood of achieving various educational goals—that is, as the causal impact of attending an institution on some outcome of education. Defining and measuring the various desired educational outcomes of higher education are major challenges to creating better quality indicators.
>
> New information about the outcomes of students attending institutions, such as the cohort completion rates, debt repayment, and median earnings found on the College Scorecard, represent a large stride forward in developing institutional quality measures. But differences in these measures represent both differences in quality as well as differences in the family income, career interests, and academic preparation of the students that institutions enroll. Isolating quality from these "selection effects" is an important challenge to resolve.
>
> Causal estimates of institutional effects on student outcomes are highly sensitive to variations in the statistical models used. Although progress has

been made, the research literature has yet to reach consensus on the best methodology to measure these causal effects.

In addition to validating methods to estimate the causal impact of institutions, more work is needed to develop measures of student outcomes aside from their labor market success. The lack of broader quality measures, such as students' learning and subjective well-being, have caused ongoing accountability efforts—such as state performance-based funding initiatives for state higher-education institutions—to focus only on earnings and completion outcomes. This poses the risk of incentivizing institutions to allocate resources toward a narrow set of educational goals.

Workshop participants considered the ideas and themes in both papers throughout the 2 days. The workshop itself included panel sessions, expert presentations, small-group discussions of key topics and themes, and large-group "report-outs" and discussions about the topics and themes. The workshop agenda is included in Appendix A.

ORGANIZATION OF THE SUMMARY

This summary is organized into major themes that arose during the workshop: defining quality, improving quality, and measuring and communicating quality. These themes should not be construed as reflecting consensus or endorsement by the committee, the workshop participants as a whole, or the Academies.

This document has been prepared by the workshop rapporteurs as a factual summary of what occurred at the workshop. The statements made in this volume are those of the rapporteurs and do not necessarily represent positions of the workshop participants as a whole, the steering committee, the Board of Higher Education, or the Academies. The workshop did not attempt to establish any conclusions or recommendations about needs and future directions, focusing instead on issues identified by the speakers and workshop participants. In addition, the planning committee's role was limited to planning the workshop.

CONTEXT

Opening remarks by planning committee chair Paul Courant (University of Michigan) framed many of the topics and questions explored during the course of the workshop. Courant articulated several questions commonly posed today to higher education: What are you doing? Why is it so expensive? Does it really work? Is it worth it? He noted that higher education asserts vigorously—and accurately, in his opinion—that there are good reasons for the high price, but higher education is not quite as good at communicating its quality and value to the public, parents and students, and government agencies.

Other participants explored some of these ideas during the course of the workshop. Paul LeBlanc (Southern New Hampshire University), for example, noted that institutions often make clear claims about their students' learning but are not able to back up their claims. In order to assess quality at the institutional

level or across higher education, he said, institutions reframe the question "What do students know?" to "What can students do with what they know?"

Several workshop participants discussed how new demands are being placed on higher education because its traditional institutions were established during an era with different expectations. Sally Johnstone (Western Governors University) noted how the current system was not purposefully designed, but rather evolved over generations. Institutions were created as places where a group of experts convened and shared their knowledge with students, functioning as students' primary sources of information. Individual participants pointed to new technologies as one major influence on the evolution of universities' roles and students' experiences and needs. For example, students today have easy access to disciplinary content on the Internet through sources such as the Khan Academy and other online content providers of instruction.

Other reasons for increased concern about quality in undergraduate education that arose during the workshop included (1) a growing concern by the federal government about the quality of instruction and the return on investment—driven in part by its spending on financial aid, which has more than doubled in recent years, (2) expressions of dissatisfaction by some employers regarding the skills and proficiencies of new graduates, and (3) an increasing concern about whether underrepresented minorities and first-generation college students have adequate access to quality undergraduate education that is designed for the social and academic challenges many face. Participants elaborated on employers' experiences and needs in particular—especially in light of the changing workforce, which seems to demand higher levels of numeracy, problem-solving, and critical thinking skills—as well as on the need for equity and inclusion as an integral part of any conversation about quality.

Individual participants cited a number of current indicators of and assumptions about quality, which they considered valuable or reasonable but insufficient. Among the sources of indicators mentioned were the College Scorecard, employer satisfaction surveys, results from the National Survey of Student Engagement, results of the College Learning Assessment (CLA), and the Gallup-Purdue Index on life satisfaction. Participants also mentioned several important initiatives to improve quality that are already completed or under way, including

- Degree Qualifications Profile (DQP), "a learning-centered framework for what college graduates should know and be able to do to earn the associate, bachelor's or master's degree"[10]
- Liberal Education and America's Promise (LEAP), "a national public advocacy and campus action initiative of the Association of American Colleges & Universities (AAC&U)"[11]

[10] See http://degreeprofile.org/read-the-dqp/dqp-cover/.

- Workcred, an affiliate of the American National Standards Institute looking "to strengthen workforce quality by improving the credentialing system, ensuring its ongoing relevance, and preparing employers, workers, educators, and governments to use it effectively"[12]
- Credentials Transparency Initiative, a joint venture of George Washington University's Institute of Public Policy, Workcred, and Southern Illinois University "to help align credentials with the needs of students, job seekers, workers and employers"[13]

Several participants also cited institutional attributes that are often treated as proxies for quality but whose causal connections to quality have not been proved. LeBlanc said that too often claims of quality have been based on an institution's having a sufficient number of faculty from reputable schools, students admitted with high SAT scores, and substantial volumes in the library, to the neglect of student outcomes. Although some of those outcomes are now being measured, they need to be fleshed out further: "How do we know? Do we have the kind of hard-nosed regular assessment that allows us to test those claims?" he continued. Scott Ralls (Northern Virginia Community College) believes that many assume the quality of an institution increases if it is more selective. Ellen Hazelkorn (Dublin Institute of Technology) highlighted the tendency to define outcomes using only the top 100 institutions in global rankings as a guide. However, she noted that currently there are 18,000 higher education institutions as defined by the Organisation for Economic Co-operation and Development (OECD) and United Nations Educational, Scientific and Cultural Organization (UNESCO), which means that the quality definitions determined by the top 100 institutions represent about 0.5 percent of the world's institutions and about 0.4 percent of the world's students.

These opening discussions set the stage for a series of panel discussions and small-group conversations focused on potential next steps for clarifying the definitions of quality, measuring the quality of student learning and mastery of skills, and developing an accountability system that communicates indicators of quality to the various stakeholders while protecting the academic freedom of postsecondary institutions.

---

[11] See https://www.aacu.org/leap.
[12] See http://www.workcred.org/About-Workcred/Default.aspx.
[13] See http://www.credentialtransparencyinitiative.org/Default.aspx.

Chapter 2
Defining Quality

WHAT IS QUALITY? WHERE DOES IT MANIFEST ITSELF? AND BY WHAT MECHANISM?

Both panelists and audience members were asked by discussion leaders to identify where quality manifests itself and by what mechanisms, and to offer their definitions of quality. In conjunction with these questions, workshop participants also offered their views of those measures that they consider to *not* reflect—or fully reflect—the quality of an undergraduate education.

A number of participants recognized that current metrics are usually focused on economics (cost of the education and graduates' earnings) or graduation rates. Many participants doubted the usefulness of economic metrics and graduation rates as measures of quality, either in conjunction with one another or in isolation. Scott Swail (Educational Policy Institute) advised against looking at earnings in isolation, noting that earnings are a result of a variety of factors beginning much earlier than the student's time at a particular institution. Jordan Matsudaira (Cornell University) discussed the example of graduation rates, one of the only indicators tracked at the federal level. He noted that in the 2-year sector, institutional completion rates are almost completely uncorrelated with the post-enrollment median earnings for institutions' former students, and the correlation is still fairly weak, though positive, in the 4-year sector. Matsudaira reasoned that earnings miss important dimensions of institutional quality, but the Scorecard earnings measures are some of the few quality indicators based on an outcome that students care about and are available for almost every higher-education institution. Paul Courant (University of Michigan) noted more generally that notions of quality tend to focus on things that are easily measured. Several participants, including Alexander McCormick (Indiana University), encouraged institutions to move from an externally focused compliance mindset, heavily influenced by the accountability discourse, to one of professional responsibility.

Is a Core Set of Metrics Possible?

A number of participants debated whether a single definition of quality that applies across the spectrum of undergraduate education is possible, or whether institutions are too diverse. Some participants believed that institutions are too diverse for a single set of metrics to be applied universally, given that postsecondary education spans from online institutions to "boot camps" to community colleges to 4-year colleges and universities. Further complicating the notion of a single set of metrics is the diversity of students, who vary considerably in their needs and expectations, as well as in their academic, social, and economic backgrounds. James Kvaal (the White House) urged the group to focus on these differences among types of students. One participant cautioned against distilling quality into a single number.

Other participants expressed the view that it is indeed possible to arrive at a core set of metrics. Josh Wyner (Aspen Institute) argued in favor of a single, integrated definition of quality and described the Aspen Institute's four-part definition of excellent colleges, which includes learning, completion, labor market, and equitable outcomes.

Courant offered a small set of quality attributes that might be applicable across a wide variety of institutions. An institution can (1) articulate its goals for student learning and regularly assesses its ability to meet those goals; (2) utilize research to better understand what contributes to better learning outcomes, for example, whether and how well students learn through hands-on activities and other experiential curricula and labs; (3) act on its measurements of quality and makes improvements where necessary (and using information technology where available); (4) provide students with the information necessary to make good choices about courses, labs, and internships; and (5) actively recruit students from underrepresented groups and provides them with necessary supports.

Ellen Hazelkorn (Dublin Institute of Technology) raised the subject of perverse outcomes and cautioned against focusing on a set of indicators that are too narrow, which could encourage types of behavior in student choice that may not fully meet societal needs. She said, "It's the broad range of things that we're seeking to look at. … What is it that we are trying to incentivize and encourage institutions to do as we look at measuring these issues around quality?"

McCormick described quality as a three-legged stool built from (1) choices that the institution makes about providing support for student success and incentivizing the right things, (2) choices that faculty make about calling upon students to apply their learning, and (3) choices that students make about spending their time and selecting courses.

QUALITY AS DEPENDENT ON THE INSTITUTION'S MISSION

Numerous participants advocated for using definitions of quality that are specific to an institution's mission. For example, participants of one small group distinguished between public and private institutions, as well as between

institutions that are and are not focused on research, stating that the criteria for quality may vary between those types. The group encouraged participants to disaggregate models of higher education—answering questions of quality in the context of missions and models of higher education, for example, a residential liberal arts model, a research university, an engineering school, a conservatory, or even a "boot camp." Kvaal distinguished between postsecondary education programs designed for specific economic outcomes and programs designed to confer broader skills and suggested that "quality outcomes" will vary strongly depending on institutional mission and purpose.

QUALITY AS DEPENDENT ON THE NEEDS OF THE END USER

Definitions of "End User"

Considerable discussion centered on quality as a function of the needs of higher education's "end users." Participants from one small group stressed that considering the viewpoints of different interested parties—such as audience, customer, consumer, or investor—affects how one looks at quality. The two primary types of end users discussed during the workshop were students and employers.

Several participants expressed interest in creating typologies of students, each having a certain set of needs and goals. Hazelkorn described how the European Union project, U-Multirank, includes multi-ranking in an attempt to address the needs of different types of students.[1] Jillian Kinzie (National Survey of Student Engagement and National Institute for Learning Outcomes Assessment) noted that her organizations incorporate students' perspectives into their assessments of the quality of educational experiences in college and universities and learning outcomes. Steve Crawford (George Washington University Institute of Public Policy) and Courant suggested informing typologies with questions such as "Where do people like me, with my attitudes, my interests, my behavioral traits, get a quality education?" and "What are my prospects for success in realizing my objectives at this institution?" Courant considers this approach to be more valuable than seeking broad measures of an institution's quality.

Mark Tuominen (University of Massachusetts–Amherst) encouraged the participants to include personal development and subjective well-being in the definitions of quality. He called attention to the multiplier value of social and emotional intelligence and the "attitudinal aspects" that complement the knowledge, skills, and abilities that students gain through their programs of study.

[1] See http://ec.europa.eu/education/tools/u-multirank_en.htm.

How Well an Institution Supports Students' Expectations and Needs

Participants characterized two major types of students as typical college-aged students (18- to 22-year-olds) and "adult learners" (people ages 23 and older who are, or have been, in the workforce). Some participants believed that more should be known about how parents and both types of students determine quality and how they base their decisions about postsecondary applications and enrollment. One participant stressed the importance of listening to today's students and offering them a platform to describe their views of a quality education.

A number of participants described how students are moving through education in new ways. Paul LeBlanc (Southern New Hampshire University) spoke of the increasing inclination of students to consume their learning in smaller units, leading to more discussions in higher education about "nano degrees" and "micro-credentials." Sally Johnstone (Western Governors University) described students "who are moving in their own directions, motivated to find the best resources out there, and using their institutions of higher education as a guide to help them get to where they want to go."

Andy MacCracken (National Campus Leadership Council) believes that many students acquire the foundational skills that prepare them for the workplace outside of the classroom. He asked, "Why are the things that are valued in the classroom, that will get you that 4.0, diverging from the things that will get you the best shot to be successful in your career?" He cited *2015 Students Speak Report on Federal Student Aid & Job Readiness,* which surveyed student leaders at U.S. institutions about important considerations for the federal investment in financial aid and the experiences that best prepare students for success outside of the classroom.[2]

Individual participants described one element of quality: the information made available to students to help them navigate through their courses of study. David Dill (University of North Carolina) stated that "students are making crucial decisions about their future life chances every day on the basis of the information we provide them, which is heavily flawed" and suggested that institutions should examine the structure of the curricula in academic programs. He asserted that the large number of possible courses, and the even larger number of combinations of courses, makes it impossible for faculty to accurately judge the quality of a student's education or to know how to improve that quality.

Jennifer Engle (Bill & Melinda Gates Foundation) highlighted the importance of providing accurate information to students who attend a community college or the local 4-year public institution or whose choice is confined by limited

[2] MacCracken, Andy, and James Scimecca. (2015). *Students Speak Report on Federal Student Aid & Job Readiness*. National Campus Leadership Council. http://www.nationalcampusleaders.org/wp-content/files/2015/07/NCLC-Report-on-Federal-Student-Aid-Job-Readiness-July-2015.pdf.

knowledge. She advocated for better structured education pathways so that students do not have to navigate the course catalogue blindly. Institutions should define sets of courses and experiences that students can follow toward their desired outcomes, she said. Engle emphasized that high-quality internal information is particularly important for students from underrepresented groups.

How Well an Institution's Graduates Succeed in the Workforce

Courant cited a recent Gallup poll that found that only 11 percent of business leaders and 14 percent of the American public believe that colleges adequately prepare students for the workforce, while 96 percent of chief academic officers believe that their graduates are workforce ready.[3] MacCracken suggested that, given how often the average person changes careers over a lifetime, the foundational elements of a quality education are the skills that apply to different occupations and fields.

Roy Swift (Workcred) stated that employers often report that universities are not listening to their input on students' preparation for the workforce. Even when higher education has defined learning outcomes, they often differ from industry needs—especially given the changing nature of the workforce. Because the world economy is changing rapidly, he believes the United States should implement a system for ongoing, systematic communication between industry and higher education. Emily Slack (Education and Labor Committee, U.S. House of Representatives) said that the responsibility for preparing students for the workforce is a shared one, not resting with only higher education or businesses or the state or the federal government. If an employer needs graduates with a specific skill set, then she believes it should communicate and form partnerships with local community colleges or universities to ensure that the appropriate education and training programs are in place—and perhaps subsidize those programs. Swail suggested that the conversation about institutions' course offerings and faculty positions, which are often based on institutional inertia and failure to consider societal and industry needs, should be reinvigorated.

During her keynote address, Carol Schneider (Association of American Colleges & Universities [AAC&U]) discussed AAC&U's extensive work with employers to advance high-quality education. She noted that "the bar is being raised for what the economy expects from people who are getting a college education" and "the rising demand for the ability to deal with complex problems, new information and the declining demand for routinized cognitive skills."[4] Ninety-three percent of employers surveyed by AAC&U value a job

[3] Busteed, Brandon. (2014). *Higher Education's Work Preparation Paradox*. Gallup. http://www.gallup.com/opinion/gallup/173249/higher-education-work-preparation-paradox.aspx.

[4] Hart Research Associates. (2013). *It Takes More Than a Major: Employer Priorities for College Learning and Student Success.* Washington, DC: AAC&U.

candidate's capacity to think critically, communicate clearly, and solve complex problems more than his or her major. Employers want to hire people who can transfer their learning from one concept or field to another. In addition, 95 percent of employers seek graduates who can contribute to innovation by thinking outside the box. Schneider also noted that 78 percent of employers believe that students need a broad foundation in the liberal arts and sciences and called workshop participants' attention to a set of learning outcomes cited often by AAC&U's members and surveyed employers. These outcomes include knowledge of human cultures and the physical and natural world, intellectual and practical skills, personal and social responsibility, and integrated learning.[5]

Several participants cited professional and basic cognitive skills as important outcomes of a quality education, noting that employers increasingly report these skills as weak among their new hires. LeBlanc relayed conversations with CEOs and human resources directors who have found that many graduates of reputable 4-year undergraduate institutions still lack basic skills such as writing and quantitative skills. Swift noted that industry is increasingly calling for graduates who can understand and identify problems, solve problems, synthesize data and information, and offer solutions. Aprille Ericsson (NASA) explained that NASA seeks new hires with not only strong engineering skills, but also collaborative and leadership skills. "I have to have people who go out there and get it, who will work to the 12th, the 13th hour solving a problem," she said.

Hazelkorn described data from the Organisation for Economic Co-operation and Development (OECD) Programme for the International Assessment of Adult Competencies (PIAAC) that shows a growing gap between credentials and competence. Matsudaira said that the National Assessment of Adult Literacy found a notable decline between 1998 and 2003 on measures of document literacy, prose literacy, and quantitative literacy among college graduates. Kinzie noted that AAC&U's employer surveys have revealed the same issues about the quality of learning skills. Julie Carnahan (State Higher Education Executive Officers Association), whose focus is on public institutions, said that employers are telling her state-level members that graduates, despite their high grade point averages, cannot perform the work required employers. In response, 12 state-level executive directors have formed a multi-state collaborative to assess learning outcomes.

---

http://www.aacu.org/leap/public_opinion_research.cfm. See also Kuh, George D. (2008). *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter.* Washington, DC: AAC&U. https://www.aacu.org/leap/hips.

[5] See http://www.aacu.org/leap. For 2009 findings, see Hart Research Associates. (2009). *Learning and Assessment: Trends in Undergraduate Education—A Survey Among Members of the Association of American Colleges and Universities*. Washington, D.C.: AAC&U. For 2015 findings, see Hart Research Associates. (2016). *National Trends in General Education Design, Learning Outcome, and Teaching Approaches.* Washington, D.C.: AAC&U.

Participants in one small-group discussion noted the interdisciplinary nature of the skills needed by employers and by students more broadly to solve the complex problems of this century. They recognized that the current university structures were not designed to foster this interdisciplinarity.

Individual participants explained that graduates must not only possess stronger basic skills but also be better able to apply their knowledge in real-world situations. MacCracken said, "A common theme that we heard from employers is that 'today's graduates are the best we've ever seen.' But they don't know how to communicate what they know or apply what they know to different settings." A basic standard in higher education should be that students can think critically, communicate their thoughts, and apply their knowledge in the field. To this end, he said that higher education should provide ample opportunities for students to gain learning experiences outside the classroom that are incorporated into multiple disciplines experiences rather than treated as extraneous events.

QUALITY FOR ALL STUDENTS: EQUITY AND INCLUSION

Several participants stressed the importance of addressing the particular needs of students from underrepresented groups. Calling attention to what she considers a distinctive element of American higher education—its commitment to excellence and inclusivity—Kinzie stated that inclusion should be a common thread in discussions of quality. Elsa Núñez (Eastern Connecticut University) encouraged higher education leaders to focus on quality for all students, not only a certain segment of the population. Schneider reminded the audience of the deepening divides in college completion: high-income students are more likely to earn college degrees. Schneider said, "We have ascended to a highly stratified tiered system of higher education in which some students are routinely steered to institutions that would give them the fullest and the most empowering possible education while others are steered off to narrow, short-term programs or to apply to those degrees which are devoid of the broad learning that people need to understand the global economy they're part of."

Ericsson highlighted a stronger background in mathematics as a key to college success for underrepresented minorities. The students whom she sees succeeding in engineering or technology coursework received strong mathematics education at an early age: "I'm seeing students taking Calculus 3 before they leave high school." Although some students' deficits may be due to deficiencies in K-12 education, she suggested maintaining focus on higher education. She believes that higher education should instill mathematical skills in its students, regardless of the level of knowledge and skill they bring to college, as a necessary component to success in STEM careers.

One participant highlighted the distinction between whether an institution *is providing* a quality education in which diverse students are well supported and

well challenged in their learning environments and whether an institution *is capable of providing* that education.

QUALITY AS INSTITUTIONS' COLLECTIVE IMPACT

Participants debated whether quality is measured—and institutions are accountable—solely during a student's matriculation or a longer, collective time period.[6] Cliff Adelman (Institute for Higher Education Policy) reminded the audience that "the temporal framework in which I have to influence anything that you [the student] do is when you're with me. After you leave me, I can't control any of that." Several other participants expressed the counter view that higher education leaders must care about a student's pre- and post-institution experiences, advocating for a focus on institutions' collective impact. McCormick, for example, described what he considers a faulty assumption that if every institution in a system is achieving its own quality objectives, then the system itself is optimal. Although this creates an inferential problem for attributing student success to a particular institution, he believes that it is important to approach the quality question from a system level rather than one institutional experience at a time.

Wyner agreed that the post-graduation success of students is an important measure of quality and argued for institutions to consider quality not only from an internal institutional perspective but also in terms of collective impact—meaning how they serve students in combination with other institutional types (such as K-12, other postsecondary institutions, and employers). He believes that a school's internal measures increasingly need to align with those of other institutional types. Wyner suggested an integrated set of quality measures for which all institutions take some ownership, recognizing that certain elements may be more important to some institutions than to others. Employers and higher education could then compare the common measures against the graduates' actual performance in the workforce.

Scott Ralls (Northern Virginia Community College) relayed the perspective of community colleges as "pathway" colleges (i.e., not "destination" colleges) that must be attuned to the requirements of the institutions that their students are trying to reach, whether universities or employers. He described how his college must be proactive about developing and using feedback loops with universities and employers and must reach out to both on an ongoing basis. Marco Molinaro

[6] The issue of to what degree institutions should be held accountable for outcomes post-graduation (and to what degree those outcomes reflect quality) is a common aspect of many discussions. For instance, the Gallup-Purdue Index has grappled with relating later life satisfaction to undergraduate experiences. The Post-Collegiate Outcomes Initiative, a collective partnership between the American Association of Community Colleges, the Association of Public and Land-grant Universities, and the American Association of State Colleges and Universities, also touched upon these issues. More information about the Post-Collegiate Outcomes Initiative is available at http://www.aacc.nche.edu/AboutCC/Trends/pco/Pages/default.aspx.

(University of California [UC], Davis) mentioned that the outcomes of UC–Davis students who started at community colleges are sometimes better than if they had started at the university. Partnerships between UC–Davis and the community colleges have been critical in the university's efforts to improve quality.

QUALITY AS SERVING THE PUBLIC GOOD

Several participants expressed the importance of integrating societal needs into definitions of quality. Wyner asked, "How can we integrate the needs of employers with what the academe's objectives in terms of civic engagement and democratic participation?" Schneider said, "It is not just that the economy is demanding more, but that our democracy is demanding more." She suggested that definitions of quality include the acquisition by students of the capabilities needed for an informed and effective citizenry. Carnahan stressed the importance of identifying the parties responsible for the good of the whole: "Is this the responsibility of faculty, institutions, states, or the federal government?"

QUALITY AS CONTINUOUS IMPROVEMENT OR AN INSTITUTION'S ADAPTABILITY

Molinaro stated that quality requires a process of continuous improvement at the course, program, and institution levels. Goals should be clearly enumerated, and approaches designed to measure the related outcomes, in his opinion. The desired result in his view is a cyclical process of continuous improvement that leads to agreement among the interested parties that the education experience is high quality. Johnstone highlighted the need for institutions to adapt and clearly determine what their students have learned. Ralls asked, "How can that adaptation be made in a way that not only preserves academic quality and learning outcomes, but also causes the institution to thrive and survive over a period of time?" Núñez believed that measures of quality are not static, but rather a delta. "It is how those numbers change over time that is an index of how the institution itself is growing and thriving to fit the realities of the student population and the external environment."

Chapter 3
Improving Quality

During the workshop, participants offered specific suggestions for improving quality. Keynote speaker Carol Schneider (Association of American Colleges & Universities) said, "We're not starting fresh and trying to define quality," noting that higher education should pay attention to what educators and employers have been saying about this topic for some time. In her view, higher education already has (1) clarity on the goals for student learning, (2) research showing that high-impact practices for engaged learning are effective, (3) a way to capture that information and examine the situation on specific campuses (made possible by National Survey of Student Engagement), and (4) tools for mapping the goals onto the curriculum and thereby supporting continuous improvement. Schneider believes that the "definition of quality is the least of our challenges" and that more imperative is the need for a broader policy discussion that incorporates what is known about the needs of employers, students, and society overall into the strategies that institutions can implement to meet those needs.

INSTITUTIONAL ATTRIBUTES AND EFFORTS

Workshop participants offered a number of suggestions for how institutions—individually and regionally—could improve the quality of the educations they provide. Several participants spoke about the need for strong leadership, both from a school's administration and from faculty champions, stressing that ongoing conversations about quality should become the norm on college and university campuses.

Affordability and Inclusion

Several participants discussed metrics based on an institution's mission. James Kvaal (the White House) suggested that an institution's affordability should be included in measures of quality. He explained that institutions that serve students from all backgrounds are "very important for the future of our society," precisely because they produce a large number of graduates in an affordable way. Several participants mentioned that inclusion—the quality of an education

provided for students from underrepresented groups—should be measured at every type of institution. One participant suggested that an institution needs both the capacity to provide quality education to diverse students as well as the resources to deliver on that capacity. Carlos Castillo-Chavez (Arizona State University) spoke highly of Cornell University's engineering school's inclusion of students from diverse backgrounds, efforts that Castillo-Chavez connected to the school's creation of an office of diversity and recruitment of prominent engineering faculty to direct it.

Network Effects

Some participants mentioned the network effects made possible within and outside of an institution as a signal of quality. For example, Paul LeBlanc (Southern New Hampshire University) discussed the value of a university's music school as extending well beyond music majors. James Grossman (American Historical Association) spoke similarly, highlighting the value of ensuring that students from diverse disciplines are in close proximity of one another—for example, the sciences and humanities. Participants cited an institution's relationships with alumni and community networks as also important, including Aprille Ericsson (NASA) and Scott Ralls (Northern Virginia Community College). Ericsson asked, "How do we maintain those relationships? How do we develop them and allow them to replicate throughout the university environment?"

A Student-Centered Approach

Several participants suggested that higher education institutions must adopt a more student-centered approach to improve the quality of the education experience. For example, Castillo-Chavez noted that it is difficult for individual faculty or an organization to excel in research while also providing a high-quality education to undergraduates: "How do we change the university so that there are good researchers that are student-centered, that are supported, encouraged, and rewarded by the institution?"

Alexander McCormick (Indiana University) highlighted the contrast between the elaborate procedures required for research involving human subjects, including certification of knowledge and ethical responsibilities, and the absence of processes that oversee the quality of new undergraduate courses. He called it an "interesting irony" that essentially no standards or systems exist for review course content and experiences.

Use of Technology to Customize Education to the End User

Several participants mentioned the growing role of technology in U.S. culture, which can be used to customize experiences around the end user but contrasts with higher education's traditional structure. Individual participants noted how

students largely adhere to the pathways created by institutions. Sally Johnstone (Western Governors University) asked whether institutions should play the role of curator of learning experiences from which students select, where disaggregated education and learning experiences are curated around individual student needs. Under this scenario, the unit of quality measurement might be the student him- or herself. She encouraged institutions to provide a "responsive education." Some participants view students as grazers (wanting the freedom to pick and choose) while others view students, especially adult learners pursing job-ready educational experiences, as looking for a clear roadmap to their desired education outcomes.

Use of Tools Designed to Improve Quality

Schneider discussed the Degree Qualifications Profile (DQP), a framework for mapping a curriculum that invites students to integrate their learning across fields and apply their learning to complex problems. Based on a recognition that the major is not enough, the DQP helps to ensure that students have ample opportunities to do real-world problem-solving and draw on perspectives from multiple disciplines—combining experiential with academic learning. The DQP, in addition to folding together those practices known to support student learning, pays specific attention to civic and global learning. Schneider also presented a matrix for mapping the DQP onto the curriculum. The matrix guides faculty in designing student assignments and mapping their programs onto the DQP, thus assisting students in achieving quality learning.

Schneider described a multi-state proof-of-concept study on 2- and 4-year college students' achievement levels in selected essential learning outcomes. Faculty members across many institutions were trained to use the VALUE rubrics, and faculty from other institutions assessed their students' learning in the areas of critical thinking, communication, and quantitative reasoning. Significant numbers of students made gains, but the picture was mixed. She highlighted competencies for which students showed the weakest performance, which included their use of evidence and communication skills.

Schneider concluded her discussion of employers' expectations and tools for improving quality (the DQP and VALUE rubrics) by asking how institutions can be held accountable for providing students with the learning outcomes required for success in today's world. Schneider noted, "Using the VALUE rubrics and using students' work is only a piece of it, because just assessing the work is only going to get us more evidence of under-achievement." She continued, "We need to hold ourselves accountable to thinking in new ways about quality." Step one is clearly defining learning outcomes, which many institutions have done. Step two is mapping these outcomes onto program offerings: "How many institutions have faculty who are prepared to work with students in an inquiry mode of learning" utilizing high-impact practices? "How many faculty can tell you where in their program people are working on ethical learning or collaborative problem solving, and how frequently this is done?"

APPROACHES IN THE CLASSROOM

Instruction Aligned with Research on Learning

A number of participants described actions to ensure alignment of faculty's teaching methods with how people learn, as revealed by research. Schneider outlined several high-impact educational practices that faculty are implementing. These practices included first-year seminars that emphasize critical inquiry, intensive writing, and collaborative learning; learning communities; writing-intensive courses; collaborative projects; diversity/global learning; community-based learning; internships; and capstone courses. She highlighted the finding that these high-impact practices, while positive for all students, have an especially strong effect on the outcomes of students who "traditionally are starting with more strikes against them," such as students from communities of color as well as students with lower SAT or ACT scores.[1]

Linda Slakey (Association of American Universities STEM initiative and Association of American Colleges and Universities) noted that alignment of what is known about student learning with current instruction methods is undergoing rapid change. Several participants, including Roy Swift (Workcred), Cliff Adelman (Institute for Higher Education Policy), and Ralls, advocated for the use of methods that connect what students know with what they can do—that is, understanding and measuring not only the acquisition of knowledge and skills but also the application of knowledge and skills to real-world problems and challenges. Emily Slack (Education and Labor Committee, U.S. House of Representatives) and other participants highlighted the value of approaches such as competency-based education.

A number of participants stressed the importance of institutions addressing broad-based competencies as well as using practices germane to student learning that are more domain-specific. Other elements of quality cited by participants included an institution's ability to (1) meet the needs of people with different expectations; (2) teach students to think critically, communicate their ideas, and apply their knowledge in the field; (3) help students apply knowledge to other occupations or fields; and (4) address students' overall well-being.

Hands-on Learning

McCormick described the Integrated Concentration in Science program at the University of Massachusetts at Amherst, in which students in STEM fields study a concentration in addition to their major. The program's goal is to engage students in real-world problems selected by the students themselves. He

[1] Hart Research Associates. (2015). *Falling Short: College Learning and Career Success*. Washington, D.C.: AAC&U., and; Finley, Ashley, and Tia Brown McNair. (2013). *Assessing Underserved Students' Engagement in High-Impact Practices*. Washington, D.C.: AAC&U.

described the transformation of students as they become more comfortable tackling open-ended, real-world problems, in a team setting, beginning in their first year. Ericsson spoke about the value of hands-on learning in STEM education from an employer's perspective, saying, "We [NASA] build things. … Students actually getting out there and making something, a product: that is what I would like to see on their resumes."

Adaptive Learning

LeBlanc encouraged the group to consider the importance of technology during its discussions, citing advances in data analytics, learning science, and immersive learning environments that can inform understanding about what students know and can do. Marco Molinaro (University of California, Davis) agreed that adaptive learning can be valuable, especially in the remedial or introductory arena and, at the other end of the spectrum, in very specific, highly detailed areas such as flight or medical simulators. To take advantage of adaptive learning, he believes higher education must accept that a model developed at one institution will be valid in other institutions—that is, institutions should adopt broadly available products rather than invest in the development of tools only for their use.

Slack believed that integrating adaptive assessment of students' progress with course material, which alters a student's pace to fit his or her learning, could work well for some students, especially adult learners. In her view, adaptive assessment could potentially revolutionize thinking about what constitutes quality in higher education and student learning.

Instruction Aligned to Employer Needs

On behalf of Adam Enbar of Flatiron School (who was unable to attend the workshop), LeBlanc described the school's strategy for achieving high levels of quality. Flatiron runs 12- to 15-week "boot camps" that provide students with intensive training in Web development, an area with high market demand in the current economy. A projected 1.4 million positions (with an average starting salary of $76,000) will become available in the Web development/IT workforce in the next 5 years, and traditional higher education will produce only about 400,000 graduates. Results obtained by an external auditing firm verify the claims the school makes for their graduates. Flatiron's graduates are receiving a quality education according to several measures, including technical Web development skills, team work, and other professional skills.

THE CHANGING ROLE OF FACULTY

Workshop participants emphasized the importance of faculty leadership in improving the quality of education within the classroom. A considerable amount of discussion focused on faculty roles and how they are changing or, in the view

of some participants, should change. Participants advocated for new ways to differentiate traditional faculty roles.

Faculty as Curators

Johnstone discussed how the current transformation of higher education relies heavily on technologies that were not available when the majority of higher education institutions were founded. She encouraged institutions to create teams of subject-matter experts and researchers who work with learning scientists and technology experts to create optimal learning environments. Individual participants, including Paul Courant (University of Michigan) and LeBlanc, suggested that the role of faculty and institutions may be shifting to one of "curating" information and learning environments. Schneider strongly disagreed with the notion that faculty's job is to curate. She asserted instead that "the intellectual talent that we have invested in and cultivated through our universities has been a driver for everything that's been productive about this country's standing in the world in the economy and creativity, even in democracy."

"Unbundling" the Faculty Roles

LeBlanc affirmed the critical role of faculty and shared his perspective that, although the displacement of faculty through the use of adjunct faculty has been long under way, faculty's role is now being "unbundled": "What are those roles that we ask faculty to do in other ways, to have those functions done more effectively and [perhaps] less expensively?" He suggested that the optimal configuration of faculty roles depends on the educational context. For example, on a residential 4-year campus where 18-year-olds are having "the coming of age experience," it might be best to have more full-time faculty who can focus on student engagement and mentoring. In contrast, faculty roles might be different at an institution with a large online program serving adults "who have had about all the coming of age they can handle" and who want to acquire very specific knowledge and/or skills to prepare for a specific career or job. He encouraged the group to focus less on the issues arising from the increasing use of adjunct faculty per se and more on the functions that faculty provide and how these might be fulfilled. He compared an undergraduate education to the manufacture of a car: a single worker could build a car, bringing to bear the skills needed for every task, or a car could be built by people with expertise in sub-tasks. A car built by one person, he said, is too expensive; likewise "making education affordable is going to increasingly require the unbundling of those roles that are very expensive to maintain."

Other participants, such as David Dill (the University of North Carolina), highlighted issues with the current unbundling of the faculty role that arise from the use of part-time instructors, that is, the lack of integration of part-time instructors' courses into academic programs and into an institution's quality

framework, and the relative absence of informal interactions with tenured faculty, which has historically been a key element of the coordination and integration of programs.

Regarding the next generation of the professoriate, Jay Labov (National Academies of Sciences, Engineering, and Medicine) encouraged the audience to consider efforts to improve the quality of undergraduate education by promoting modifications to graduate education, because most undergraduate professors have graduate degrees. "In many ways, the future of undergraduate education depends upon the future of graduate education," he said.

Supporting Faculty Efforts to Improve Student Learning

Denise Simmons (Virginia Tech) described a need to provide faculty with resources to assist them in making positive changes to their courses and pedagogy. Resources that might be readily available on campuses, for example, instructional designers, could support faculty who both recognize the need to modify their courses and are willing to make those changes.

INSTRUCTION OUTSIDE OF THE CLASSROOM

Several participants spoke about the need for a quality undergraduate education to include connecting students with real-world problems, including providing them with experiences in organizations and businesses in the community. Ericsson, Andy MacCracken (National Campus Leadership Council), Elsa Núñez (Eastern Connecticut University), and others noted the importance of compensating students for their time devoted to off-campus learning experiences, which is a significant consideration for some students from underrepresented minority groups and others who struggle to pay for their educations. MacCracken asked "How do we make sure that students who are already having a hard time paying for this education are getting the experiences that we know will give them a better sense of quality and a better platform to be successful later on in the workplace?" He discussed the potential use of the federal work-study program to connect students with relevant, real-world work experience, which he called "a tremendous resource that is vastly underutilized and under resourced." He explained that a part of the federal program now allows universities to have host sites off campus and described communities that are working with industry to invest in the development of programs similar to federal work-study programs.

Several participants advocated for a stronger relationships between higher education and industry. Swift believes that higher education should acknowledge that learning occurs in many environments and needs to bring industry's real-world challenges to the classroom. He cited a recent meeting hosted by the Business-Higher Education Forum around data science and data analytics that examined the competencies needed by students need in these areas. Participants at that meeting stated that students who encounter industry's

problems in the classroom are able to function in the workplace much more quickly.

ELEMENTS OF A QUALITY EDUCATION THAT ACTIVELY SUPPORT UNDERREPRESENTED MINORITIES

Several participants described elements of a quality education necessary to ensure the inclusion of all students. Schneider discussed the positive effect of research-based, high-impact practices on the experiences of underrepresented minorities. Ericsson argued for the necessity of paid internships, noting, for example, that many bright students at Historically Black Colleges and Universities (HBCUs) struggle to maintain a high grade point average (GPA) because they must work part-time while pursuing their studies. If their GPA suffers, then they are less likely to secure a higher-paying job or a job with a minimum GPA requirement (such as those at NASA) upon graduation. Ericsson acknowledged the value of the Pathways program, which is mandated across federal agencies, but noted that it does not allow the full volume of students to utilize those internship opportunities because of the hiring bottleneck it has created.[2] Castillo-Chavez cited Cornell University's successful efforts to increase diversity in its engineering program through the creation of an office of diversity in engineering. Castillo-Chavez attributed the program's success in part to the leadership of very distinguished faculty, who would not have accepted the positions without Cornell's guarantee that they would be empowered to make significant change. He highlighted the importance of inclusion of strong leadership at the university president's level as well.

A ROLE FOR THE ACCREDITATION PROCESS

Several participants believe that the United States' accreditation process can play a significant role in improving the quality of undergraduate education. Slack stated that the accreditation system is very valuable and that she thought the attention paid to it during the current reauthorization cycle will assist in shifting the process away from an inputs-driven model to one that considers student learning outcomes in light of an institution's mission. She recommended that "we make sure that we are strengthening the accreditation process, not dismantling it," adding that "accreditors, because they are peers, know about quality and about student learning and have a much better perspective on that than some other outside folks that could be in charge of quality." Kvaal voiced the concern that the accreditation process is "too tough and too loose at the same time." For example, an institution's accreditation does not ensure that it provides quality; however, students rely on accreditation as a seal of approval of

[2] See https://www.opm.gov/policy-data-oversight/hiring-information/students-recent-graduates/.

institutional quality. On the other hand, Kvaal believes that concerns that accreditation prevents innovation are legitimate.

ROLES OF PROFESSIONAL ORGANIZATIONS

Several participants discussed how professional organizations play a role in adoption by faculty of teaching strategies known to connect with how students learn. One participant, a member of the Accreditation Board for Engineering and Technology (ABET), suggested that the role of professional organizations in the accreditation process in their own disciplines could be strengthened. For example, the involvement of professional organizations in the accreditation of electrical and computer engineering programs has helped to define student outcomes that are more practical in the private sector.

Another participant, a member of the American Society for Mechanical Engineers, highlighted a role for professional societies in program design. She described experiences with students who completed substantial hands-on projects in high school and entered engineering programs in universities with great enthusiasm. However, after spending the first 2 years not making anything, they leave the major. The American Society for Mechanical Engineers is working with ABET to explore integration of more project-based learning into the first 2 years of the major course of study.

Participants in one small-group discussion acknowledged the presence of James Grossman, Executive Director of the American Historical Association, at the workshop, applauding his leadership and noting the value of the association's involvement in "tuning," a process that utilizes the DQP in a discipline-specific way. Participants in another small-group discussion mentioned a similar project by the American Association for the Advancement of Science, identifying this initiative as an opportunity for collaboration.

Chapter 4
Measuring and Communicating Quality

QUESTIONS AND CHALLENGES RELATED TO MEASURING QUALITY

Quality measures may be useful for internal institutional improvement or for external benchmarking, and often the same quality measures are not useful for both purposes. Wide-ranging discussion occurred at the workshop around the challenges of measuring quality, including how to determine the appropriate level, time frame, and attribution; how to interface with the public good; and how to measure the roles of faculty and institutions.

At What Level Can Quality Be Measured?

Participants debated the appropriate level for best measuring quality—institutional, program or department, or classroom and faculty.

The institution as the focus. Alexander McCormick (Indiana University) noted that there are powerful cultural beliefs in the United States that the institution matters most, but that there is persuasive evidence that the quality of the educational experience and student learning varies more between particular programs or departments within institutions than between institutions. He asked the audience, "Do you think that you experienced quality uniformly at the institution throughout your experience? Did your peers experience quality at that institution in a uniform way?" The belief that quality is an attribute of an institution, he said, is reinforced not only by ranking, but also now by governments asking for evidence of quality and return on investment.

Because educational quality is often delivered program by program, Josh Wyner (Aspen Institute) suggested that the institution and its senior leaders are some of the essential actors who need to better understand quality across programs if the institution is to improve at scale. McCormick noted that although institutions are *an* actor, they are not *the* actor. He observed that institutions can provide and encourage certain conditions for educational

effectiveness, "but no president, provost, or dean can walk up to a dimmer on the wall and turn a switch and ratchet up the quality of education." Paul LeBlanc (Southern New Hampshire University) and James Grossman suggested that an institution-level aspect of quality may emanate from diversity in an institution's program offerings, as in the example of a music school benefiting students other than music majors. These network effects—interactions outside the classroom among students with widely different interests—may add quality to the educational experience. Grossman suggested that the benefit is not only personal, but also public: "How much does the public benefit from political science students and future lawyers and future cooperation executives interacting on a daily basis with artists, musicians, and future clergymen? That's a public good."

The program as the focus. Several participants believe that quality can best be measured at the program level. Jessica Howell (College Board) described how quality in the health care environment is viewed as a specific program, rather than an overall hospital, issue. LeBlanc suggested that the quality discussion could be situated at the program level given the variability between an institution's programs. McCormick explained that National Survey of Student Engagement has identified the program as the primary driver of the student experience. Scott Ralls (Northern Virginia Community College) also believes a focus on the program level to be most appropriate.

The classroom as the focus. Several participants connected quality to teaching methods guided by research on student learning, including hands-on learning, inquiry-based learning, and student connections to real-world problems. Quality, in this case, would be measured at the classroom level. McCormick noted institutions have reliable systems for tracking coursework and credit hours, but it is much more difficult to measure student learning. A number of participants discussed the contributions that adaptive learning assessments could make to efforts to determine quality.

When Can Quality Be Measured, and How Can It Be Attributed?

The group discussed the question of when quality can be measured, whether immediately upon graduation or several years thereafter. One participant noted that feedback will be very different if captured 5 years versus 10 years after graduation. Elsa Núñez (Eastern Connecticut University) pointed out that as information on quality is gathered year after year, higher education's questions, values, and concerns will evolve, as will its data-collection tools.

Individual participants discussed the limitations of assessing the experience of graduates and making inferences about an institution's contribution to their success later in life. Cliff Adelman (Institute for Higher Education Policy) noted that greater than 50 percent of students attend more than one institution and greater than 30 percent attend more than two institutions. Ralls noted that students are not randomly assigned to institutions. Some institutions can select

their students, who tend to arrive with many of the experiences that contribute to their future success, making attribution of their knowledge and proficiencies to a particular institution difficult.

QUESTIONS AND CHALLENGES CONCERNING DATA

Workshop participants discussed the quality and usefulness of existing data and the ways to improve and coordinate data collection.

How Can Data Quality Be Improved or Made More Relevant to the Quality Discussion?

Several participants acknowledged dueling needs for quality data: the need for contextualization (related to an institution's mission or to a particular type of student's needs) and the need for comparability across institutions. For example, Jennifer Engle (Bill & Melinda Gates Foundation) noted, "Even as we want to contextualize, we also have to balance that with a need to provide students with the information that they can compare. Both of those impulses are valid."

Engle described how the Gates Foundation has undertaken a number of initiatives to improve data quality, for example, collecting data through completion initiatives such as Complete College America[1] and Achieving the Dream.[2] She believes that scaling the data collection and analysis process is crucial for expanding innovation. She stressed that quality information is needed for all students (including "nontraditional" and remedial students) and all institutions. She recognized that "it doesn't seem innovative to count all students, and yet that's what's underlying a lot of discussions about why the data are not sufficient" for quality improvement efforts. Referencing the lack of data on nontraditional students, Emily Slack (Education and Labor Committee, U.S. House of Representatives) observed, "It would revolutionize higher education data if we would just count the other 50 percent of students that are out there." Engle noted that current data systems were not designed to capture the experiences of nontraditional students, which in her view, is one of the most prevalent problems, "but also the most easily fixed because we know the students are there. We're already counting their outcomes, but we're not making them part of how we publicly express the outcome of an institution."

How Can Dissimilar Systems Work Together, and What Is the Infrastructure We Might Want?

Several participants highlighted the need to coordinate data-gathering systems. Engle noted, "We have a lot disconnected data systems that were

[1] See http://completecollege.org/.
[2] See http://achievingthedream.org/.

created for their own purposes but none of which were exclusively created for the purposes that we're talking about here, in terms of understanding how students are moving through the college experience." In particular, she believes the outcomes of nontraditional students should be communicated: "How do we start to change the publicly available data systems so that we can better capture those students?" This issue is also relevant to measuring the quality of education delivered to students enrolled in the nontraditional postsecondary education and training programs.

Engle advocated for careful thought about the optimal data infrastructure, including communication between the various data systems, noting that "we need to decrease burden and increase utility." She urged that state systems should communicate better with federal systems, federal systems should communicate better with one another, and private systems—such as a national student clearinghouse—should play a role as well. Engle described the Gates Foundation's data infrastructure working group that is writing papers (released in early 2016 through the Institute for Higher Education Policy), one of which focuses on what institutions need to do to improve data quality and recommendations for action at the state and federal levels.[3]

How Can the Existing Data Be Put to Better Use and Be Coordinated?

Engle described how she believes existing data could be put to better use, including by more audiences such as students and institutions themselves. She asked how institutions might better use the data they are already collecting and connect those data to their campuses' operational data—"How can institutions link performance metrics to what is happening in terms of individual students?" She explained that some institutions make this connection by using technology-enabled advising systems to examine why students are not making sufficient progress toward their degree.

Participants in one break-out group noted the obstacles created by regulations that restrict data-sharing across institutions (e.g., institutional review boards, or the Family Educational Rights and Privacy Act). Sharing these data is important to determining which interventions improve student outcomes and which do not. Increased data-sharing would strengthen trans-institutional conversations about quality.

[3] Engle, Jennifer. (2016). *Answering the call: Institutions and States Lead the Way Toward Better Measures of Postsecondary Performance*. Bill & Melinda Gates Foundation. http://postsecondary.gatesfoundation.org/wp-content/uploads/2016/02/AnsweringtheCall.pdf.

Chapter 5
Suggested Actions

Individual participants offered suggestions for future actions to further the conversation around quality in undergraduate education.

IMPROVING DATA GATHERING, SHARING, AND COMMUNICATING

Jessica Howell (College Board) suggested that institutions should make better use of faculty in departments of economics, education, and sociology who are trained to work with administrative datasets and can determine the causal connections between universities' efforts to improve quality and student outcomes. Participants in one small-group discussion believe that accrediting bodies require institutions to collect a lot of data that are not relevant to true measures of quality and advocated for influencing these bodies to require data that are more meaningful, in quality terms. Several participants, including Emily Slack (Education and Labor Committee, U.S. House of Representatives), suggested that further studies on how to encourage accreditors to focus more on student learning outcomes and less on input factors would be a valuable outcome of this workshop.

Individual participants discussed how institutions should communicate about educational quality to the general public and to policymakers. Participants in one small-group discussion highlighted what they determined to be the need for institutions to identify the key proficiencies that students acquire through their undergraduate experiences (both those that are general in nature and those specific to a discipline), ensure that these are clearly stated, and clearly connect them to the outcomes of an individual student. Scott Ralls (Northern Virginia Community College) stressed the importance of communicating in a focused way to each of the many different "publics." He encouraged institutions not to assume that posting a Collegiate Learning Assessment (CLA) score on their website is contributing to either transparency or improved decision-making on the part of students, parents, or employers. Elsa Núñez (Eastern Connecticut University) believes that how an institution changes over time, adapting to the

realities of the student population and the external environment, is important to "creating a story the public can understand."

Cliff Adelman (Institute for Higher Education Policy) suggested that the U.S. Department of Education Baccalaureate and Beyond Longitudinal Study[1] be dramatically expanded by adding questions about students' activity such as civic participation and cultural participation and by administering it to 500,000 people.

Several participants suggested that co-curricular transcripts and diploma supplements can communicate students' experiences and abilities beyond the classroom. Andy MacCracken (National Campus Leadership Council) advocated for co-curricular transcripts as a way to bridge the divide perceived by students between what is valued in the classroom versus out of the classroom. This transcript would accompany the traditional transcript and list out-of-classroom experiences. One challenge, he noted, is in determining how the university would validate the student's transcript as an indicator of success in the workplace. Adelman described how diploma supplements might work, noting similarities to efforts in Europe that focus on the institution rather than the student. A diploma supplement might include a description of the senior project (certified by the chief academic officer); a summary of how the student contributed to the university or its surrounding communities—local, regional, national, international (certified by the dean of students); a summary of the student's foreign language skills; and a listing of completed courses.

PUBLISHING BEST PRACTICES AND GUIDELINES

Participants shared several ideas about publications to assist institutions in improving the quality of the undergraduate education. Jennifer Engle (Bill & Melinda Gates Foundation) suggested production of a report oriented toward higher education's consumers (i.e., parents and students) to inform their decision-making.

Individual participants suggested that a report of best practices would help faculty, deans, and other administrators to leverage the abundant knowledge that already exists about high-quality undergraduate education, as has been done on other subjects.[2] Marco Molinaro (University of California, Davis) suggested a

[1] See https://nces.ed.gov/surveys/b&b/.

[2] For example, National Academies of Sciences, Engineering, and Medicine. (2016). *Promising Practices for Strengthening the Regional STEM Workforce Development Ecosystem*. Washington, DC: The National Academies Press. doi:10.17226/21894; National Research Council. (2011). *Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics Education: Summary of Two Workshops*. Washington, DC: The National Academies Press. doi:10.17226/13099; and National Academies of Sciences, Engineering, and Medicine. (2016). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. Washington, DC: The National Academies Press. doi:10.17226/21739.

periodical volume that could be reviewed regularly and possibly tied to the College Scorecard as long as the Scorecard is focused on student learning. He cited the example of the Academies' series on disciplinary-based educational research, in which a first report outlined the evidence while a second report functioned as a practitioner's guide.[3]

STRATEGIES FOR IMPROVING QUALITY (EXAMPLES)

Several participants described efforts by specific institutions to improve the quality of their undergraduate education. James Kvaal (the White House) described the positive results obtained by City University of New York's (CUNY's) Accelerated Study in Associate Programs (ASAP) through combining counseling, full-time enrollment, tuition waivers, and active remediation.[4] Graduation rates of CUNY ASAP cohorts are approximately double those of non-ASAP comparison group students. Although not inexpensive, the program costs less per graduate than traditional higher education. Kvaal cited work on boosting success rates in developmental education as also very important. He described the cognitive tutor model that employs technology to ensure that students are progressing. The data generated by this approach allows the instructor to randomize methods for teaching different concepts, observe what works best for student learning, and build in continuous improvement.

Sally Johnstone (Western Governors University) outlined the approach used by Western Governors University, an institution described by Adelman as having a clear-cut focus and mission. Many of its students, almost 70 percent of whom are Pell-eligible, did not succeed in traditional models of higher education. To define the appropriate competencies or proficiencies and the right learning outcomes, Western Governors University engages employers, external academics well-known in their field, and instructional designers. The university employs a system for continuous measurement of learning outcomes online, looking at what learning resources students are using and not using, what is predictive of student success, and where students are having problems. Through these processes, she described how the university can identify successful pathways or clusters of pathways through course materials and uses that information to improve student success.

Núñez described the decade-long process of quality improvement at Eastern Connecticut University, Connecticut's only public liberal arts university, and

[3] National Research Council. (2012). *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. Washington, DC: The National Academies Press. doi:10.17226/13362. and National Research Council. (2015). *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. Washington, DC: The National Academies Press. doi:10.17226/18687.

[4] See http://www1.cuny.edu/sites/asap/.

how she learned what was needed to support the university's claims of quality. The faculty's philosophy, as well as hers, is one of engaged learning—that students learn best when they can apply their classroom learning in internships for credit, co-ops, or undergraduate research or community-based projects. The incentive for the faculty is release time, which she acknowledged is costly.

Since its efforts directed at engaged learning, the university has improved on 70 percent of the National Survey of Student Engagement's (NSSE's) 42 measures. It scores higher than Research I institutions[5] on measures such as engaging in internships and learning communities and conducting undergraduate research with the faculty. However, because NSSE is self-reported data, Eastern Connecticut University administers the Collegiate Learning Assessment (CLA) to groups of freshmen and seniors to collect externally validated evidence of quality. Núñez highlighted faculty participation in the Current VALUE Project: Multi-State Collaborative and SHEEO (the State Higher Education Executive Officers' association).[6] A pilot project begun in Massachusetts and involving nine states, the program is currently housed at the Association of American Colleges and Universities (AAC&U). The project measures broad intellectual skills as well as integrated applied learning and evaluates actual artifacts of students' academic course work. Faculty have embraced the project because its rubric measures the artifacts from their courses. The rubric is evaluated by faculty from out-of-state institutions, which lends strong support to her funding requests to the state legislature and the university's board. She believes that this project is the first in the nation to use a nationally recognized rubric and an external review process to assess student's actual products.

PLACING THE QUALITY DISCUSSION IN AN INTERNATIONAL CONTEXT

Ellen Hazelkorn (Dublin Institute of Technology), policy advisor to the Irish government, noted that the discussion about quality in the United States resembles the one that is happening internationally. She cited efforts to assess quality and rank institutions in the United Kingdom, Australia (myUni), and the European Union (U-Multirank). A commonality among these efforts, in her view, is an increasing focus on the quality of teaching. She sees an international trend to move beyond issues of improving quality to measuring and comparing performance and productivity, and linking this information to resource allocation. Institutions in other countries are often compared on specific objectives, generally national social objectives.

[5] There are 115 Research I: Doctoral Universities–Highest research activity, as measured by research expenditures, number of research doctorates awarded, number of research-focused faculty, and other factors, as of the 2015 Classification update. More information is available at http://carnegieclassifications.iu.edu/.

[6] See https://aacu.org/value/msc.

David Dill (University of North Carolina) mentioned "quality enhancement," a term used widely outside of the United States to describe a regulatory procedure for evaluating the quality assurance of college and universities, particularly in England, Hong Kong, and several Scandinavian countries. He noted that what distinguishes quality enhancement regulatory procedures from U.S. accreditation is the focus on the internal processes and governance by which colleges and universities ensure, measure, and design academic courses and maintain the quality of the academic standards.

Andrew Crews, an architect, discussed the global marketplace for academic credentials. In his work in higher education and accreditation for architecture he has seen a surge in a number of non-U.S. institutions seeking accreditation as a mark of quality outside the United States. Conversely, as the leader of human capital strategy for a global architecture and design firm, he has also observed that students and people wishing to enter his field are "recognizing that a non-U.S. credential is frequently faster and cheaper." He therefore encouraged the group to keep the global marketplace for academic credentials in mind during discussions of quality in undergraduate education.

LEARNING FROM OTHER INDUSTRIES

Howell suggested that the discussion about quality in education could benefit from the literature from other industries such as in health care. Industries in which consumers make high stakes choices, such as choosing hospital and doctors, are also grappling with defining, measuring, and communicating quality to end users. She cited some researchers at the University of California, Davis, who are looking at measures of hospital quality and modeling it in the context of higher education.

Another participant referred to an interesting model for measuring quality in health care that grew out of the Institute of Medicine's (IOM's) National Roundtable on Health Care Quality, which was convened in 1995. That roundtable of representatives from industry, academia, government, and other stakeholder groups was tasked with identifying a set of health outcomes to be measured.[7] The IOM group determined the most reliable hospital- and doctor-level measures of quality and then curated the methodological development of risk-adjusted measures. These measures are now utilized in the pay-for-success matrix behind some reimbursement policies and used in the Affordable Care Act.

During a luncheon keynote speech, Nigel Croft, chairman of the International Standards Organization's (ISO) Committee for Quality Systems, described concepts of quality across fields as they have evolved and been applied around

[7] A report issued by the National Roundtable on Health Care Quality is: Institute of Medicine. (1998). *Statement on Quality of Care.* Washington, DC: The National Academies Press. doi:10.17226/9439.

the world. The ISO's formal definition of quality is "the degree to which a set of inherent characteristics of an object fulfills requirements." He highlighted common challenges between industry's and education's needs to define and achieve quality.

Croft noted that such characteristics can have very different dimensions that "may or may not fulfill the needs and expectations of the different interested parties." In addition, different products have different characteristics, and institutions of higher education have multiple relevant interested parties, from students to employers. A common challenge for industry and higher education is how to substantiate claims for the products and services offered.

As an example of addressing the quality needs of a broad set of interested parties, Croft described the work of the technical subcommittee responsible for the ISO 9000 quality management system standard.[8] The subcommittee of representatives from government, education, and industry in 83 countries focused on finding high-level, principles-based commonalities. From those high-level principles, the subcommittee drilled down to examine quality in specific contexts from business-to-consumer situations to more abstract contexts such as health care in which the transaction is not linear.

He emphasized the importance of two contexts: (1) the organization, that is, "What are the external and the internal factors that allow that organization to achieve its objectives?" and (2) the interested parties. He observed an absence of students at the workshop and encouraged participants to account for their needs and expectations.

Croft told participants that ISO is developing a new standard, ISO 21001, on management systems in educational organizations and invited their participation in this effort.[9] The discussion is currently at a high level and generic, involving all of the ISO member bodies. The goal is to create a common, high-level structure from which people working in different educational contexts can drill down.

CONTINUING TO DO MORE WITH LESS

Gabriela Weaver (University of Massachusetts at Amherst) reminded the audience that faculty, and institutions themselves, are being asked to do more with less. She described the "big gap between knowing what [quality] is and having the ability to provide it, when your institution is dealing with decreased levels of state funding, federal funding, and research money and increased levels of accountability on metrics that it hasn't been held accountable for in the past." "The question of quality," she said, "has to intersect with the question of institutions having the ability to support the faculty to do those things." Roy Swift (Workcred) cautioned against leaving smaller institutions out of the

[8] See http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm.
[9] See http://www.iso.org/iso/catalogue_detail.htm?csnumber=66266.

quality discussion; institutions lacking abundant resources may need assistance in collecting, storing, and analyzing the data.

CONVENING A WIDE RANGE OF INTERESTED PARTIES TO DEVELOP A FRAMEWORK FOR QUALITY

Individual participants described a desire to bring the evidence to bear on the problem of quality in undergraduate education to help a wide range of stakeholders, including institutional actors, policymakers, and business leaders, to understand why certain steps must be taken. Participants in one small-group discussions suggested that the Academies could take a lead role in the quality discussion as convener of the process. In that capacity, Jillian Kinzie (National Survey of Student Engagement and National Institute for Learning Outcomes Assessment) urged that inclusive excellence be an integral part of the Academies' efforts and that the Academies might bring together different audiences to address this issue more directly.

Other participants suggested that the Academies would serve the nation well by establishing a framework on which quality models can be developed and that the Academies are well positioned to serve the role as integrator and synthesizer of data. Hazelkorn asked, "Is there a core set of principles to guide the development of the framework for measuring quality?"

Some participants identified the driver of quality metrics to be one important component of the framework. Some institutions may require assistance in following the thread from their missions to the detailed data that must to be collected to assess whether they are fulfilling those missions.

Regarding the potential role of accrediting organizations in improving quality in undergraduate education, a participant suggested that the Academies could help to shape reform how accreditation is conducted.

Appendixes

Appendix A
Workshop Agenda

QUALITY IN THE UNDERGRADUATE EXPERIENCE

What Is It? How Should It Be Measured? Who Decides?
A Workshop Hosted by the
The National Academies of Sciences, Engineering, and Medicine
December 14-15, 2015
National Academy of Sciences Building,
2101 Constitution Ave. NW, Washington, DC

Sponsored by The Lumina Foundation

Day 1—December 14, 2015

3:00-3:15 Welcome: Paul Courant, Committee Chair

3:15-5:30 "What Is Quality? How Should It Be Measured? Who Should Decide?" Moderated panel discussion involving project committee members (3:15-4:15) followed by small group discussions (4:15-5:30); each discussion facilitated by a committee member. Panelists and group discussion facilitators are:

Paul Courant, Committee Chair, Professor of Public Policy, Professor of Economics, Professor of Information, and Faculty Associate in the Institute for Social Research at the University of Michigan.
Ellen Hazelkorn, Policy Advisor to the Higher Education Authority (Ireland), Emeritus Professor and Director, Higher Education Policy Research Unit (HEPRU), Dublin Institute of Technology
Paul LeBlanc, President, Southern New Hampshire University
Alexander McCormick, Associate Professor of Educational Leadership and Policy Studies, Indiana University Bloomington; Director of the National Survey of Student Engagement (NSSE)

Marco Molinaro, Assistance Vice Provost for Educational Effectiveness, University of California, Davis
Roy Swift, Executive Director, Workcred, an affiliate of the American National Standards Institute
Jordan Matsudaira, Assistant Professor, Policy Analysis and Management, Cornell University; Former Chief Economist, White House Council of Economic Advisers.
Moderator: Jillian Kinzie, Associate Director, Center for Postsecondary Research and National Survey of Student Engagement (NSSE), Indiana University

5:30-7:30 Reception and dinner. Keynote Address by Carol Geary Schneider, President, Association of American Colleges and Universities

Day 2—December 15, 2015

9:00-10:30 Panel: How Do the "Consumers" of Higher Education See Quality? Perspectives from students, the federal government, employers, and foundations.

Andy MacCracken, Executive Director, National Campus Leadership Council
Emily Slack, Professional Staff Member, Education & Labor Committee, U.S. House of Representatives
James Kvaal, Deputy Director, Domestic Policy Council, The White House
Jennifer Engle, Senior Program Officer, Bill & Melinda Gates Foundation
Aprille Ericsson, SBIR/STTR Program Manager, Innovative Technology Partnerships Office, NASA Goddard Space Flight Center
Moderator: Paul LeBlanc, President, Southern New Hampshire University

10:45-12:15 Panel: What Is Quality? Perspectives from a 4-year university, a 2-year college, an online institution, and a web development immersive school

Elsa Núñez, President, Eastern Connecticut State University
Scott Ralls, President, Northern Virginia Community College
Sally Johnstone, Vice President for Academic Advancement, Western Governors University
Stanley Ikenberry, President Emeritus of the University of Illinois; Former President of the American Council on Education; a Co-Principal Investigator of the National Institute for Learning Outcomes Assessment (NILOA)
Adam Enbar, President, Flatiron School
Moderator: Alexander McCormick, Associate Professor of Educational Leadership and Policy Studies, Indiana University Bloomington; Director of the National Survey of Student Engagement (NSSE)

12:15-1:00 Lunch. Remarks by Nigel Croft, Chairman of the ISO Technical Committee for Quality Systems—"A Global Perspective of Quality"

1:00-2:15 Small Group Table Discussions. Key Questions to Consider:

What actions are required in the next year or two to move us from current models (e.g., VALUE Rubrics, PULSE, and DQP) that are being implemented on an ad hoc basis to a system of quality measurement whereby a group of like institutions adopts a standard set of indicators and reports their results?

Now that the College Scorecard has been released, what further steps should the federal government (and, possibly, state governments) take to improve public information about the quality of undergraduate institutions? Are there improvements to the College Scorecard that are feasible and desirable in the near term? If so, who should be responsible for implementing them? What structures need to be put in place to assure that the College Scorecard is well-curated and that it can improve over time?

Should a group be assembled to create a core set of principles to guide the development of a general framework for measuring quality in undergraduate education--one that can be adopted by nearly any type of institution, e.g., 4-year university, 2-year college, online institution, "boot camp," etc.? If so, who should be involved in that process, who should lead it, and who should fund it?

What might be the most appropriate role, if any, for the Academies? Could it, for example, serve an integration and synthesis role, bringing together and leveraging the good work that is under way (including the DQP, VALUE, VSA, and perhaps other emerging programs)? Might it also seek to broaden the emphasis from defining competencies and outcomes to working out the quite thorny assessment and consumer information components?

2:15-3:00 Report outs from small groups, and full group discussion of short-term next steps

3:00-3:15 Concluding Remarks: Paul Courant, Committee Chair

Appendix B
Quality in the Undergraduate Experience—A Discussion Document

By the Planning Committee for Quality Higher Education:
What Does It Mean, How Is It Measured, and Who Decides? A Workshop
December 4, 2015

This framing document is intended to stimulate thinking about key issues related to the quality of undergraduate education. It elaborates on the five themes identified in the workshop invitation: the measurement of student learning; qualitative factors often cited as important outcomes of undergraduate education; the importance and challenges of assessment; federal policy implications of assessing quality; and the importance of context with regard to institution type, learning environments, and student goals. We do not intend to address all facets of the quality challenge, nor do we mean to suggest that this is the only way to unpack it. Indeed, we are interested in identifying and filling gaps and hearing other perspectives. We hope advance thinking about these (and other) issues will help stimulate a rich and generative discussion at the December 14-15 workshop hosted by the National Academies.

MEASURES OF STUDENT LEARNING

1. *Much of the focus on "quality" in undergraduate education has been on input factors or a variety of outcome measures: reputation, entrance examination scores and admissions selectivity, financial resources, graduation rates, graduates' employment and earnings, and other attributes that can easily be measured but that say little about student learning—that is, the acquisition of important and relevant market-valued knowledge, skills and abilities (KSAs) and the ability to apply those KSAs in real-world settings. How can we change that? Are there approaches and metrics that can accurately speak to student learning?*

When reflecting on measures of student learning, one can focus on the specific course experience, formative in nature, or on the overall academic

experience, a more summative departmental or institutional perspective. In this section we choose to focus on the latter. Traditional measures such as graduation, retention, campus resources, graduate employment and more, while relatively straightforward to measure and of value, are limited in their ability to provide evidence of student learning. Broader and deeper methods and approaches are needed to help departments and institutions clearly define the value they provide their graduates. Here we highlight three specific approaches, noting that there are many others. We discuss the Association of American Colleges and Universities (AAC&U) VALUE Rubrics, the PULSE Vision and Change Rubrics, and the Degree Qualifications Profile (DQP). These three approaches, outlined below, can help us think about the dimensions of the problem and how a substantial number of institutions are moving forward in an iterative journey of self-exploration to define the value they bring to the student learning dimension.

The AAC&U sponsored VALUE Rubrics (Valid Assessment of Learning in Undergraduate Education) provide tools to help assess students' work produced across the students' varied learning pathways and institutions, "to determine whether and how well they are progressing toward graduation-level achievement in learning outcomes that both employers and faculty consider essential." Dimensions considered by VALUE include intellectual and practical skills (inquiry and analysis, critical thinking, creative thinking, written communication, oral communication, reading, quantitative literacy, information literacy, teamwork, and problem solving); personal and social responsibility (civic engagement, intercultural knowledge and competence, ethical reasoning, foundations and skills for lifelong learning, global learning); and integrative learning. Rubrics were developed by faculty and other professionals from more than 100 institutions. According to the project website, "Each rubric was developed from the most frequently identified characteristics or criteria of learning for each of the 16 learning outcomes."[1] A recent multi-state collaborative is looking at student work from 69 participating 2- and 4-year campuses. In addition, the Voluntary System of Accountability (VSA)[2] includes two VALUE rubrics on the list of assessment tools that participating institutions can use to demonstrate student learning.

The PULSE Vision and Change Rubrics provide a structure for departmental reflection, self-assessment and discussion regarding a host of topics relevant to program transformation. While the focus has been on the life sciences, the process is equally applicable for any STEM field. The current rubrics have criteria immediately applicable to all STEM fields except for the disciplinary core concepts, which are available for life sciences only. The rubrics and suggested activities suggest a process by which faculty and appropriate learning

[1] See http://aacu.org/value/rubrics.
[2] See http://www.voluntarysystem.org.

and technology staff work collaboratively to maximize their collective transformative change.

The Degree Qualifications Profile (DQP) helps frame an institution's mission and overall goals when granting degrees. The DQP identifies five essential areas of learning that should be incorporated in any post-secondary degree, with increasing complexity based on the degree obtained. The five areas are specialized knowledge, broad and integrative knowledge, intellectual skills, applied and collaborative learning, and civic and global learning. Specialized knowledge outlines what students in any specialization should demonstrate with respect to the specialization, or major, with proficiencies within each field determined by each field through a process called "Tuning" to describe the particular concepts, knowledge areas, methods, skills and accomplishments necessary. Broad and integrative knowledge asks that students are able to consolidate and utilize knowledge across multiple areas to discover and explore questions that span multiple fields of study. Intellectual skills are defined as evidence-based reasoning across fields of study and include: analytic inquiry and operations, use of information resources, engaging diverse perspectives, ethical reasoning, quantitative fluency, and communicative fluency. Applied and collaborative learning focuses on how students can utilize what they know to innovate and move beyond classroom level work as individuals and in groups. Civic and global learning refers to student preparation to engage and contribute to political, social, environmental and economic challenges. Overall the DQP asks university stakeholders to engage in a process of ensuring that students are both competent (can demonstrate a certain level of skill in a course/experience) and proficient (summative ability gained through multiple course experiences with commitment to ongoing learning) as relevant to their field of study and level of degree attainment.

These are just three example approaches, by no means meant to be limiting. There is a good deal of other work on 21st century learning skills, including the Next Generation Science Standards and the P21 Framework for 21st Century Learning, among others. Traditional standardized test approaches are also being used to assess more diffuse and generic learning outcomes. For example, in addition to VALUE rubrics, institutions participating in the VSA can demonstrate learning gains using ACT's Collegiate Assessment of Academic Progress, the Council for Aid to Education's Collegiate Learning Assessment, or ETS's Proficiency Profile.

The assessment of learning is complicated by the emergence of new providers and new approaches to provision such as competency-based learning, problem-focused field experiences such as internships, and other programs, Do these change the ways we should approach the definition and assessment of learning? If we say we are assessing "competency" and "proficiency," how do they differ and how does that affect the validity of the assessment tools that are chosen?

Regardless of the specific approach, there is increasing consensus that student learning needs to be at the core of our thinking about educational quality as educators, administrators, taxpayers and global citizens.

QUALITATIVE FACTORS

2. *"Not everything that can be counted counts, and not everything that counts can be counted." – attributed to Albert Einstein. How do we begin to define, identify and measure the qualitative elements of a high-quality undergraduate education?*

A high-quality undergraduate education involves more than the accumulation of factual knowledge and intellectual skills. There is wide agreement that it should also inculcate a range of diffuse skills and habits of mind that prepare students for lives of engaged citizenship, intercultural competence, social responsibility, and continued intellectual growth. While some of these capacities and habits are addressed by the VALUE rubrics described above, their measurement defies precise and consensually accepted methods. Indeed, the list itself is subject to debate. Excluding these outcomes from the quality discourse risks marginalizing them at a time when there is increasing recognition of the importance of so-called "soft skills." But if we are to tackle these behavioral traits that go beyond traditional declarative and procedural knowledge, we have to become more knowledgeable about how individuals learn these "behaviors" and about how to measure such learning.

How might these important outcomes be incorporated into the quality discourse? One approach is to gather information about the activities and experiences of alumni at designated time points (say, one, five, and ten years after graduation). But graduates don't live their lives in a bubble—they are exposed to other influences after graduation through employment, further education, family formation, and so on. This introduces substantial inferential challenges: Can colleges and universities properly claim credit for their graduates' achievements five, ten, or more years after graduation? Attributing institutional responsibility for alumni outcomes may require unrealistic assumptions—or alternatively, sophisticated analyses that attempt to rule out confounding factors (and that undermine simplicity and transparency from a consumer information perspective).

Another important qualitative dimension involves the student population itself. As the U.S. progresses toward a pluralistic "majority minority" society, can any institution be deemed high quality if (a) it does not serve a student population that is reflective of the broader population, and (b) distinctive educational opportunities and salutary outcomes are not enjoyed across student populations? How should we measure an institution's achievement of these important goals beyond coarse measures of compositional diversity? Should we try to measure the impact of the diversity on student interaction and student learning?

An additional dimension for consideration is whether there is a role for other interested parties, such as those who work with an institution's graduates (employers and graduate program faculty) to provide information about graduates' level of preparation. We often hear of employer complaints regarding

student preparation, and more rigorously collecting and analyzing this information could help us understand the specific needs that are unmet. In using such information it is important to be mindful of when in the graduates' work and life the information is acquired. It is perfectly plausible to conclude that learning outcomes that prepare students well for jobs right out of college are not the best preparation over longer time periods, and vice versa.

ASSESSMENT

3. *In a system strongly guided by norms of professional judgment, peer review, and evidentiary support, quality is closely linked to processes of diagnosis and improvement. For an institution to be judged high quality, should criteria include the presence of a rigorous program of outcomes assessments and continuous improvements that are found in other industries? What should such a program look like, and who should judge its adequacy? What should be the role of faculty, employers, governments, and students/parents in establishing desired outcomes?*

Assessment is a dish best served formatively—asking what went well, what did not and where can improvements be made. Unfortunately assessments are often used summatively to judge, sort and separate, not to promote growth, especially for students. Individual faculty, curricular committees, and departments are often in a similarly poor situation where assessments are either non-existent, non-actionable or of poor quality. In many of our institutions, adjunct faculty are hired or fired based on teaching evaluations that emphasize student satisfaction with instructor "performance" rather than clear measures of student learning and capability. While appointment, promotion, and tenure decisions for tenure-line faculty involve a wider range of performance criteria, the teaching quality component typically relies on the same evaluations. So what can be done? One potential approach at the course level calls for: 1) clear, measurable learning goals that are agreed upon between instructors of the same course and communicated to students, 2) multiple forms of feedback, including low (quizzes, graded or ungraded assignments) to high (midterm and final exams) stakes assessments, that help students gauge their learning and guide their improvement, 3) instructional approaches appropriate to the student population and the learning outcomes to be achieved and based on evidence of educational effectiveness, 4) agreed-upon approaches for instructors to reflect on their instruction and their students' learning, and to engage in a process of continuous improvement.

Instructional quality can also be assessed at the departmental level, which would involve an expansion of assessment of individual courses. The individual course level elements considered can be summed up for all the courses offered by a department to gauge variation and engagement with evidence-based teaching practices at the level of a course series/sequence and through the entire degree program. Additional elements worthy of evaluation can include TA

instructional preparation and assessment, articulation between courses within a department and with related courses in other departments, and other factors that bear on student learning.

At the end of the day, if learning is regularly measured in a systematic way, analyzed, acted upon, and the cycle iterated with an emphasis on increasing student learning and capabilities, chances are the system is working. The "system" can be considered an individual instructor and her/his instructional team, the instructional teams that teach a given course, the teams involved in providing a course series, all the way to the department's educational mission.

One can easily see how this approach can be summed up for a collection of departments, or a college, and these groups can be brought together to reflect an entire institution. What is deemed "fair," "standard," or "exceptional" in terms of overall quality and outcomes is open to debate but as long as there is continuous improvement guided by student outcomes within and outside of the institution, as well as evidence-based instructional practices, chances are high that quality is present. Communicating to stakeholders the evidence that an institution is actively engaged in measuring its aspirations against its actions and is using that information for a process of improvement is likely to be very powerful information for all relevant stakeholders as long as it is sufficiently detailed and transparent. Indeed, one can argue that the quality discourse would benefit from a greater focus on the rigor and consistency of assessment and improvement processes in place than on the specific outcomes (scores) on any given assessment. But this may challenge the system to create quality performance assessments to demonstrate that students are indeed acquiring the knowledge and skills that a department, program or college has promised to deliver.

In addition, numerous approaches exist to help institutions gauge their overall impact on students, including student surveys that assess the learning experience (NSSE, CCSSE, UCUES/SERU) and the use of electronic portfolios to document student progress. These approaches help provide the big picture and can help to identify areas of strength and areas in need of improvement. It may also be helpful to look at other industries, such as manufacturing and health care, which have extensive experience in continuous quality improvement to produce better products and outcomes with less "recall" or "error."

UNDERSTANDING QUALITY FROM A FEDERAL POLICY PERSPECTIVE

4. *For understandable reasons, federal policymakers concerned with quality focus on measurable quantities (e.g., completion rates of Pell recipients, employment and starting salaries of graduates). This is important but insufficient and can have perverse consequences.*

The recently updated College Scorecard represents an attempt by the federal government to improve the consumer information available to students planning

to attend a 2- or 4-year institution.[3] It displays summary information drawn from existing federal data sources on program offerings; student body composition (achievement test scores, enrollment status, race/ethnicity, and percent receiving Pell grants as an indicator of socioeconomic diversity); net cost of attendance (broken down by income bracket); financial aid and debt (limited to college completers); retention and graduation rates; and earnings ten years after entry (limited to federal aid recipients; including all who ever attended the institution, regardless of duration, whether they graduated, or where they graduated from). Where relevant, displays compare an institution's result to the national average. Several states have undertaken similar transparency-focused efforts. For example, Indiana produces "College Completion" and "Return on Investment" reports comparing results for public institutions in the state.

The indicators used in these efforts suffer from limitations, but the more important question for the present purpose is: What do they tell us about the quality of education delivered? A college's outcomes are highly influenced by who attends and what they study. Students are not randomly assigned to colleges and universities, so to a considerable extent institution-level variation in outcomes such as retention, graduation, and earnings reflect differences in the characteristics of those who enroll. Employment and earnings are affected by major field of study, so the mix of majors produced by an institution also accounts for differences in these outcomes. But those who turn to resources like these for authoritative consumer information may not be sensitive to these nuances, leading to improper inferences about institutional effectiveness and "quality." Thus, one aspect of any effort to help students choose among institutions should be the implementation of programs to help students and their families to decode and interpret information about institutions and programs.

The federal government is involved in quality assessment in another important way: by recognizing accreditation agencies that employ a peer-review process to ensure that institutions satisfy a designated set of quality standards. Accreditation has been criticized for lack of transparency, failure to motivate optimal performance or adequately penalize poor performance, and inadequate attention to student learning outcomes. Nevertheless, it remains the nation's formal quality assurance process for higher education. It attends to facets of institutional performance that do not lend themselves to easy measurement in a comparative framework, and the use of external peer reviewers is a common approach to performance assessment in professional domains that rely on expert judgment to navigate complexity and specialized knowledge. Can the accreditation process be modified to provide more useful consumer information and to provide federal policymakers with better information about institutional performance, while continuing to serve its core purpose? This is a key question that needs significant attention by organizations and individuals who seek

[3] See https://collegescorecard.ed.gov/.

significant federal policy influences on improving the quality of undergraduate education.

THE IMPORTANCE OF CONTEXT

5. *Does the meaning of quality depend on what students and other payers "hire" colleges and universities to do for them?*

Higher education—whether in a community or technical college, a private liberal arts college, or a public research university—represents a significant investment by students (both direct expenditure and opportunity cost) and by taxpayers (whether direct institutional subsidy or student financial aid). Students and society deserve to be confident that those investments pay off.

But the diversity of U.S. higher education—encompassing both institution types and students' reasons for attending—means there can be no "one size fits all" solution to the quality question. Across this variety, however, we can and should ask whether colleges and universities provide their students access to quality experiences and outcomes that correspond to their goals. At its most basic, this means asking whether colleges and universities are delivering what students and other payers expect—or as some have put it, what they are hired to do.

This raises the question of the proper level of analysis for the quality question. Where do students experience educational quality? An increasing share of the student population attends more than one institution on the way to a credential. More research is needed to understand the reasons for attending multiple institutions. Research examining the student experience and student outcomes finds far more variability between students within institutions than it does between institutions. These facts call into question whether the institution is the right focus for the quality discussion. Should we look higher, at the level of a state's collection of public and private providers? Or should we be assessing quality at the program level, which is where students engage most directly with educational experiences? Or should we look at all three levels, and be careful to describe what it is that we are discussing and how the levels interact, so that assessment and information can be well used?

CONCLUSION

Our purpose in this document is to raise a number of the key issues that have surfaced over the past several years among educators, policymakers and others who are seeking to advance both our thinking and our institutional and public policies around defining and measuring quality in undergraduate education. We offer these concluding questions to help guide the discussions during the December 14-15, 2015, Academies workshop. These represent only a few of the core questions that need attention, and are intended as a starting point for

action among the various stakeholders, particularly during the next 24-36 months:

1. What actions are required in the next year or two to move us from current models (e.g., VALUE Rubrics, PULSE, and DQP) that are being implemented on an *ad hoc* basis to a *system* of quality measurement whereby a group of like institutions adopts a standard set of indicators and reports their results, keeping in mind the work of the Voluntary System of Accountability and the related community college effort, the Voluntary Framework of Accountability?[4] What are the next steps in the process of implementing such a system, even on a pilot basis?
2. Now that the College Scorecard has been released, what further steps should the federal government (and, possibly, state governments) take to improve public information about the quality of undergraduate institutions? Are there improvements to the College Scorecard that are feasible and desirable in the near term? If so, who should be responsible for implementing them? What structures need to be put in place to assure that the College Scorecard is well-curated and that it can improve over time?
3. Can—and should—a group be assembled to create a core set of principles to guide the development of a general framework for measuring quality in undergraduate education—one that can be adopted by nearly any type of institution, e.g., 4-year university, 2-year college, online institution, "boot camp," etc.? If so, who should be involved in that process, who should lead it, and who should fund it? How could such an entity build on many of the existing rubrics and tools that have recently been developed?
4. What might be the most appropriate role, if any, for the Academies? Could it, for example, serve an *integration* and *synthesis* role, bringing together and leveraging the good work that is under way (including the DQP, VALUE, VSA, and perhaps other emerging programs)? Might it also seek to broaden the emphasis from defining competencies and outcomes to working out the quite thorny assessment and consumer information components?

[4] See http://vfa.aacc.nche.edu/.

Appendix C
Defining and Measuring Institutional Quality in Higher Education

Jordan Matsudaira Cornell University
November 2015

"Is college worth it?" In the past several years nearly every major news outlet in the United States has run a story posing some variation of the question.[1] Rising tuition prices and student debt levels have increased the public's concern over whether college investments are worth the money, and recent revelations of fraud and deceptive practices by large college chains have shed light on the reality that not all universities contribute positively to students' success, and fueled calls for greater accountability for institutions receiving public funds. It may seem paradoxical that skepticism over the quality and value of postsecondary education has intensified in recent years, as the college earnings premium has risen over the past several decades and is currently near record levels.[2] But while a deep body of research has answered the value proposition "is college worth it?" with a resounding "yes" *on average*—that is, *on average* the return to a college degree is not just positive but very high relative to other investments a person might make—this is a comforting answer to the wrong

[1] Among many others, see http://www.nytimes.com/2014/05/27/upshot/is-college-worth-it-clearly-new-data-say.html; http://www.washingtonpost.com/news/wonkblog/wp/2015/01/30/college-is-worth-it-if-you-graduate-on-time/; http://www.economist.com/news/united-states/21600131-too-many-degrees-are-waste-money-return-higher-education-would-be-much-better; http://www.forbes.com/sites/michakaufman/2015/03/20/is-college-still-worth-it/; http://content.time.com/time/interactive/0,31813,2072670,00.html; and http://www.newyorker.com/magazine/2015/09/07/college-calculus.

[2] The college earnings premium is the ratio of the average earnings of college graduates to the average earnings of those with only a high-school degree. See Autor (2014) and Baum (2014) for recent tabulations. The college earnings premium rose from the late 1970s to about 2000, but has been relatively stable since.

question. For families choosing a college and for policymakers targeting support or sanctions to poorly performing institutions, the relevant question is whether a *particular* college is "worth it," or which of a set of colleges is most "worth it."

Unfortunately, there are currently few institution-specific indicators of quality suitable either for informing college choice or for use in policies aimed at improving accountability for institutions of higher education. In part, the lack of institution-specific quality information may be because the notion that attending college is anything but "the surest ticket to the middle class"[3] has only recently been called into question. It is telling, for example, that the primary federal measure of institution "quality"—the graduation rate—measures only whether students complete their course of study, effectively taking for granted the benefits associated with doing so. In recent years a burgeoning private industry of college rankings and some state higher education administrative database tools have made richer information available about college attributes and performance. However, a) too little of this information is currently based on student outcomes—which I argue below is essential for measuring quality—is available; b) *causal effects* of institutions on outcomes inform almost none of these measures; and c) many data elements are not comparable across institutions due to different conceptual definitions and populations covered in different data sources.

There are reasons for the limited progress in developing college quality measures. Colleges and the students they serve have myriad and diverse goals, and many of these are intangible and not readily subject to measurement or quantification. But this is changing: new administrative data is increasingly available through the efforts of federal and state governments, private data collections, and institutional consortia that link various student outcomes to the educational institutions they attended. It is thus an opportune time to attempt to take stock of the current landscape of quality information available, and to begin to outline a research agenda to fill in the considerable gaps that exist in information to guide both college choice and public policies.

This paper aims to provide background for thinking about such an agenda. Below, I first suggest a framework for defining institutional quality that acknowledges both that quality is multidimensional and can differ for students at the same institution. Based on this definition, I discuss some of the properties relevant for evaluating the utility of different quality indicators. Next, I provide a brief overview of research, primarily in the field of economics, on college quality. Finally, I conclude with a sketch of some priority areas that future research should address to build better indicators and discuss some of the associated challenges. As I clarify below, I focus primarily on the need and challenges of developing better quality indicators to guide college choice and to

[3] President Obama, in remarks announcing a proposal to make community college free. See https://www.whitehouse.gov/the-press-office/2015/09/12/weekly-address-new-college-scorecard.

inform accountability efforts, and thus consider primarily quality information that could be produced consistently for broad sets of institutions. Institutions may also benefit from better quality information to guide benchmarking and improvement efforts, but I comment only briefly on quality metrics and issues for this purpose. Additionally, in this paper I focus on institutional quality measures, but the framework here is general and could in principle be applied to smaller units within institutions, such as programs of study or even courses. Whether measuring quality at a more disaggregated level is desirable or feasible, however, deserves further discussion, as I discuss below.

DEFINING QUALITY IN HIGHER EDUCATION

In a 2001 report, *Crossing the Quality Chasm,* the Institute of Medicine (IOM) proposed a definition for health care quality that, suitably adapted, provides a useful starting point for defining quality in higher education: "The degree to which education services increase the likelihood of desired education outcomes." The heart of this formulation is that quality is defined in terms of the causal impact that exposure to some educational experience (e.g., attending college A, or studying engineering at college B) has on some set of outcomes that is valued by the student or society more broadly (e.g., deepened knowledge of a subject area of interest; higher earnings; a reduced probability of criminal victimization, etc.).[4] The IOM definition of health care quality also included consideration of whether health services "are consistent with current professional knowledge." While process-oriented considerations might be valued *per se* (i.e., regardless of whether they affect student outcomes) in some conceptions of education quality, I ignore such issues below.

The most important feature of this view of college quality is that it is measured based on student (and broader societal) outcomes. Two observations about the current landscape of commonly used quality information follow immediately. First, much of the information about college performance currently available is based on input measures, such as faculty-student ratios, expenditures, or student test-scores. As I discuss below, some of these inputs may indeed have a causal impact on student outcomes and thus be useful proxies for quality. It is nonetheless important to keep inputs and practices (e.g., such as whether institutions provide counseling, link financial aid incentives to performance, etc.) conceptually distinct, and to verify their link to outcomes before accepting their value as quality measures. Second, measures of program completion are not direct measures of quality unless completion is valued as an

[4] Education researchers commonly refer to this causal impact as an institution's "value-added." An important issue for measuring value-added is that the causal effect of attending college A is defined relative to some counterfactual, such as not attending college at all, or perhaps attending some other 'reference' institution. In this discussion, I assume the counterfactual is well posed and well understood (e.g., with quality measured as the effect of attending each institution relative to stopping one's education at a high-school degree).

outcome *per se*. Graduation rates convey important information about the exposure of students in a cohort to an institution's educational offerings, but treating them as quality information is akin to equating the efficacy of a pain medication with the fraction of those prescribed it that take the medicine. It may be that graduation rates convey quality information indirectly, insofar as students more satisfied with the quality of their education may be more likely to complete it, but the existence of this link should not be taken for granted.

Several other aspects of this definition deserve mention. First, college quality is multidimensional since there are myriad outcomes that students and society desire colleges to affect. For example, a recent survey of prospective and newly enrolled students conducted by Harris for New American Foundation asked respondents about the importance of 12 different "reasons to go to college."[5] A majority of students responded that 11 of the 12 options were "important" or "very important" (the top two of four response options). Of the options presented, the top three reasons were all related to labor market outcomes: "to improve my employment opportunities," "to make more money," and "to get a good job" with between 89 and 91 percent of respondents saying those were "important" or "very important." Learning "more about a favorite topic or area of interest" or "more about the world" also rated highly, with 85 and 74 percent of respondents, respectively, saying those were "important" or "very important." Finally, personal development ("to become a better person" and "to improve my self-confidence") and improving the lives of their children also rated as important reasons to attend college. Other surveys, such as the CIRP Freshman Survey, asking why current students chose the college they did reveal similar priorities.[6]

Similarly, the impact of attending a particular college on a given outcome (say, the probability of becoming a doctor) may depend on a student's academic preparation, interests, etc. and so an institution's quality may be heterogeneous across different types of students. For example, some institutions might devote extra resources to serving students lacking college preparation and thus have higher quality—a greater impact on their future earnings, for example—for such students, even if they might have a less positive impact on more affluent students relative to other institutions.[7] There are many student characteristics that might interact with institution quality, but the student's academic preparation, interests, and career goals seem particularly relevant.

---

[5] Details about the methodology and the survey instrument can be found here: http://dev-edcentral.pantheon.io/wp-content/uploads/2015/05/Harry_Poll_Survey_Instrument-FINAL.pdf.

[6] See for example Tables on pages 38 and 41 of Eagan, Kevin, et al. (2014). Los Angeles: Higher Education Research Institute, UCLA. *The American Freshman: National norms fall 2014.* http://www.heri.ucla.edu/monographs/TheAmericanFreshman2014.pdf.

[7] For example, a recent *New York Times* article suggests that Xavier University particularly excels in helping black students to get into medical school and become doctors. See http://www.nytimes.com/2015/09/13/magazine/a-prescription-for-more-black-doctors.html.

With only the considerations mentioned thus far, we already have a complex framework for defining college quality. If we imagine an exhaustive list of J education outcomes and a set of K types of students, then the quality of a particular institution has $J \times K$ dimensions. That is, to fully describe an institution's quality we need to consider the causal impact that attending that institution has on each of the J different outcomes for each of the K different types of students. In many cases, a single summative measure of an institution's quality is called for. It is natural to view such a measure as a weighted average of these $J \times K$ measures of institutional quality. Restating this with just a bit of formalism, we can measure institution i's quality, Δ^i, as

$$\Delta^i = \sum_j \sum_k w_{jk} \delta^i_{jk}$$

where δ^i_{jk} represents the causal impact of attending institution i on outcome j for students of type k, and w_{jk} is the weight assigned to that outcome in constructing the overall measure of quality. These weights will depend on context and the intended use of the quality information, and will generally require judgements about the relative merit of each outcome.

In discussions of quality, issues of affordability and value are sometimes blurred with outcomes so I clarify use of those terms here. The "value" of a college is a function of its quality (i.e., its causal impact on outcomes over a student's life) relative to its price (including opportunity costs like foregone earnings)—that is, what you get relative to what you pay. Since the price of college to students can differ from the true resource cost of providing education, and not all benefits of education accrue to the individual receiving the education, the value of college or any other educational experience to students can differ from the value provided to society.

DESIRABLE PROPERTIES OF QUALITY INFORMATION

The ultimate goal of developing better quality indicators for higher education is to enable better decision-making on the part of prospective students, higher education officials, and policy makers to improve the quality of education that institutions offer, or to guide students to institutions offering better quality. Each of these consumers of quality information—that is, students vs. higher education officials vs. policymakers—is likely to find different sorts of information useful.

Prospective students and their families are looking for information to guide their college choices, by considering which institution will best contribute to attaining their educational goals. The relevant information for students is thus an estimate of the set of causal effects on various outcomes of each institution for students like them (i.e., with the same "type" k). This is a forward looking measure, in the sense that data on past students' outcomes must be used to forecast the outcomes each type of student is likely to experience at each institution. "Good" quality information in this setting is information that

provides an accurate forecast (e.g., low mean-squared forecast error) of the outcomes a student can expect if they attend each institution. It is worth noting the obvious points that quality information is more valuable the more the *outcomes are salient for students,* and the more *consistently quality is defined for all institutions being considered*.

Policy makers, on the other hand, may want to use quality information to build accountability schemes that explicitly tie the receipt of public funds to institutions' performance. Recent state efforts to develop "performance based funding (PBF)" systems are perhaps the most developed initiative of this kind, with articulated formulas linking the distribution of funds to different institutions as a function of credit accumulation, degrees awarded, grant funding received, and job placements among other outcomes. Similarly, eligibility to receive federal Title IV payments is dependent on cohort default rates for all institutions and on average debt-service payments relative to average annual earnings for "gainful employment" programs. These types of information are backward looking, in the sense that rewards (or sanctions) depend on whether an institution's past performance met some quantifiable benchmarks. The information needs of such schemes might overlap with the quality information useful to prospective students, but the main criteria for "good" information is whether it *provides incentives for institutions to improve their quality* (and avoids inducing undesirable behavior changes). Moreover, while savvy consumers might be best served by having a variety of measures corresponding to different dimensions of each institutions' quality to weigh according to their own preferences, accountability schemes generally require an explicit method of combining these measures into a single index of quality—for example, in the extreme, "eligible" or "not eligible" for Title IV participation.

From the standpoint of both prospective students and policymakers, we can enumerate several other desirable properties of quality measures. The core part of our definition of quality is that it is *based on a causal effect of an institution on education outcomes*. If measures of college performance mirror the predetermined characteristics of the students those colleges enroll, then that information would provide a misleading (or "biased") forecast of expected outcomes to prospective students whose characteristics might differ. Moreover, we generally do not want to reward (sanction) institutions for selecting more (less) economically advantaged or academically prepared students who are likely to experience better (worse) outcomes. One reason is that measures that reflect such "selection" of students encourage institutions to alter the people they serve to achieve gains in measured performance, potentially undermining access goals. Just as the outcomes in high-stakes health care accountability schemes are risk-adjusted to mitigate doctors' incentive to avoid sicker patients (i.e., to "cream-skim"), and teacher quality measures in K-12 education take into account the prior achievement levels of the students they serve, quality measures in higher education need similar kinds of adjustments in order to isolate the causal impacts of institutions on students' outcomes.

To illustrate the issue, data from the College Scorecard show that Cornell University's alumni (who received Title IV) have median earnings of nearly $71,000 thousand 10 years after beginning their studies, while SUNY Cortland's former students earn about $45,000. Should a student who has been accepted to both believe his expected difference in earnings is $26,000? Almost surely not. SUNY Cortland students are much more likely to come from low-income families (eligible for Pell grants) and have much lower SAT scores, and the factors behind these disparities also contribute to their lower earnings. The raw outcome differences can thus mislead prospective students about the quality differences between schools—a more useful set of quality measures might be the predicted earnings of a student with average characteristics (e.g., family income, SAT scores, etc.) at each institution from a regression model. This would likely lead to a higher earnings measure for SUNY Cortland and a lower one for Cornell, compressing the difference.

This type of adjustment is controversial: many higher education stakeholders believe that "risk-adjustment" institutionalizes lower standards for subgroups of students likely to have lower outcomes—often low-income and minority students. This is a valid concern that should be considered in developing thoughtful accountability schemes, but it should not prevent the development of quality measures.

A second point is that quality indicators should be *reliable*—in the sense that they are stable from year to year. Analyses by the Council of Economic Advisers (CEA, 2015) suggest institution-level outcomes are highly reliable since most institutions have large numbers of students in the cohorts generating the data for most performance metrics. But smaller institutions or metrics covering programs within institutions can be highly variable if they are based on few students. Researchers have noted that this kind of variability can both be misleading to potential students and cause accountability schemes to focus penalties or rewards on small institutions whose outcome measures are most variable.

Information on student outcomes must balance being relevant with being *timely*. For example, earnings measures based on the initial years after students leave college may not be indicative of their lifetime earnings outcomes, and may be distorted by student's decisions to enter graduate school or pursue other education programs that are likely to improve their long-term earnings. On the other hand, if earnings are measured many years in the future then they will capture institutional quality with a significant lag. If institutions and the students they serve are changing, the resulting measures may thus not provide a good forecast of prospective students' outcomes. Moreover, for accountability purposes if an institution is not performing well we want to be able to have measures capable of detecting problems rapidly, so that resources can be targeted for improvement, or so sanctions can be applied in a timely fashion.

A final dimension to consider is whether the quality indicators *provide guidance to facilitate improvement* efforts. While information on student outcomes is fundamental to understanding institutional quality, value-added

measures can have a black-box character to them. Documenting inputs and processes across institutions hold the promise of pointing to specific areas that institutions might target to improve their students' outcomes.

AN OVERVIEW OF RESEARCH ON COLLEGE QUALITY

In this section I provide a brief overview of the literature, primarily in economics, that has attempted to assess the importance of college quality in determining student outcomes. What progress has been made in this literature developing measures of college quality that have the properties described above? Unfortunately, the answer is relatively little until quite recently, with most attention from economists having been devoted to addressing the methodological challenge of isolating the causal impact of postsecondary institutions on outcomes. Differences in students' academic preparation, family background, career interests, and differences in geographic factors like the strength of local labor markets can all lead to differences in student outcomes in and after college that may have nothing to do with institutional quality. To address these selection biases, the literature has struggled to identify settings that approximate experiments where similar groups of students end up at different institutions, allowing a comparison of their outcomes to shed light on differences in the institutions' quality.

Early research on college quality in economics focused on estimating the effect of particular dimensions of college quality on student outcomes. Using student-level data on college attended, graduation, and measures of earnings from a survey (such as *High School and Beyond*, *College and Beyond,* the *National Longitudinal Survey of Youth*, or the *National Education Longitudinal Study*, etc.), researchers attempted to establish a correlation between proxies for college quality and the students' outcomes. Several proxies have been examined in the literature, including freshman SAT/ACT scores, selectivity rankings (e.g., from *Barron's*), information on price of attendance, average professor salaries, faculty-student ratios, and indices combining combinations of these measures.

The main methodological challenge confronted by this literature is that more affluent students and those with stronger academic backgrounds are likely to sort into more selective (or higher in other dimensions of quality) schools. Since such students are likely to have high graduation rates and labor market outcomes regardless of where they attend college, there is a tendency for more selective institutions to have better average student outcomes simply because they enroll more advantaged students. To disentangle this student "selection effect" from institutional quality effects, most studies in economics have relied on either multiple regression analysis (James, Alsalam, Conaty, and To, 1989; Loury and Garman, 1995; Brewer and Ehrenberg, 1996; Monks, 2000; Long, 2008) or matching techniques (Black and Smith, 2004) to control for differences in the types of students attending different institutions. The particulars vary depending on the information available in the data, but most studies control for some measure of student academic preparation such as student age, gender, race, high

school GPA and/or SAT or ACT scores and family income background. Using one of the richer datasets available, Long (2008) additionally controls for parents' marital status and education levels, family income, number of siblings, a variety of neighborhood characteristics, and an index of the student's high school quality. To the extent that students do not differ in unobservable ways that affect their outcomes conditional on the set of information included in the regression model, the estimates of college quality will be unbiased.

Studies employing versions of this research design have found evidence that attending colleges with higher student SAT scores increases students' graduation rates but has more mixed effects on wages. The pattern of these results is mixed across data sets, measures of quality, empirical strategy, and outcome measure. Moreover, an influential pair of studies by Dale and Krueger (2002, 2011) suggests that the apparent quality effects in the analyses are driven by unobserved differences in students' academic background across schools. Using data from *College and Beyond*, Dale and Krueger are able to control for students' "choice sets"—that is, the set of institutions to which students apply and are accepted—which they argue better controls for unobserved differences in both students' academic background and interests that might affect their future outcomes. Adding such control variables results in estimates of the effect of college quality (measured by student SAT scores) on post-college earnings that are not significantly different from zero for most students, with the exception of black and Hispanic students and students whose parents have lower education levels, for whom the effects remain positive and statistically significant.

The work above generally provides support for the notion that college quality matters, and suggests a candidate set of attributes that may be correlated with institutional quality. Black and Smith (2006) critique this literature, however, pointing out that the methodology used by these studies rarely establishes whether a particular input has a causal impact on student outcomes, or the extent to which overall institution quality affects outcomes. Since these studies tend to use only a small set of college attributes, it is unclear whether studies are measuring the causal impact of a particular college characteristic—such as the faculty student ratio—or the combined impact of other college inputs that might be correlated. And since these inputs are only noisy proxies for overall quality, estimates of the impact of quality will tend to be attenuated due to measurement error

The majority of work attempting to quantitatively measure college quality—in the causal sense used in this paper—estimates average outcomes for students at each institution, adjusted for differences in preexisting characteristics that might affect their outcomes as in the literature described above. Most of this work relies on aggregate institution-level outcomes, like graduation rates or average earnings, and adjusts those measures using aggregate institution-level student characteristics, such as median SAT scores, the fraction of students eligible for Pell grants, etc. In these studies, regression or matching techniques are used to estimate the predicted relationship between student and institution

characteristics and student outcomes, and institutional quality is measured as the difference between their actual and predicted performance. Many college rankings publications use this methodology to estimate regression adjusted graduation rates, including *U.S. News and World Report*, *Washington Monthly*, and *Forbes*, and more recently *Money* and *The Economist*, among other outlets, have used similar methods to estimate adjusted earnings (or "value-added") measures for colleges.

To give a heuristic, if overly simplified, sense for how the technique works, Figure C-1 depicts the relationship between 6-year graduation rates for first-time, full-time students and the percentage of students receiving Pell grants at 4-year institutions based on IPEDS data. As might be expected, there is a strong negative correlation between the family income background of the student body and their graduation rates: for example, institutions where 80 percent of students receive Pell grants have graduation rates that are more than 35 percentage points lower on average than institutions where only 20 percent of students receive Pell grants. Since it would be unfair, and perhaps create incentives to reduce enrollment of lower-income students, to attribute graduation rate difference between these groups of institutions to differences in their quality, regression adjustments are used to "level the playing field." That is, regression analysis is used to predict the graduation rate we would expect given the percentage of Pell recipients the institution has, and institutions are judged based on whether they exceed this "expected" level of performance. For concreteness, the figure highlights two institutions—Tulane University and University of Texas at El Paso (UTEP)—and the solid line in the Figure displays the predicted values from a regression of graduation rates on percentage Pell.[8] Despite having a 37 percentage point higher graduation rate, Tulane has lower measured quality (i.e., the vertical distance from the solid regression line), since the regression suggests that its graduation rate should be even higher than UTEP's, given how advantaged its student body is in comparison.[9]

In practice, researchers and other college ratings publishers adjust outcomes for a wide range of student and college characteristics beyond just the fraction of students receiving Pell grants. The appropriate set of variables to use for adjustment should include all predetermined student characteristics that affect student outcomes, such as academic ability, family income, and demographic measures. Most studies use some permutation of SAT or ACT scores, the admission ratio, or other (e.g., *Barron's*) selectivity ranking; information about students' high school (HS) or HS performance (e.g., average class rank); the percentage of students in the college receiving Pell grants or another measure of

[8] The relationship is modeled as quadratic in the figure. It should be apparent from inspection that different assumptions about the functional form of the relationship can dramatically alter institutions' adjusted performance.

[9] In particular, it should be the same vertical distance above the regression line (at Pell equal to 11 percent) as UTEP.

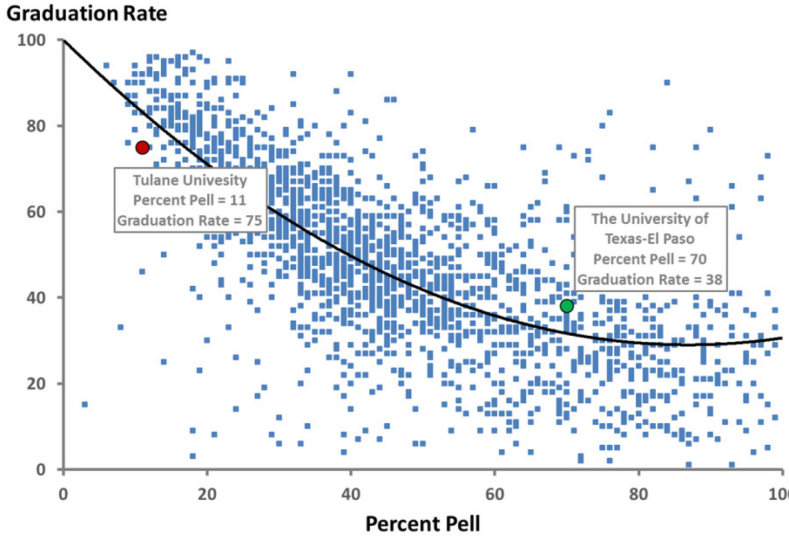

FIGURE C-1 Cohort 6-year graduation rates and the percent of full-time first-time undergraduate students receiving Pell Grants for four-year institutions
SOURCES: U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS) 6-year Graduation rate-bachelor's degree within 150 percent of normal time (2014) and percent of full-time first-time undergraduates receiving Pell grants (2011-2012).

financial aid receipt or family income; and demographic information such as the gender and race-ethnicity makeup, and average age of students (see for example, Bailey et al., 2005; Kelchen and Harris, 2012; Rothwell and Kulkarni, 2015). Estimates of college value-added based on these techniques yield dramatically different rankings of institutional quality when compared to raw graduation rates or average earnings.

Researchers have identified several limitations with this methodology and the ways it is commonly applied. First, many analysts additionally control for institutional factors, including the share of students enrolled online, the proportion of degrees awarded in various majors, overall enrollment, faculty salaries, and measures of education expenditures or endowment size. Arguably, however, all of these indicators represent strategies that can be manipulated by institutions in order to affect student outcomes. To the extent these factors represent key inputs into institution quality (or mechanisms through which colleges affect outcomes), then including them in the regression model will "over-adjust," or eliminate part of the true signal about institutional quality. A related flaw is that estimates of college quality can change dramatically due to small changes in the sets of variables used for regression adjustment, or the exact way in which they enter the regression (e.g., whether the relationship between graduation is assumed to be linear or quadratic).

A more subtle yet fundamental set of concerns with this approach stems from the use of aggregate student data. As Bailey and Xu (2012) note, controlling for aggregate student characteristics such as SAT scores and family income adjust for two separate influences on student outcomes. The first is the impact that students' academic family socioeconomic status (SES) background has on their own outcomes, and purging this part of the variation in institutional outcomes is the primary goal of regression adjustment. But aggregate student characteristics also capture the effect of the characteristics of a student's peers on his or her outcomes, and peer effects ought to be considered part of an institution's value-added.

A more important limitation is that it is likely that selection on the part of both students and college admissions offices leads to a correlation between institutional quality and aggregate student characteristics. To the extent this type of selection occurs, then the relationship between student characteristics such as the percentage receiving Pell grants and outcomes will reflect the impact of SES on outcomes, but also the higher quality of institutions with fewer poor students. In other words, in the presence of selection student demographics may "over control" and eliminate differences in outcomes across institutions due to quality differences. Figure C-2 provides evidence this may indeed be a serious flaw of this methodological approach using data on average earnings of students who receive Title IV aid at 4-year institutions. A regression of aggregate institution-level earnings on the fraction of Title IV students receiving Pell grants suggests that Pell students earn an average of nearly $25,000 less than non-Pell students, as shown by the slope of the regression line in the figure. But looking at the earnings outcomes of students within institutions reveals that in fact Pell students earn only $5,000 less than non-Pell students, and that both Pell and non-Pell students earn less at institutions with higher shares of Pell recipients. This is suggestive evidence that college quality measures based on regression adjustments using aggregate data may be very misleading, and in particular might overly adjust the outcomes of the highest and lowest quality schools toward the average quality school.

Both of the limitations just discussed can be addressed using individual-level data. In recent years, newly available administrative databases linking students' background characteristics to their college attendance and progression outcomes, and to their labor market earnings has enabled new progress in measuring college quality. This newer literature estimates the value-added of colleges using methods isomorphic to those used in the teacher quality literature, relying on the relationship between student characteristics and their outcomes *within* (rather than across) colleges to adjust institutional outcomes.[10] For example, Kurlaender et al. (2015) estimate the quality of community colleges in

[10] That is, the regression model $Y_{is} = X_{is}\beta + \mu_s + \epsilon_{is}$ is estimated, where Y_{is} is the outcome of student i attending school s, and X_{is} is a vector of student characteristics. μ_s are treated as fixed effects whose coefficients represent estimates of each school's quality.

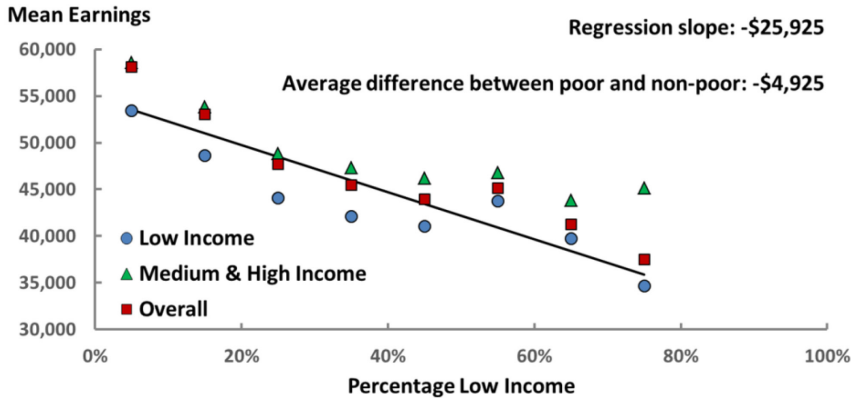

FIGURE C-2 Mean 10 year earnings for the 2001-2002 cohort by family income status (Low, Medium and High, and Overall) and percent of low income students at the institution.
NOTE: Low income is defined as less than $30,000, and medium/high income is defined as greater than $30,000.
SOURCE: U.S. Department of Education, College Scorecard data (2011).

California in terms of their impact of several progression outcomes, controlling for demographics, parental education, and eleventh grade test scores from students' high school transcripts.[11] Similarly, Cunha and Miller (2014) use rich data on students' academic and demographic background, as well as data on their application and acceptance history to estimate the value-added of 30 4-year institutions in the University of Texas system on graduation and earnings outcomes. Finally, the Council of Economic Advisers (2015) used student-level information from students' FAFSA forms, including their parental education, family income, and the SAT scores of institutions to which they sent their FAFSA form to estimate regression-adjusted earnings measures for every degree-granting institution in the United States.

All of these studies find substantial differences in the ranking of institutions using regression-adjusted rather than raw average outcomes. The studies also confirm that the set of variables used to adjust outcomes can matter a great deal for measuring the relative quality of institutions. Figure C-3 below adapts data from Cunha and Miller (2014) to show the estimated difference in average earnings between a flagship (Texas A&M) and regional (University of Texas [UT], Pan-American) institution based on five different models. The first column of the figure shows that UT Pan-American alumni earn 52 percent less than students who attended Texas A&M, but subsequent columns use different

[11] Kurlaender et al. (2015) report results from estimating the equation in footnote 9 as a random effects model, but note their results are similar if they estimate using the fixed effects approach described in the text.

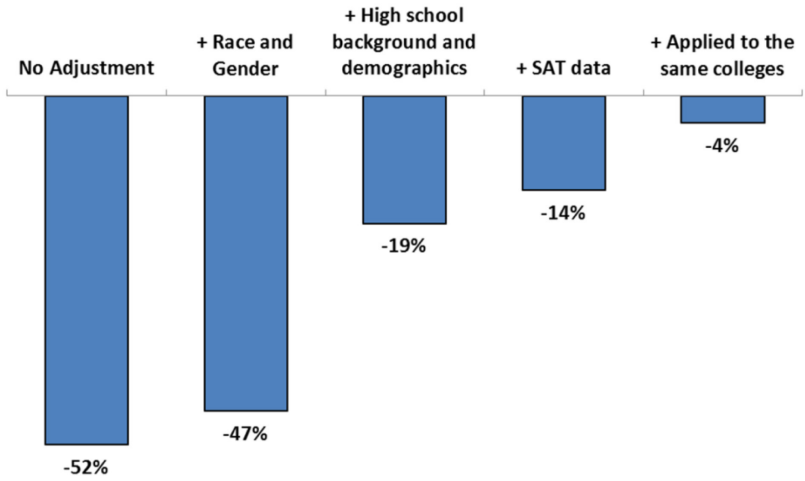

FIGURE C-3 Difference in average earnings for enrollees at a regional Texas university and a flagship after various regression adjustments.
SOURCE: Adapted from Table 3 in Cunha, Jesse M. and Trey Miller (2014). "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data." *Economics of Education Review* 42: 64-77.

input variables to adjust this difference for the differences in the students who attend. The second column shows that when comparing students who are similar in terms of their gender and race the difference falls to 47 percent. Adding student demographics and high school and SAT test scores further reduces the difference in earnings to 14 percent. And finally, for students similar on all of the preceding variables and who applied to the same set of colleges the students who attend the regional university have only 4 percent lower earnings.

Cunha and Miller's results highlight the importance of accounting for differences in student characteristics when measuring institution quality. A student growing up near UT Pan-American choosing whether to go to school locally or to Texas A&M might make a very different college choice if he believed his earnings would be only 4—not 52—percent higher if he attended A&M. But the results also show the frailty of these estimates, and how they can be very sensitive to the precise set of variables used for adjustment. This sensitivity begs the question of whether the results remain biased by unobserved differences in student characteristics across institutions even in the model that includes all available controls.

The gold standard for assessing causal impacts is to conduct a randomized control trial, but this is clearly infeasible in the context of measuring college quality. That said, in the past several years a handful of researchers have identified quasi-experiments that result in "as good as" random assignment of subsets of students to different institutions in a way that allows credible evaluation of their relative quality. Hoekstra (2009) was the first paper in this

literature, and exploited the fact that admission to a flagship state university in Florida depended on having SAT score above a threshold. Due to the admissions policy, the probability of enrolling at the flagship jumped by about 40 percentage points for students scoring just above the threshold relative to just below. In this setting, so long as other factors affecting outcomes do not change discontinuously for students who score just above and below the SAT threshold, the casual effect of attending the flagship university can be measured by comparing the outcomes of students scoring in a narrow window above and below the admissions cutoff. Hoekstra used this "regression-discontinuity design" to show large effects: attending the flagship increased the earnings of 28- to 33-year-old men by about 20 percent. Subsequent research by Zimmerman (2014) and Goodman, Hurwitz, and Smith (2015) found similar differences in institution quality among other schools using the same methodology.

These quasi-experimental studies are perhaps the most credible estimates available of the relative quality of institutions. They are, however, very limited in that they identify quality differences for only a handful of institutions and thus cannot hope to inform broader college choice or policy initiatives. In an ambitious recent effort, however, Hoxby (2015) uses a research design motivated by the regression-discontinuity studies to estimate the earnings value-added of all postsecondary institutions in the United States. Hoxby links individual-level SAT data from the College Board, college enrollment from the National Student Clearinghouse, and earnings data from the Internal Revenue Service. Using these data she employs two complementary research designs aimed at eliminating the effect of two types of selection from college quality estimates. First, to isolate similar students who enroll at institutions with different selectivity, she identifies students "on the bubble" of admissions at each school by finding applicants to each school whose admission probability based on their SAT score is between 40 and 60 percent.[12] Among this subset of students, she argues that whether they are admitted and enroll is "random," determined by college admissions responses to idiosyncratic aspects of students' application file that are unlikely to be determinants of their outcomes. To the extent this assumption holds, the effect of attending the institution relative to the alternatives chosen can be estimated by comparing the outcomes of students who are in this "bubble range" for each school.

The second research design Hoxby employs is meant to address "horizontal selection," and attempts to isolate situations where similar students might "randomly" choose among institutions with similar selectivity. To do so, she identifies every pair of colleges with very similar student SAT scores, and all students who apply to both institutions with a high probability of being admitted to each as a function of their own SAT score. Again, this allows the relative

[12] Admissions are not directly observed so this is estimated by the ratio between the share of students who enroll and the maximum enrollment rate across SAT scores.

quality of the pair of institutions to be estimated by comparing the outcomes of students meeting these conditions who attend each institution.[13] Each of Hoxby's methods yields estimates of the relative quality of a pair of institutions for particular types of students (defined by their test scores). Paired comparison methods (Langville and Meyer, 2013) are used to generate a consistent quality measure on common scale for all institutions.

Hoxby's results are preliminary and she has not yet published institution-specific estimates, but summary tables reveal large differences in quality across institutions. For example, relative to the least selective schools, schools in the top selectivity tier (defined by the SAT scores of their students) have an average value added that is more than $90,000 (i.e., causal impact on yearly earnings) higher. There is wide variation even within selectivity tiers as well: for example, among institutions with SAT math and verbal scores between the 25th and 30th percentiles, the school at the 90th percentile of the value-added distribution increases yearly earnings by roughly $9,000 more than the school at the 10th percentile of the distribution. Whether the assumptions underlying the methods Hoxby proposes survive further investigation is yet to be seen, but the data she has assembled (especially on students' pre-test scores) are the most complete available that can support value-added estimates for nearly the full universe of postsecondary institutions.

NEXT STEPS AND CHALLENGES

As described above, only limited progress has been made in developing quality measures that have the desirable properties described above. Below, I comment on a set of next steps that seem necessary building blocks in the construction of better quality indicators and some challenges and considerations that will need to be addressed.

Determining and measuring the outcomes that matter most.

The most fundamental limitation in our knowledge about college quality is a lack of consistent information about various student outcomes for broad groups of colleges. As noted above, in surveys the reasons students give for choosing to attend college tend to cluster in three main themes—to improve their employment opportunities, to learn more about the world and subjects of interest, and personal development or to become a better person. Research has provided little guidance about what specific measures in these domains might be most salient for prospective students, but this is in part since few institution metrics have been developed to test whether students respond to that

[13] A variant of this method is used to incorporate nonselective schools. Only comparison schools are based on the set of schools attended by a significant fraction of students from the past four classes at each student's high school.

information. What types of indicators should and could be developed? An initial and partial list follows.

A. **Labor market outcomes.** The number one reason to attend college offered by students in surveys is often related to improving their employment outcomes. There are several aspects of individuals' labor market outcomes that might be useful.
   1. Earnings. Newly available administrative data, either from tax data (generally W-2 and self-employment earnings) or from unemployment insurance data, allow earnings to be linked to individuals and the institutions they attended. Any function of earnings can be estimated in addition to simple averages, which may help allay concerns about incentivizing institutions to reduce enrollment in less lucrative programs. A variety of such earnings measures were recently released by the Department of Education in its updated College Scorecard, and growing numbers of state higher education systems are reporting the earnings outcomes of their graduates.
   2. Employment in a specific occupation, industry, or company. For many students, the goal of getting a "good" job may reflect a desire to attain a job in a particular occupation, industry, or company rather than (just) to get higher earnings. Unfortunately data reflecting this area are currently limited. Perhaps the most promising source of occupation data in the long-term is the (currently self-reported and hence at least somewhat unreliable) data on occupation reported on tax filings to the Internal Revenue Service. Practically speaking, pursuing this at scale would require modifications to the way this information is collected on tax forms (it is currently manually typed or hand-written into the form), but this seems technically feasible. Though falling short of measuring occupations, coarser information on types of employment is available in other data. For example, it is possible to identify the industry in which individuals work through employer identifiers on W-2 forms and unemployment insurance records, and also possible to know whether individuals work in the public sector via the tax status of their employer. Unemployment insurance data might also be used to identify the top company employers of each institution's students, but to my knowledge this has not been done to date.[14]
   3. Graduate degree attainment. While not a direct measure of occupational attainment, many undergraduate students report attending their institution to prepare them for graduate study. For students who receive

[14] LinkedIn, the online professional networking site, has computed the fraction of students in each institution working for "desirable" employers for select occupations heavily represented on their site based on the employment transitions of their users, but these measures are unlikely to be representative or even available for the majority of schools.

federal loans in their graduate studies, it is possible to observe whether they eventually obtain a graduate degree. But the fact that outcomes cannot be measured for students who might be fully funded or self-paying presents a significant challenge. An alternate source of data is the National Science Foundation's Survey of Completed Doctorates, which allows each Ph.D. recipient to be linked to her undergraduate institution.

4. Direct "productivity" estimates. For some types of postsecondary training aimed at preparing individuals for a particular occupation it may be feasible to measure the work performance of alumni, and compare this performance across institutions. This is the logic, for example, behind recent proposed regulations on teacher preparation programs, which seek to tie eligibility for some federal funding to the student test scores of teachers from the program. While data may exist in some sectors to measure raw performance (e.g., medical schools), the methodological challenges involved in estimating the value added of these programs are substantial due to selection into different types of employers.
5. Other outcomes. Policymakers may put extra value on institutions that foster entrepreneurship or innovation. In principle, we can measure entrepreneurship via self-employment income in tax data. The same data can probably also be used to create finer measures of business creation, with details about the number of employees. In similar fashion, the degree to which each institution's former students foster innovation might also be measured by linking data on patent recipients to where they attended college, as Chetty and coauthors have recently explored. Other outcomes that might be feasible to link to institutions' former students include social welfare program participation (e.g., Food Stamps or Temporary Assistance for Needy families) to measure family-level economic distress.

B. **Less tangible outcomes.** There is broad agreement among higher education officials, policymakers, and students that the quality of a college education should not be narrowly equated with labor market success. Myriad other goals for the sector include developing a passion and aptitude for lifelong learning, producing better global citizens, and becoming a better person. Unfortunately, there have been relatively few systematic attempts to measure these outcomes at scale since they do not yield easily to quantification. Nonetheless, future research ought to address which, if any, indicators might capture some of the less tangible benefits of education. A few types of measures that might be promising as "catchall" measures of quality are:

1. Subjective well-being (SWB). Many countries have recently promoted SWB measures to complement Gross Domestic Product as a yardstick to measure the performance of the economy to capture non-pecuniary aspects of well-being. Though researchers remain divided on the issue

of how appropriate such measures are for overall assessments of welfare, the measures certainly capture aspects of welfare that cannot be measured through employment outcomes. Though relatively little work has been done to date, several organizations, including PayScale and Gallup, collect information on measures related to SWB on a large scale that might provide insight to researchers.

2. Student satisfaction. Many institutions conduct surveys of their students' satisfaction or engagement, and some of these measures have been shown to correlate with higher graduation rates. In their current incarnation, generally asked while students are enrolled or as part of an exit survey, these measures are probably most useful in helping assess what institutional practices are associated with quality (i.e., the institution's impact on outcomes). It might be more useful to ask satisfaction questions well after students exit from institutions when they can reflect on how well their education helped them attain their goals. One could imagine such a survey being administered at scale to loan recipients, for example, through Federal Student Aid.

C. **Learning outcomes.** At least since the Spellings Commission, many have advocated and investigated measuring institution quality through learning value-added metrics. While there is currently no wide-scale deployment of an assessment that would allow the learning gains of large shares of students to be compared across many institutions in the United States, many efforts are under way to develop assessments in a wide variety of learning domains. As noted above, I largely ignore the body of research on such measures here, but a few points are worth making. These efforts have many advantages, such as being able to measure learning and growth in areas not likely to be reflected in labor market outcomes or other data, and the ability to provide high-frequency feedback about institutional performance with little lag due to measurement. On the other hand, researchers and policymakers should be explicit about whether the learning outcomes captured on these tests are valued *per se*, or because they are believed to lead to other valued outcomes. If the latter, then wherever possible direct measurement of the outcome is preferred. For example, rather than measuring learning valued because of its instrumental value in enhancing students' employment prospects, direct measurement of employment outcomes seems preferable.[15] When direct measurement is not possible, the validity of tests as proxies for the outcomes ultimately valued should be investigated. In a recent study, Melguizo, Zamarro, Velasco, and Sanchez (2015) does just this by leveraging the fact that in Colombia there is a mandatory exit exam meant to test general skills

[15] Of course, linking the two sources of information together brings other benefits, such as pointing to particular competencies that are valued in the labor market to guide improvements to curricula, etc.

(based on the CLA) to compute institution value-added measures for both learning outcomes and early career earnings. Melguizo et al. show that for many programs the correlations between these measures are low, and for some programs they are negative. At a minimum, these results suggest caution in treating measures of student learning outcomes as valid quality measures without further validation. A second concern is whether it is realistic to field the necessary assessment tests at large scale both within and across institutions, given the current climate surrounding testing. Researchers should also investigate whether students will find this type of quality information to be salient when making college choice decisions, or whether the information will primarily be useful for accountability schemes or institutions seeking to improve.

Developing and validating methods to measure the causal effects of institutions and the mechanisms through which they affect outcomes.

As discussed above, researchers have yet to determine which of several methods available best measure the quality of postsecondary institutions. In recent years, however, high-quality quasi-experimental research designs have yielded credible estimates of the impact of attending particular colleges on student progression and earnings outcomes. Unfortunately, relatively few institutions have institutional features—such as test score cutoffs used in their admissions decisions—that would permit this type of evaluation of their quality. Estimates of quality from these studies may nonetheless prove useful in future research, by providing a way to validate quality estimates from other research designs with less *ex ante* validity. For example, with quasi-experimental quality estimates for enough institutions we could compare the quality estimates based on non-experimental methods (such as the regression-based value-added measures presented in Cunha and Miller (2014) or CEA (2015)) and evaluate which, if any, provide accurate estimates of institutions' "true" quality. This type of work could help identify, for example, whether individual-level data is critical, and what core sets of student characteristics need to be included in regression adjustment models for them to produce accurate quality measures.[16]

Researchers also need to develop information about which institutional practices contribute most to quality to help inform improvement efforts. Once quality estimates exist, correlational studies can help identify candidate factors that appear to be the most important determinants of quality. Subsequent studies

[16] For example, McClellan and Staiger (2000) discuss how in the history of hospital quality measures, methods once thought too coarse to sufficiently eliminate selection effects were eventually revealed to perform nearly as well as models based on much richer information when richer data became available. Similarly, in the literature on teacher value-added, several papers have suggested that non-experimental value-added estimates can provide unbiased measures of true (e.g., experimentally estimated) teacher quality (Kane and Staiger, 2008; Chetty et al., 2014).

can help distinguish whether these practices have causal impacts on student outcomes, or perhaps are simply correlated with other practices that warrant further investigation.

Are different measures appropriate for setting minimum standards of quality as opposed to an overall assessment of quality?

One of the most pressing sources of demand for quality information about postsecondary institutions is to inform public policy efforts to ration public funds away from institutions with unsatisfactory outcomes. For example, the Higher Education Act proscribes participation in Title IV financial aid programs for institutions with high default rates on federal loans, and the Gainful Employment regulations do the same for vocational training programs whose graduates have high debt-to-earnings ratios. Neither of these performance measures reflects the causal impact of institutions, but proponents of the measures have argued that an absolute minimum standard may be more appropriate. Leaving the particulars of these policies aside, it is important for researchers and the higher education sector to consider what levels of performance on various indicators should be considered unacceptable for various purposes, especially in determining eligibility for participation in various types of public aid programs. It may be that some measures of institutional performance are valuable for setting minimum standards—e.g., by reliably identifying low-quality institutions—even if they might not be suitable proxies for institution quality across the full distribution of quality.

Should summative measures of quality be constructed from various measures of particular dimensions of quality? Under what circumstances and how?

The discussion above has focused on separately estimating various dimensions of college quality, or the causal impact of institutions on various outcomes of interest. For a variety of purposes, however, summarizing this information in a single quality rating may be desirable. Many consumer-oriented ratings sites use a "star system" or some other one-dimensional rating in order to convey a summary of a variety of different indicators in a way that is simple and easy for consumers to understand and act upon. Moreover, accountability schemes such as the performance-based funding systems that many states are currently experimenting with presume a single index of quality that maps into the amount of funds that should be allocated to each university. While indices of overall quality can be formed by averaging together multiple measures of different dimensions of quality (suitably normalized, for example by expressing measures in standard deviation units relative to the mean), there is little guidance in the literature on how heavily to weight each dimension of quality.

Is there a principled way to combine measures of institutional quality to inform either college choices (for particular student types), or for accountability efforts such as performance based funding schemes?

What level of quality information is most useful for various purposes?

Finally, as noted above, it is likely that the quality of an institution varies considerably for different types of students, and for different programs within the institution. This begs the question of what level should quality information be constructed to be most useful for informing college choice or other purposes. For example, some states have begun to report outcomes data, such as average earnings, for graduates separately by program of study. On the one hand, this might provide valuable information to students to inform what major they should pursue, and might diminish concerns that institutions might drop programs with low-earning graduates to improve their measured performance. On the other hand, many students do not complete degrees in the fields they initially express interest even when they do earn a degree. It may be preferable, then, to report outcomes (or quality estimates) for students based on their stated program of interest at enrollment, or perhaps to combine estimates of the fraction of students who graduate in each program by their initial stated interests in conjunction with the outcomes of graduates. In practice this may be complicated since many institutions do not require students to declare a program of study in their first years of study.

Another dimension of the question is whether outcome or quality information should be computed separately for different types of students. Again, research has shown that different types of students (e.g., low-income vs. higher-income) can experience different impacts on their outcomes from attending the same institution, so tailoring the information might provide more accurate information to inform choices. At the same time, some observers have expressed concerns about whether presenting tailored information to students showing that their particular subgroup's outcomes tend to be lower than average might (overly) discourage them from enrolling in college, or induce poorer performance perhaps by exacerbating anxieties about belonging or other behavioral channels. Experimentation to test how students might respond to various types of information content and presentation is needed to investigate this complex set of issues.

REFERENCES

Autor, David H. (2014). "Skills, Education, and the Rise of Earnings Inequality Among the 'Other 99 Percent.'" *Science*, 23: 344 (6186), 843-851.

Bailey, Thomas and Di Xu. (2012). "Input-Adjusted Graduation Rates and College Accountability: What Is Known from Twenty Years of Research?" HCM Strategists Working Paper.

Bailey, Thomas, Juan Carlos Calcagno, Davis Jenkins, Gregory S. Kienzl, and D. Timothy Leinbach. (2005). "The effects of institutional factors on the success of community college students." New York: Columbia University, Teachers College, Community College Research Center.

Baum, Sandy. (2014). *Higher Education Earnings Premium: Value, Variation, and Trend*. The Urban Institute.

Black, Dan and Jeffrey Smith. (2004). "How robust is the evidence on the effects of college quality? Evidence from matching." *Journal of Econometrics* 121: 99-124.

Black, Dan and Jeffrey Smith. (2006). "Evaluating the returns to college quality with multiple proxies for quality." *Journal of Labor Economics* 24(3): 701-728.

Brewer, Dominic J., and Ronald G. Ehrenberg. (1996). "Does It Pay to Attend an Elite Private College? Evidence from the Senior High School Class of 1980." *Research in Labor Economics* 15:239-71.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. (2014). "Measuring the impact of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104(9): 2633-2679.

CEA (Council of Economic Advisers). (2015). "Using Federal Data to Measure and Improve the Performance of U.S. Institutions of Higher Education." Washington, DC.

Cunha, Jesse M. and Trey Miller. (2014). "Measuring value-added in higher education: Possibilities and limitations in the use of administrative data." *Economics of Education Review* 42: 64-77.

Dale, Stacy Berg and Alan B. Krueger. (2002). "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." *Quarterly Journal of Economics* 107(4): 1491-1527.

Dale, Stacy Berg and Alan B. Krueger. (2011). "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data." NBER Working Paper No. 17159.

Goodman, Joshua, Michael Hurwitz, and Jonathan Smith. (2015). "College Access, Initial College Choice and Degree Completion." NBER Working Paper No. 20996.

Hoekstra, Mark. (2009). "The effect of attending the flagship state university on earnings: A discontinuity-based approach." *The Review of Economics and Statistics* 91(4): 717-724.

Hoxby, Caroline. (2015). "Value-Added of American Postsecondary Institutions." Stanford University mimeo.

Institute of Medicine. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press.

James, Estelle, Nabeel Alsalam, Joseph C Conaty, and Duc-Le To. (1989). "College quality and future earnings: Where should you send your child to college?" *American Economic Review* 79(2): 247-252.

Kane, Thomas J. and Douglas O. Staiger. (2008). "Estimating Impacts on Student Achievement: An Experimental Evaluation." NBER Working Paper No. 14607.

Kelchen, Robert and Douglas N. Harris. (2012). "Can 'Value Added' Methods Improve the Measurement of College Performance? Empirical Analyses and Policy Implications." HCM Strategists.

Kurlaender, Michal, Scott Carrell, and Jacob Jackson. (2015). "The Promises and Pitfalls of Measuring Community College Quality." University of California, Davis mimeo.

Langville, Amy and Carl D. Meyer. (2013). *Who's #1? The Science of Rating and Ranking*. Princeton, NJ: Princeton University Press.

Long, Mark C. (2008). "College quality and early adult outcomes." *Economics of Education Review* 27(5): 588-602.

Loury, Glen and David Garman. (1995). "College selectivity and earnings." *Journal of Labor Economics* 13(2): 289-308.

McClellan, Mark and Douglas Staiger. (1999). "The Quality of Health Care Providers." NBER Working Paper 7327.

McClellan, Mark and Douglas Staiger. (2000). "Comparing the Quality of Health Care Providers," in Alan M. Garber, ed., *Frontiers in Health Policy Research, Volume 3*, Cambridge, MA: MIT Press.

Melguizo, Tatiana, Gema Zamarro, Tatiana Velasco, and Fabio J. Sanchez. (2015). "How Can We Accurately Measure Whether Students Are Gaining Relevant Outcomes in Higher Education." University of Southern California mimeo.

Monks, James. (2000). "The returns to individual and college characteristics: Evidence from the National Longitudinal Survey of Youth." *Economics of Education Review* 19(3): 279-289.

Rothwell, Jonathan and Siddharth Kulkarni. (2015). "Beyond College Rankings: A Value-Added Approach to Assessing Two- and Four-Year Schools." Brookings Metropolitan Policy Program.

Zimmerman, Seth. (2014). "The returns to college admission for academically marginal students." *Journal of Labor Economics* 32(4): 711-754.

Appendix D
Planning Committee and Staff Biographies

PAUL COURANT (*Chair*) is Arthur F. Thurnau Professor, Harold T. Shapiro Collegiate Professor of Public Policy, Professor of Economics, Professor of Information, and Faculty Associate in the Institute for Social Research at the University of Michigan. From 2007 to 2013 he was University Librarian and Dean of Libraries at the University of Michigan. From 2002 to 2005 he served as Provost and Executive Vice-President for Academic Affairs, the chief budget officer and chief academic officer of the university. He has also served as the Associate Provost for Academic and Budgetary Affairs, Chair of the Department of Economics, and Director of the Institute of Public Policy Studies (which is now the Gerald R. Ford School of Public Policy). In 1979 and 1980 he was a Senior Staff Economist at the Council of Economic Advisers. Dr. Courant has authored 6 books and more than 70 papers covering a broad range of topics in economics and public policy, including tax policy, local economic development, gender differences in pay, housing, radon and public health, relationships between economic growth and environmental policy, and university budgeting systems. More recently, he has been studying the economics of universities, the economics of libraries and archives, and the changes in the system of scholarly communication that derive from new information technologies. Dr. Courant holds a BA in history from Swarthmore College (1968), an MA in economics from Princeton University (1973), and a PhD in economics from Princeton University (1974).

ANGELA BYARS-WINSTON is an Associate Professor in the University of Wisconsin (UW), Department of Medicine, Division of General Internal Medicine. Her research interests include the examination of cultural influences on career development, especially for racial and ethnic minorities and women in the sciences, engineering, and medicine. Specifically, she has focused on testing the validity of theoretical models to explain and predict academic and career outcomes using social cognitive theoretical approaches. In 2010, Dr. Byars-Winston was awarded a multi-year R01 grant from the National Institutes of Health as Principal Investigator to measure and test critical factors in research

training interventions for mentors of ethnically diverse mentees in biological science. She was recently selected as a Champion of Change by the White House through President Obama's *Winning the Future* initiative for her research efforts to diversify science fields. Dr. Byars-Winston completed a predoctoral clinical internship at the University of Maryland, College Park and received her PhD in counseling psychology from Arizona State University. She was a KL2 scholar at ICTR in the UW School of Medicine and Public Health and is a Researcher in the UW Center for Women's Health Research.

ELLEN HAZELKORN is Director of Research and Enterprise, and Dean of the Graduate Research School, Dublin Institute of Technology, Ireland. She also leads the Higher Education Policy Research Unit. She works closely with the International Association of Universities (IAU) and is a consultant to the Organisation of Economic Co-operation and Development (OECD). She is a member of the Higher Education Authority (Ireland) and incoming President of EAIR (the European Higher Education Society). Prof. Hazelkorn has been/is a member various governmental and international review teams. In addition, she has worked/is working with universities and university associations around the world, including Australia, Belgium, Denmark, Egypt, Ethiopia, Jamaica, New Zealand, the Netherlands, Romania, Serbia, Saudi Arabia, and Sweden, and with private organizations on issues of higher education assessment and evaluation. Prof. Hazelkorn is Visiting Professor at the University of Liverpool and a member of the International Research Committee of the American Education Research Association (2012-2015). She is/has been a member of various editorial boards: *Higher Education Policy*, *International Journal for Researcher Development*, and *Higher Education Management and Policy*. She was Chairperson of the Dublin Regional Higher Education Alliance (2011-2012) and Visiting Fellow at the University of North London (now London Metropolitan University), OECD, Paris, and Centre for the Study of Higher Education, University of Melbourne, Australia. She was rapporteur and lead author for the report of the European Union (EU) Expert Group, *Assessing Europe's University-based Research* (2010), and a member of the Arts, Humanities and Social Sciences Foresight Working Group (Ireland). Prof. Hazelkorn has more than 18 years of senior management experience in higher education, previously holding the position of Vice President and founding Dean of the Faculty of Applied Arts, Dublin Institute of Technology (1995-2008). She was Company Secretary, Contemporary Music Centre (2001-2007); Board Member on Higher Education Equality Unit (1996-2002) and the European League of Institutes of the Arts (2000-2002); and Deputy Chairperson, Centre for International Technology and Education—a consortium of European academic, research, and industrial institutions developing EU research programs in intelligent digital content (1996-2000). She was awarded a BA and PhD from the University of Wisconsin, Madison, and the University of Kent, UK, respectively. She is leading a study of "The Impact of the Global Economic Crisis on Higher Education: Higher Education Leadership and Management Challenges" with the

IAU and co-leader of an European Science Foundation project measuring the societal impacts of universities' research into arts and the humanities (HERAVALUE), the result of which will be published by Palgrave in 2014.

PAUL LeBLANC is President of Southern New Hampshire University (SNHU). Under the 10 years of his direction, SNHU has more than tripled in size and is the largest provider of online higher education in New England, one of the five largest in the country, and the first to have a full competency-based degree program (untethered to the credit hour or classes) approved by a regional accreditor and the U.S. Department of Education. In 2012 the university was #12 on *Fast Company* magazine's "World's Fifty Most Innovative Companies" list and was the only university included. Dr. LeBlanc won a New England Higher Education Excellence Award in 2012 and was named one of "New Hampshire's Most Influential People" by New Hampshire Business Review. In 2012 *Forbes* magazine listed him as one of its 15 "Classroom Revolutionaries," and he was featured on Bloomberg TV's "Innovators" series. He speaks frequently to industry, Institutions of higher education, national policymakers, and other higher education stakeholders and often appears in the media. Dr. LeBlanc immigrated to the United States as a child, was the first person in his extended family to attend college, and is a graduate of Framingham State University (BA), Boston College (MA), and the University of Massachusetts (PhD). He directed a technology start-up for Houghton Mifflin Publishing Company (1993-1996), was President of Marlboro College in Vermont (1996-2003), and became President of SNHU in 2003.

ALEXANDER C. McCORMICK succeeded George D. Kuh as the Director of the National Survey of Student Engagement (NSSE) in January 2008. He also holds a faculty appointment in the Indiana University School of Education's Educational Leadership and Policy Studies department, where he teaches in the Higher Education and Student Affairs program. Through his work with NSSE, Dr. McCormick aims to enrich the national discourse about quality and accountability in higher education, while also providing institutions with tools to diagnose and improve undergraduate teaching and learning. His research interests center around assessment and accountability in higher education, as well as organizational change and improvement in higher education. Prior to joining Indiana University, Dr. McCormick served as Senior Scholar at The Carnegie Foundation for the Advancement of Teaching, where he directed a major overhaul of the Foundation's widely-used Classification of Institutions of Higher Education and also served as director of survey research. McCormick began his career in higher education as an admissions officer at Dartmouth College, where he subsequently served as Assistant Dean of the College. Dr. McCormick holds a PhD (education and sociology), a master's degree (educational administration and policy analysis) from Stanford University, and a bachelor's degree (French) from Dartmouth College.

MARCO MOLINARO has a dual BS in biophysics and chemistry from Wayne State University in Detroit, Michigan, and a PhD in biophysical chemistry from the University of California (UC), Berkeley. Since the early 1990s, he has been strongly involved with education at all levels and technology. From 1994 to 1999, Dr. Molinaro was involved in various national efforts (ModularChem Consortium and ChemConnections) to reform the undergraduate curriculum in chemistry utilizing problem-based approaches and technology. During that period, he spent 1 year as a Research Fellow at the University of Wisconsin, Madison, to research faculty use of technology in instruction. From 1998 to 2003, he was the Founder and Director of the ScienceVIEW educational multimedia design, research, and development group at the Lawrence Hall of Science (LHS) at UC Berkeley, specializing in creating multimedia materials aimed at teaching and learning science in formal and informal settings. Between his earlier chemistry work and LHS, he has developed more than 15 major CD-ROM and Internet-based products for teachers, students, and families. During his tenure as ScienceVIEW Director, he also led various research efforts related to educational technology effectiveness including learning-optimized use of molecular simulations in the classroom, understanding the potential of computer-based data collection for formative assessment in formal and informal learning environments, and developing usability guidelines for creating age-appropriate interactive activities on the Internet. From 2002 to 2005 he led the Windows on Research: Focus on Nanotechnology public exhibit project (Nanozone.org) aimed at communicating nanoscience to visitors ages 8-14 years. He is currently focusing his attention to communicating the latest research results, and the science behind them, to students of all ages in both formal and informal settings with an emphasis on actively engaging participants in "doing" science. Now at UC–Davis, he is the Chief Education Officer for the Center for Biophotonics, where he coordinates all of the Center's educational activities, including those aimed at K-12, higher education, and the public. One of latest projects is titled Biophotonicsworld.org—a biophotonics knowledge base for education, research, and industry. Dr. Molinaro is a member of the graduate group in the School of Education at UC–Davis. His current and ongoing research interests involve social interactions around technology use in informal and formal science settings, methods for facilitating public understanding of research, integration of cutting-edge scientific research and researchers with the formal and informal educational enterprises, and approaches to attracting, engaging, and retaining underrepresented students in science.

INDIRA NAIR recently retired from Carnegie Mellon University after 32 years. During her last 12 years there, she was the Vice Provost for Education and Professor in the Department of Engineering and Public Policy. Her research covered risk assessment, policy and risk communication, green design, bioelectromagnetics, education in general, engineering and innovation education, education assessment, and pedagogies for the modern-day literacies such as scientific, environmental, and global literacy. She designed and taught

several interdisciplinary courses including ethics of science and technology; environmental science; technology and decision-making; and radiation, health, and policy. Dr. Nair currently chairs the national Global Learning Leadership Council of the American Association of Colleges & Universities (AAC&U). She advises several universities and colleges on incorporating global and environmental literacy and ethics throughout the curriculum, on faculty development and on interdisciplinary education. She has received several National Science Foundation (NSF) grants and served on numerous national committees. She founded the Carnegie Mellon Chapter of Student Pugwash to encourage students to think about the social responsibility of science and technology. Her current quests and involvements include developing a new scheme for general education including the new literacies; investigating pedagogies for educating for innovation; increasing the inclusion of underrepresented minorities across all segments of education; and improving K-12 STEM education and bioelectromagnetics. She holds a PhD in physics from Northwestern University and a Pennsylvania teachers certificate for high school science teaching.

ROY SWIFT is currently the Executive Director of Workcred. Prior to joining Workcred, he served as the Chief Workforce Development Officer and Senior Director of Personnel Credentialing Accreditation Programs at the American National Standards Institute (ANSI). Prior to ANSI, he was a consultant to educational, certification, licensure, and health care organizations. From 1993 to 1998, he was Executive Director of the National Board for Certification in Occupational Therapy (NBCOT). This appointment followed a 28-year career in the U.S. Army Medical Department. In his last position, he was Chief of the Army Medical Specialist Corps in the Army Surgeon General's Office with policy responsibility for Army occupational therapists, physical therapists, dietitians, and physician assistants throughout the world. He has served on many national committees, nonprofit boards of directors, and federal and state government advisory committees. He has served as Chair of the Assembly of Review Committee Chairs of the former Committee on Allied Health Education and Accreditation of the American Medical Association; Chair of the American Occupational Therapy Association Accreditation Committee (Academic Accreditation); and member of the Secretary of the Department of Veterans Affairs Advisory Committee for Certification. Dr. Swift recently served on an Institute of Medicine panel dealing with Provision of Mental Health Counseling Services under TRICARE, and a planning committee for the future of Allied Health Practice. In addition, Dr. Swift recently chaired an international working group within the International Accreditation Forum (IAF) to recognize personnel certifications among member countries through the development of multilateral recognition arrangements. He is also active on working groups related to personnel credentialing in the International Organizational for Standardization (ISO) in Geneva, Switzerland, and is a guest lecturer at the University of Geneva on credentialing. Dr. Swift holds a BS in occupational

therapy from the University of Kansas, an MS Ed from the University of Southern California, and a PhD in continuing and vocational education with an emphasis in continuing competency in the professions from the University of Wisconsin, Madison. He has also successfully completed the University of Chicago's 3-week management development course. Formed in 2014, Workcred connects credentials, competencies, careers, and customers. Its mission is to serve industry needs and the public by creating new knowledge about industry credentials and enhancing the quality, transparency, market value, and portability of competency-based and industry-endorsed credentials to increase the quality, productivity, and performance of workers. An ANSI affiliate, Workcred is currently undergoing incorporation as a 501(c)3 not-for-profit corporation. Workcred's independence as a legal entity will maintain separation from and respect the impartiality of ANSI's accreditation services.

STAFF

THOMAS RUDIN is Director of the Board on Higher Education and Workforce at the National Academies of Sciences, Engineering, and Medicine—a position he assumed in August 2014. Prior to joining the Academies, Mr. Rudin served as Senior Vice President for Career Readiness and Senior Vice President for Advocacy, Government Relations and Development at the College Board from 2006 to 2014. He was also Vice President for Government Relations from 2004 to 2006 and Executive Director of Grants Planning and Management from 1996 to 2004 at the College Board. Before joining the College Board, Mr. Rudin was a policy analyst at the National Institutes of Health in Bethesda, Maryland. In 1991, Mr. Rudin taught courses in U.S. public policy, human rights, and organizational management as a visiting instructor at the Middle East Technical University in Ankara, Turkey. In the early 1980s, he directed the work of the Governor's Task Force on Science and Technology for North Carolina Governor James B. Hunt, Jr., where he was involved in several new state initiatives, such as the North Carolina Biotechnology Center and the North Carolina School of Science and Mathematics. He received a BA from Purdue University, and he holds master's degrees in public administration and in social work from the University of North Carolina at Chapel Hill.

AQILA COULTHURST is an Associate Program Officer with the Board on Science, Technology, and Economic Policy (STEP) at the National Academies of Sciences, Engineering, and Medicine where she works on a broad range of policy issues ranging from workforce development and immigration to intellectual property and innovation. She joined STEP in the fall of 2011 after serving as marketing specialist at the National Academies Press (NAP) for 2 years. Prior to her work at the Academies, Ms. Coulthurst developed and directed a leadership development and civic engagement program for at-risk youth in the DC area. She has also worked in various capacities at Smithsonian Enterprises, the National Community Action Foundation, and SRI International.

Ms. Coulthurst earned a BA in economics, a BA in Spanish, and a certificate in markets and management from Duke University. She received an MS in foreign service with a concentration in international development from Georgetown University.

MARIA LUND DAHLBERG is an Associate Program Officer with the National Academies of Sciences, Engineering, and Medicine. She works with a number of groups across the institution, including the Board on Higher Education and Workforce, the central Office of Communications, and the Committee on Science, Engineering, and Public Policy, specializing in report review, production, and dissemination. She has contributed to more than 18 reports and 10 different boards since joining the Academies in January 2012. She came to the Academies by way of a Christine Mirzayan Science and Technology Policy Fellowship, which she received after completing all requirements short of finalizing the dissertation for her doctorate in physics at the Pennsylvania State University. Ms. Dahlberg holds a BA in physics from Vassar College and an MS in physics from the Pennsylvania State University.

IRENE NGUN is a Research Associate with the Board on Higher Education and Workforce at the National Academies of Sciences, Engineering, and Medicine. Before joining the National Academies, Ms. Ngun was a congressional intern with the House Science, Space, and Technology Committee's Democratic Office and also served the office of her district congresswoman, Ms. Eddie Bernice Johnson (TX-30). Ms. Ngun holds an MA in international relations from Yonsei Graduate School of International Studies and a BA in biochemistry and economics from Goshen College.

ADRIANA NAVIA COUREMBIS joined the National Academies of Sciences, Engineering, and Medicine in January 2012 as a Financial Associate for the Policy and Global Affairs Division. As a Financial Associate, she collaborates with the financial management for the Board on Higher Education and Workforce, the Committee on Women in Science, Engineering and Medicine, the Science & Technology for Sustainability Program, the Committee on Human Rights, and the Board on Research Data and Information. Prior to the Academies, Mrs. Courembis worked with the American Bar Association—Rule of Law Initiative as a Program Associate and Bay Management, LLC as an Accounts Payable Associate. Mrs. Courembis holds a BA in International Economics from American University.