

Methodology of Educational Measurement and
Assessment

Olga Zlatkin-Troitschanskaia
Miriam Toepper · Hans Anand Pant
Corinna Lautenbach · Christiane Kuhn
Editors

Assessment of Learning Outcomes in Higher Education

Cross-National Comparisons and
Perspectives

 Springer

Methodology of Educational Measurement and Assessment

Series Editors

Bernard Veldkamp, Research Center for Examinations and Certification (RCEC),
University of Twente, Enschede, The Netherlands

Matthias von Davier, National Board of Medical Examiners (NBME), Philadelphia,
USA

This book series collates key contributions to a fast-developing field of education research. It is an international forum for theoretical and empirical studies exploring new and existing methods of collecting, analyzing, and reporting data from educational measurements and assessments. Covering a high-profile topic from multiple viewpoints, it aims to foster a broader understanding of fresh developments as innovative software tools and new concepts such as competency models and skills diagnosis continue to gain traction in educational institutions around the world. *Methodology of Educational Measurement and Assessment* offers readers reliable critical evaluations, reviews and comparisons of existing methodologies alongside authoritative analysis and commentary on new and emerging approaches. It will showcase empirical research on applications, examine issues such as reliability, validity, and comparability, and help keep readers up to speed on developments in statistical modeling approaches. The fully peer-reviewed publications in the series cover measurement and assessment at all levels of education and feature work by academics and education professionals from around the world. Providing an authoritative central clearing-house for research in a core sector in education, the series forms a major contribution to the international literature.

More information about this series at <http://www.springer.com/series/13206>

Olga Zlatkin-Troitschanskaia
Miriam Toepper • Hans Anand Pant
Corinna Lautenbach
Christiane Kuhn
Editors

Assessment of Learning Outcomes in Higher Education

Cross-National Comparisons
and Perspectives

 Springer

Editors

Olga Zlatkin-Troitschanskaia
Johannes Gutenberg University
Mainz, Germany

Miriam Toepper
Johannes Gutenberg University
Mainz, Germany

Hans Anand Pant
Humboldt-Universität zu Berlin
Berlin, Germany

Corinna Lautenbach
Humboldt-Universität zu Berlin
Berlin, Germany

Christiane Kuhn
Johannes Gutenberg University
Mainz, Germany

ISSN 2367-170X ISSN 2367-1718 (electronic)
Methodology of Educational Measurement and Assessment
ISBN 978-3-319-74337-0 ISBN 978-3-319-74338-7 (eBook)
<https://doi.org/10.1007/978-3-319-74338-7>

Library of Congress Control Number: 2018930984

© Springer International Publishing AG, part of Springer Nature 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Introduction: Assessing Student Learning Outcomes in Higher Education

Until the last decade, objective information on student learning and student learning outcomes in higher education at the national and international levels was scarce. This area was largely underrepresented in comparison to other areas of formal education such as school. In the context of current developments in higher education such as internationalization of study programs and ever-increasing student mobility and the ensuing increase in heterogeneity of students' learning conditions, the need for objective, valid, and reliable assessment tools that adhere to the Standards for Educational and Psychological Testing set out by the American Educational Research Association (AERA) has become urgent. This has led to intense research efforts being made within and across many countries, which are of great practical and political importance.

This book presents the most significant of these initiatives and developments in order to highlight the tremendous work national and international research communities have done in this area over the past decade. A broad range of national and international assessment research projects and curricular innovation initiatives in higher education focusing on both domain-specific and generic student learning outcomes are presented in this volume. Results and lessons learned from various research programs such as the German Modeling and Measuring Competencies in Higher Education (KoKoHs) and feasibility studies such as the Assessment of Higher Education Learning Outcomes (AHELO; an international comparative study by the Organisation for Economic Co-operation and Development (OECD) of students' generic skills and economic and engineering competencies) form the basis of several ongoing initiatives by testing institutes to make assessments suitable for use in higher education abroad. Examples include the Educational Testing Service's (ETS) Heighten Outcomes assessment and the Council for Aid to Education's (CAE) Collegiate Learning Assessment CLA+. At the European level, the CALOHEE initiative on Measuring and Comparing Achievements of Learning Outcomes in Higher Education in Europe aims to develop a joint basis for learning outcomes in higher education as well as curricula in five disciplines, including education. One of the most current international initiatives, the International Collaborative for Performance Assessment of Learning in Higher Education (iPAL),

focuses on developing performance assessments of learning that meet high standards in psychometric quality criteria and are suitable for use at higher education institutions across nations.

The compilation of this work in the present book shows where we stand today and the progress that has been made in this field of research with newly developed theoretically conceptualized approaches to modeling and measurement instruments for empirical studies. It also illustrates which issues have not yet been thoroughly addressed by the – indeed very active – research community measuring student learning in higher education. Therefore, this book offers a sound basis for further research, highlighting the current challenges and future perspectives in measuring learning and learning outcomes in higher education we need to deal with in the next decades.

The contributions in this book are organized according to content and topic and are divided into three parts (Conceptual Development and Advances, Domain-Specific Student Learning Outcomes – National and International Perspectives, and Generic Student Learning Outcomes – Cross-National Comparative Approaches), giving the reader a structured overview of the wide range of student learning outcomes assessment in higher education.

The book begins with an outline of the contributions of the first thematic block, **Conceptual Development and Advances**.

In the first chapter in this part, *Research and Governance Architectures to Develop the Field of Learning Outcomes Assessment*, Hamish Coates gives an overview of research in the field and presents a conceptual approach, which indicates a possible path of development for this research area. Furthermore, he develops a framework for research and paints a picture of what assessment will look like in a decade's time.

In the second chapter, *Documenting and Improving Collegiate Learning in the United States*, Timothy Reese Cain and James C. Hearn describe the historical development and current features of learning outcomes assessment in higher education in the United States. They discuss the changing context of and increased interest in learning outcomes in the late twentieth and early twenty-first centuries, including the roles of various external actors and stakeholders. They conclude with considerations of on-campus actors at the hub of assessment processes and the evidence of changing practices that has been revealed through recent national surveys by research and volunteer groups.

In the third and last chapter in this part, *Information Management Versus Knowledge Building: Implications for Learning and Assessment in Higher Education*, Patricia A. Alexander builds a highly significant bridge between the concepts of *assessment* and *learning* in the sense of the Assessment Triangle. In the process, the author introduces another theoretical concept, which is highly relevant for learning in higher education in the information age and distinguishes between information management and knowledge building in the sense of key learning outcomes.

The second part of this book, **Domain-Specific Student Learning Outcomes – National and International Perspectives**, explores current empirical work in the

field of domain-specific learning outcomes. Research in the areas of teacher education in medicine and, as of late, in business and economics can be regarded as particularly advanced. This part gives an exemplary presentation of the work in this field and complements it with an outlook on a new European initiative for assessing domain-specific student learning outcomes.

Entrance diagnostics is an integral part of student learning outcomes assessment in higher education. The domain of medicine has had a pioneering role in this field of assessment. In their chapter *Challenges for Evaluation in Higher Education: Entrance Examinations and Beyond – The Sample Case of Medical Education*, Christiane Spiel and Barbara Schober discuss entrance diagnostics in higher education using the example of medical education. They highlight the potential limitations of typical entrance examinations for medical education; summarize the changes, concepts, and goals of entrance examinations in medical education have undergone in recent years; propose a comprehensive evaluation model for competence-based teaching; and explore implications for education.

Diagnostic competences are of vital importance not only in medicine but in teacher education practice as well. In their chapter *Teachers' Judgments and Decision Making: Studies Concerning the Transition from Primary to Secondary Education and Their Implications for Teacher Education*, Sabine Krolak-Schwerdt, Ineke M. Pit-ten Cate, and Thomas Hörstermann focus on accuracy in assessing academic achievement and potential as a core facet of teachers' diagnostic competence. Their findings regarding teachers' information processing emphasize the need to include situational and process-oriented components into models of diagnostic competence. The authors conclude with a discussion of important implications for teacher education and assessment practice.

Assessments in business and economics education are the focus of three further chapters in this part. In their chapter, *Threshold Concepts for Modeling and Assessing Higher Education Students' Understanding and Learning in Economics*, Sebastian Brückner and Olga Zlatkin-Troitschanskaia focus on student learning in economics and introduce a novel approach to modeling and measuring competences, which has great potential and is highly relevant particularly for process diagnostics over the course of studies.

While Brückner and Zlatkin-Troitschanskaia concentrate on the development of knowledge and domain-specific competences as student learning outcomes, in the next contribution, *Rescue an Enterprise from Failure: A Revolutionary Assessment Tool for Simulated Performance*, Fritz Oser, Susan Mueller, Tanja Obex, Thierry Volery, and Richard J. Shavelson present a novel approach to assessing performance in the area of entrepreneurship. They discuss measures for capturing the two competence constructs *sense of failure* and *rescue an enterprise from failure*, which are based on the initial validation results. They employ this innovative performance-oriented test instrument to measure entrepreneurial competence as a way to prevent entrepreneurial failure.

For decades, higher education in South Korea has had an excellent reputation. In the next chapter in this part, *Assessment of Economic Education in Korea's Higher Education*, Jinsoo Hahn, Kyungho Jang, and Jongsung Kim provide a comprehensive

overview of economics education and the established assessment practices in Korea, which originate in part from the United States or other countries and were adapted for the Korean context. The authors demonstrate a path toward increased international research in this area.

The part concludes with the chapter *What Do We Know – What Should We Know? Measuring and Comparing Achievements of Learning in European Higher Education: Initiating the New CALOHEE Approach* by Robert Wagenaar, who presents the new research initiative for assessing student learning outcomes in five domains at the European level. CALOHEE is developing the instrument's conditionalities for establishing cross-national diagnostic assessments, which can be applied Europe-wide. CALOHEE delivers three types of outcomes, which are outlined in this chapter: state-of-the-art reference points (benchmarks) for five academic sectors/subject areas, detailed assessment frameworks for these disciplines, and a multidimensional assessment model that does justice to the mission and profile of individual higher education institutions and degree programs.

The third part, **Generic Student Learning Outcomes – Cross-National Comparative Approaches**, encompasses contributions from the research area of interdisciplinary, generic student learning outcomes. One central innovative development in assessing generic student learning outcomes has been performance-oriented assessment. In the first chapter of this part, *International Performance Assessment of Learning in Higher Education (iPAL) – Research and Development*, Richard J. Shavelson, Olga Zlatkin-Troitschanskaia, and Julián P. Mariño present the conception of, rationale, and theoretical framework for this approach, which forms the basis of the new international research program iPAL.

The currently most widespread and most commonly used instrument for performance-based assessment of generic student learning outcomes in higher education is CLA+ by the Council for Aid to Education. It is the subsequent version of CLA, which was used in the AHELO study by the OECD. The next two chapters focus on the results of the validation studies and pilot implementations of CLA+ in many countries. In their chapter, *International Comparison of a Performance-Based Assessment in Higher Education*, Doris Zahner and Alberto Ciolfi present results from the validation of the tool for use in the United States and Italy. They conclude that this performance-based assessment enables comparative studies in higher education, and, therefore, international assessment of generic learning outcomes is feasible.

In the following chapter, *Adapting and Validating the Collegiate Learning Assessment to Measure Generic Academic Skills of Students in Germany – Implications for International Assessment Studies in Higher Education*, Olga Zlatkin-Troitschanskaia, Miriam Toepper, Dimitri Molerov, Ramona Buske, Sebastian Brückner, Hans Anand Pant, Sascha Hofmann, and Silvia Hansen-Schirra describe the adaptation and validation of CLA+ for use in Germany. The authors critically explore both the potentials and challenges of the implementation of this assessment instrument in higher education practice and the continuation of comparative studies based on the results of the validation study.

In their chapter *Validating the Use of Translated and Adapted HEIghten® Quantitative Literacy Test in Russia*, Ou Lydia Liu, Lin Gu, Prashant Loyalka, Amy Shaw, and Jane Wang present a validation study of another new assessment tool for generic learning outcomes. This test was developed in the United States and, like CLA+, has since been adapted for use in many countries. The authors provide an exemplary report on the validation of the Russian HEIghten Quantitative Literacy (QL) assessment with a representative group of students in Russia.

In addition to objective measurement of student learning outcomes, subjective measurement, for example, through self-reports, continues to be a research pillar in assessing generic student learning outcomes in higher education. In the final chapter in this part, *Comparative Study of Student Learning and Experiences of Japanese and South Korean Students*, Reiko Yamada introduces this type of assessment in the context of a comparative study of learning and experiences of students in Japan and South Korea. The findings indicate that student and faculty engagement variables appear to play important roles in the acquisition of knowledge and skills such as globalized skills, interpersonal skills, and cognitive ability.

Overall, this anthology introduces and explores important types of assessment in higher education as well as national and international developments. As a whole, this volume offers a broad overview of a relatively new field of research, which is of great significance for higher education, and demonstrates that more in-depth and extensive work in this field is necessary for developing appropriate approaches to assessing student learning in the twenty-first century. Along with other recent publications and with leading international studies in this field of research such as the AHELO study cited in these chapters, this volume offers a valuable foundation for further development of this emerging field.

This volume, which contains documentation on the current state of international research, would not have been possible without the tremendous collaboration of several researchers and experts from various disciplines and fields. We warmly thank all the authors for their active support and excellent contributions. We also thank all the reviewers and series editors for their extensive and helpful feedback and advice. Finally, we thank our graduate students at the University of Mainz for providing continuous support in preparing this volume, namely, Jennifer Fischer, Katja Kirmizakis, Mirco Kunz, and Mareike Magel.

Mainz, Germany

Olga Zlatkin-Troitschanskaia

Miriam Toepper

Christiane Kuhn

Berlin, Germany

Hans Anand Pant

Corinna Lautenbach

September, 2017

Contents

Part I Conceptual Development and Advances

- 1 **Research and Governance Architectures to Develop the Field of Learning Outcomes Assessment** 3
Hamish Coates
- 2 **Documenting and Improving Collegiate Learning in the USA** 19
Timothy Reese Cain and James C. Hearn
- 3 **Information Management Versus Knowledge Building: Implications for Learning and Assessment in Higher Education**..... 43
Patricia A. Alexander

Part II Domain-Specific Student Learning Outcomes – National and International Perspectives

- 4 **Challenges for Evaluation in Higher Education: Entrance Examinations and Beyond: The Sample Case of Medical Education** 59
Christiane Spiel and Barbara Schober
- 5 **Teachers' Judgments and Decision-Making: Studies Concerning the Transition from Primary to Secondary Education and Their Implications for Teacher Education** 73
Sabine Krolak-Schwerdt, Ineke M. Pit-ten Cate, and Thomas Hörstermann
- 6 **Threshold Concepts for Modeling and Assessing Higher Education Students' Understanding and Learning in Economics**..... 103
Sebastian Brückner and Olga Zlatkin-Troitschanskaia

7	Rescue an Enterprise from Failure: An Innovative Assessment Tool for Simulated Performance	123
	Fritz Oser, Susan Mueller, Tanja Obex, Thierry Volery, and Richard J. Shavelson	
8	Assessment of Economic Education in Korea’s Higher Education	145
	Jinsoo Hahn, Kyungho Jang, and Jongsung Kim	
9	What Do We Know – What Should We Know? Measuring and Comparing Achievements of Learning in European Higher Education: Initiating the New CALOHEE Approach	169
	Robert Wagenaar	
 Part III Generic Student Learning Outcomes – Cross-National Comparative Approaches		
10	International Performance Assessment of Learning in Higher Education (iPAL): Research and Development	193
	Richard J. Shavelson, Olga Zlatkin-Troitschanskaia, and Julián P. Mariño	
11	International Comparison of a Performance-Based Assessment in Higher Education.....	215
	Doris Zahner and Alberto Ciolfi	
12	Adapting and Validating the Collegiate Learning Assessment to Measure Generic Academic Skills of Students in Germany: Implications for International Assessment Studies in Higher Education.....	245
	Olga Zlatkin-Troitschanskaia, Miriam Toepper, Dimitri Molerov, Ramona Buske, Sebastian Brückner, Hans Anand Pant, Sascha Hofmann, and Silvia Hansen-Schirra	
13	Validating the Use of Translated and Adapted HEIghten® Quantitative Literacy Test in Russia.....	267
	Lin Gu, Ou Lydia Liu, Jun Xu, Elena Kardonova, Igor Chirikov, Guirong Li, Shangfeng Hu, Ningning Yu, Liping Ma, Fei Guo, Qi Su, Jinghuan Shi, Henry Shi, and Prashant Loyalka	
14	Comparative Study of Student Learning and Experiences of Japanese and South Korean Students.....	285
	Reiko Yamada	

Part I
Conceptual Development and Advances

Chapter 1

Research and Governance Architectures to Develop the Field of Learning Outcomes Assessment



Hamish Coates

Abstract This chapter articulates new research and governance architectures forming internationally to frame the future of learning outcomes assessment. It begins with a historical tour of the field, taking stock of the last 30 years by examining signature initiatives and geopolitical developments. The next section uses these foundations to extrapolate future technical, practical and substantive dimensions of a research framework. To give life to this framework, a picture is painted of what assessment will look like in a decade's time. The chapter's final section clarifies government arrangements, which would have the capacity to spur the kind of progress required to propel the field.

1.1 Introduction

This chapter articulates new research and governance architectures that are forming internationally and framing the future of student learning outcomes assessment. It takes stock of this establishing field and advances arrangements for spurring development. Important work must be done to reform assessment, and it is critical that researchers do not get lost into their own conversations or sidelined from the to and fro of major developments in policy and practice.

It seems common to hear in newspapers, reports and conferences that 'higher education is changing rapidly'. Higher education is becoming more central to socio-economic prosperity spurring intensification and proliferation of change (Coates 2017). Online technologies have undoubtedly changed access to much curriculum and the mechanics of much teaching, and more students than ever before move internationally to advance their academic and professional prospects. Interesting institutional variants are emerging, giving rise to new forms of governance and

H. Coates (✉)
Tsinghua University, Beijing, China
e-mail: hamishcoates@tsinghua.edu.cn

commercial opportunity. Collaborative technologies have fundamentally reformed core facets of much research and changed the way research is created, constructed and disseminated. At an aggregate level, change does indeed abound. On the ground, change is more complex and slower than espoused.

Yet as observation at almost any higher education reveals, much of how teachers and institutions assess student learning has not changed for over a century. Within changing institutional settings, new generations of faculty are of course interacting with new computing technologies to provide diverse students with information and experiences intended to prepare students with capabilities for tomorrow's world of professional work. Given such change, it is surprising that much assessment in higher education has not changed materially for a very long time and that economically and technically unsustainable practice is rife. As other chapters in this book affirm, there are an enormous number of innovative and high-quality developments, including those associated with technology advances (e.g. Shavelson et al., Chap. 10 in this volume). Still, every day around the world, students write exams using pen and paper, sitting without talking in large halls at small desks in rows. It is possible that this reflects the pinnacle of assessment, but given the lack of reflective advance over an extended period, this seems unlikely. Rather, given the enormous changes reshaping core facets of higher education, and pressures and prospects surrounding assessment, it is more likely that the 'transformational moment' is yet to come.

The assessment of higher education student learning outcomes is very important. Assessment provides essential assurance to a wide variety of stakeholders that people have attained various knowledge and skills and that they are ready for employment or further study. More broadly, assessment signposts, often in a highly distilled way, the character of an institution and its educational programmes. Much assessment is expensive, making it an important focus for analysis. Assessment shapes education and how people learn in powerful direct and indirect ways, influencing teaching and curriculum. Assessment is highly relevant to individuals, often playing a major role in defining life chances and directions (see also Cain and Hearn, Chap. 2 in this volume).

This chapter is posed at a formative time of the development of this field. As the field of higher education assessment research and reform takes shape, it is timely to step back and examine the broader developments in play, how to structure an understanding of emerging trends and what kind of governance arrangements would help transfer research into practice. This chapter tackles each of these areas, along the way articulating possible futures for assessment in higher education.

1.2 Signature Developments in Recent Decades¹

Assessment has forever played an integral role in higher education, but the most relevant antecedents for analysing contemporary development can be traced back over the past few decades (see also Cain and Hearn, Chap. 2 in this volume). This

¹The following text builds on Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment and Evaluation in Higher Education*, 41(5), 662–676 (Taylor and Francis).

section examines signature initiatives and how these have been shaped by various geopolitical developments. Clearly, taking critical stock of a field as large and diverse as higher education assessment is a useful, though challenging task—there are an enormous number of actors and initiatives, each at varying stages of maturity and diffusion. Rather than conduct an exhaustive review, it is feasible to conduct a review of a series of signature case studies, which have sought to shift policy and practice.

One broad line of development has involved specifying qualification-level outcomes. Examples include the European Qualifications Framework (European Commission (EC) 2015), the UK's Qualifications and Credit Framework (Ofqual 2015), the Australian Qualifications Framework (AQFC 2015) and the US Degree Qualifications Profile (Lumina Foundation 2015). As such titles convey, this work is developed and owned by systems, and such initiatives have served as important policy instruments for shifting beyond an anarchic plethora of qualifications, generating conversations about finding more coherence and indeed articulating the general outcomes graduates should expect from a qualification (Chakroun 2010). These system-wide structures can suffer from unhelpful collisions with fruitfully divergent local practice, but their inherent constraint is that they go no further than articulating only very general graduate outcomes (Allais et al. 2009; Wheelahan 2009). They offer little beyond broad guidelines for improving the assessment of student learning.

Going one step deeper, another line of work has sought to specify learning outcomes at the discipline level. The tuning process (González and Wagenaar 2008) is a prominent example which has been initiated in many education systems and across many diverse disciplines. Broadly, tuning involves supporting collaboration among academics with the aim of generating convergence and common understanding of generic and discipline-specific learning outcomes (see also Wagenaar, Chap. 9 in this volume). Canada adapted this work in an innovative way, focusing the collaborations around sector-oriented discipline clusters rather than education fields (Lennon et al. 2014), while in Australia a more policy-based and compliance-focused approach was deployed (Australian Learning and Teaching Council (ALTC) 2010). Such collaboration travels several steps further than qualification frameworks by engaging and building academic capacity within disciplinary contexts. Like the qualification frameworks, however, the work usually stops short of advancing assessment resources or sharing data and tends to focus instead on advancing case studies or best practice guidelines.

A slightly deeper line of development involves shared rubrics to compare assessment tasks or student performance. Moderation in assessment can play out in many ways (Coates 2010) as indeed has been the case in recent higher education initiatives. The moderation of resources has involved rudimentary forms of peer review through to slightly more extensive forms of exchange. Mechanisms have also been developed to help moderate student performance. In the United States, for instance, the Association of American Colleges and Universities (AAC&U) (Rhodes and Finley 2013) has developed VALUE (Valid Assessment of Learning in Undergraduate Education) rubrics for helping faculty assess various general skills. This has been

progressed in most recent cross institutional moderation work (AAC&U/State Higher Education Executive Officers (SHEEO) 2015; see also Cain and Hearn, Chap. 2 in this volume). The UK's external examiner system (Quality Assurance Agency (QAA) 2014) is a further example. Several such schemes have been launched in Australia, including a Quality Verification System and a Learning and Teaching Academic Standards Project, both of which involve peer review and moderation across disciplines (Marshall et al. 2013). This work travels more widely than qualification- or discipline-level specifications, for it involves the collation and sharing of evidence on student performance, often in ways that engage faculty in useful assurance and development activities. Such moderation work is limited, however, in being applied in isolation from other assessment activities and materials.

Collaborative assessments build from the developments discussed so far to advance more coherent and expansive approaches to shared assessment. As with other developments addressed here, such work plays out in myriad ways. For instance, medical progress testing in the Netherlands (Schuwirth and van der Vleuten 2012) involves the formation of shared assessment materials and administration of these in a longitudinal sense (for medical education, see also Spiel and Schober, Chap. 4 in this volume). Other assessment collaborations have focused on the development of shared tasks and analytical or reporting activities, for instance, the Australian Medical Assessment Collaboration (AMAC) (Edwards et al. 2012) and the German initiative showcased in other chapters with the umbrella title Modeling and Measuring Competencies in Higher Education (KoKoHs) (Zlatkin-Troitschanskaia et al. 2014, 2017). In 2015, the Higher Education Funding Council for England (HEFCE) funded a suite of mostly collaborative projects to assess learning gains in higher education (HEFCE 2015), and the European Commission funded a large-scale collaboration titled Measuring and Comparing Achievements of Learning Outcomes in Higher Education in Europe (CALOHEE) (EC 2015). Such work is impressive as it tends to involve the most extensive forms of outcome specification, task production, assessment administration, analysis and reporting and at the same time develop faculty capacity. Work plays out in different ways, however, shaped by pertinent collegial, professional and academic factors. This can mean, for instance, that extensive work is done that leads to little if any benchmarking or transparent disclosure.

Standardised assessment is easily the most extensive form of development and would appear to be growing in scope and scale. Licencing examinations are the most long-standing and pervasive forms of assessment, though their use is cultural and they tend to be far more common in the United States than Europe (see also Cain and Hearn, Chap. 2 in this volume). Other related kinds of national effort are evident in certain countries, for instance, in Brazil (Melguizo 2015), Colombia (Shavelson et al. 2016) and the United States (Shavelson 2007; Educational Testing Service (ETS) 2014). A series of international graduate outcome tests have also been trailed in recent years, such as the OECD's Assessment of Higher Education Learning Outcomes (AHELO) (Coates and Richardson 2012), the International Association for the Evaluation of Education Achievement Teacher Education and Development Study (IEA TEDS) assessment of teachers (Braeken and Blömeke

2016), the HEIghten assessment (Liu et al. 2016; see also Liu et al., Chap. 13 in this volume), a cross national assessment of engineering competence (Loyalka 2015), and the Collegiate Learning Assessment (CLA) (CAE 2016; see also Zahner and Ciolfi, Chap. 11 in this volume). Standardised assessments are also promulgated via commercial textbooks (Pearson 2014). As implied by the term ‘standardised’ and by the external sponsorship of such work, such assessment often proceeds without engaging with academics. Though such exogenous intervention may in the longer run inject the shock required for assessment reform, it also tends to balkanise internal from external interests and has not yet been shown to have large impact on learning or teaching practice.

A variety of these practices are used by higher education providers around the world, but it must be said, in varying and inconsistent ways. Traditional assessment practices are rife in older more established universities. Such practices are baked into academic policy and procedures and more particularly into well-tenured workforces. More recently established institutions have the opportunity to leapfrog and set up more modern approaches to education design, which better express contemporary assessment ideas and practices. It is surprising that tertiary institutions are not playing a greater leading role in assessment reform given their involvement in such work over centuries. A review of the initiatives discussed above shows that most institutions are participating at the margins or in spasmodic ways, with many yet to embrace comprehensive assessment reform.

1.3 A Framework to Structure Future Trends

What do these developments over the past few decades, but particularly in recent years, tell us about the shape of things to come? This section advances a normative framework that can be used to extrapolate future technical, practical and substantive research trends. This three-dimensional framework is proposed as a mechanism for advancing principles for reforming the field of learning outcomes assessment. These dimensions are described, and the framework’s value is teased out via a number of illustrative change areas. One dimension of this framework divides change into those aspects which are substantive in nature, another which are technical in nature and the third which are practical in nature.

Substantive—policy, disciplinary and conceptual—considerations are the most significant forces shaping learning assessment. Assessment is of little use unless it is relevant to students, to policymakers, to institutional leaders and managers, to academics or to the general public. Establishing such relevance is tricky, as it involves not just identifying but also then defining what counts, and of course stakeholder interests play a role in this. Power plays a key role that manifests through the formal or informal authority of individuals or institutions. It is not uncommon to hear of conflicts regarding what should be assessed between educators, professional associations and industry or accreditation agencies. More broadly, the oligopolistic character of most established higher education systems has limited the extent to

which change has been driven by research and technological development, though appetite for research-driven change appears to be increasing with the increasingly competitive nature of higher education markets.

From a normative perspective, though evidently not always in practice, it is imperative that assessment is cogent technically. This means that assessment resources and approaches should aim to be valid and measure and report what is intended. Assessment should be reliable, which means that assessment should provide consistent measurement of the target focus area. There are a host of methods for assessing and reporting these kinds of technical properties, which of course are the focus of active scientific debates within specific communities. At a minimum, it might be expected that explicit consideration has been given to measurement considerations, but ideally a set of statistics should be provided as with professionally validated assessment instruments. Students and other key stakeholders have a right to know that assessment is producing information which pertains to people's competence in the measured area as opposed to measurement noise.

Substantive relevance and technical integrity are not sufficient to spur change in assessment. Practice is critical in that it must be feasible to collect, analyse and report data. Though institutional budgets are getting tighter, many entrenched assessment methods have high fixed costs and limited economies of scale. It is vital that more viable options are explored. Really important changes in assessment might be costly or slow to deliver. They may waste students' time and hinder learning experiences and outcomes. Indeed, such practical constraints are often claimed as impediments to progress. What matters is not just only fixed start-up costs but also ongoing costs of deployment over a prescribed time period. In building financial equations, decisions must be made about which costs are direct and indirect. A key reason for resisting change may well be that much of the cost of current assessment approaches is hidden within undifferentiated faculty roles. But the opaque nature of such costing does not make it cheap. Rather, the lack of scientific management of assessment implies all sorts of inefficiencies and scope for improvement. Of course, cost is not the only practical facet of assessment though it offers a means for summarising important decisions and uncertainties.

Each of these three dimensions plays out at varying levels. The Organisation for Economic Cooperation and Development (OECD) (2015, p. 15) notes the importance of 'distinguish[ing] between the actors in education systems: individual learners and teachers, instructional settings and learning environments, educational service providers, and the education system as a whole'. The level at which information is reported is not the same as the level at which information is collected (typically the student with assessment). Data is often collected at a lower level and then aggregated and often also combined with other data for reporting. Similarly, the interpretation level might be different again and will likely vary with the interests and concerns of stakeholders. Many current institution rankings, for instance, aggregate information on individual researcher performance and report this at the institution level, and then the information is interpreted in all sorts of ways, including in relation to fields of education. Assessment change is required for those involved in education such as students and teachers, and that change is required by broader

communities, including the general public, business and industry and people associated with planning education policy and strategy.

A series of framing ideas can be evoked from this three-dimensional framework. Substantively, it is important for assessment to be relevant or authentic to students and teachers. This often means that a diversity of assessment practice is required. At the same time, stakeholders more removed from everyday practice seek evidence which is more general in nature. Hence, a substantive idea which might be derived is that future reform should ensure that assessment is locally relevant and externally generalisable. A technical idea is that reform should advance transparency regarding the validity and reliability of assessment. The most well-designed and validated assessments are meaningless unless they are feasible to implement. Hence, a further idea for reform is that assessment must make efficient use of money and time. In terms of practice, emphasis might be placed on delivering feasible and efficient assessment to large student cohorts given tight resource constraints, whereas those more removed from the process may give more regard to the technical veracity of the evidence produced. Stereotypical remarks made by employer groups can suggest a lack of confidence in the everyday assessment by institutions of students' knowledge and skills.

Such ideas could be nuanced differently or elaborated more exhaustively, but the above formulations are sufficient to tease out the main points at play. None of the three dimensions or the kinds of ideas that they motivate are particularly surprising or controversial, though they can provoke substantial complexity and be difficult to implement. Part of the trouble arises from the conundrums provoked by attempts to harmonise or jointly optimise the dimensions in unison. Further trouble flows from negotiating the dialectic between internal and external interests. Broader considerations flow from complexities associated with generalising the assessment of complex higher-order skills across national and cultural contexts. Resolving such issues offers a chance to unlock substantial progress in the assessment of student learning outcomes. Hence, the dimensions provide a useful normative rubric against which to evaluate current progress and change dynamics and to forecast insights and frontiers for reform.

1.4 Guiding Transformation into Practice

What do the normative framework and the earlier insights regarding signature initiatives convey in terms of progress and strategies for future development? This section deploys the framework to take brief stock of the emerging field before turning to focus on potential steps ahead. A picture is painted of prospects for future development, building on earlier analyses of Coates (2014) and researchers from many different systems.

What evidence is there that the current initiatives are helping to ensure that assessment is locally relevant and externally generalisable? Large-scale qualification frameworks do very little to achieve this, but the more practice-focused initia-

tives do appear to be driving progress in this direction. The suite of programmes that invite academics to focus on organising their practice around more generalisable principles—such as the tuning process and the VALUE rubrics—provides signposts for change (see also Cain and Hearn, Chap. 2 in this volume). Simultaneously, the externally driven initiatives are themselves benefitting from technological advances in assessment (Bennett 2015) which give new insights into what assessment can look like and deliver. There would appear to be some way to go, however, in transcending the internal/external dialectic that appears to simultaneously spur and hinder progress. As well, there is much work to be done to bridge the reducing but still large gap between large-scale policy and technical development and everyday practice.

On the technical front, is progress regarding outcomes assessment advancing transparency regarding the validity and reliability of assessment? Diversified and large-scale initiatives such as KoKoHs reveal the extent of work required to validate higher education assessment (e.g. Zlatkin-Troitschanskaia et al., Chap. 12 in this volume). Larger-scale assessments such as AHELO illustrate the extent of technical validation and transparency that can be achieved. But with the workforce capacity bottlenecking progress even with funded large-scale initiatives, it is unlikely that teaching institutions will be positioned anytime soon to validate student assessment in ways that meet address psychometric standards and criteria. There is a high risk that students are assessed using insufficiently validated tasks which yield spurious information about performance and potential (Coates 2015). Inadequate information of this kind carries risk for graduates and also for institutions and countries. The need for development in this front is important, and a suggestion is given below.

Are contemporary advances helping to enhance the efficiency of assessment? Given that most large-scale initiatives tend to be relatively expensive (e.g. crude estimates of around \$10,000 per finished professionally produced multiple-choice item are not uncommon), such work is itself unlikely to be offering any intrinsic signals for how to make assessment more efficient. Through innovation, however, large-scale initiatives do carry potential to initiate new technologies and approaches and to model how new efficiencies may be achieved. The risk, of course, is that change is shaped more by factors, which are exogenous to assessment. That is, given ambiguous budget constraints and unclear technical and substantive expectations, explicitly identifiable assessment costs become a real target for savings, particularly compared with more visible staffing resources and facilities. The consequence of such disinvestment is obvious—cheaper and lower quality forms of assessment will be used that are less authentic and robust. Understanding the trade-offs linked with differential levels of direct and indirect resourcing is important, which hinges on the kind of productivity evaluations exemplified via the National Center for Academic Transformation (NCAT) course redesign initiatives (Twigg 2003). Broadly, it could be expected that in any decomposition of assessment costs it is development and implementation as opposed to the planning, analysis and reporting phases, in which new techniques carry potential to spur new economies of scale.

What does this albeit brief stocktake imply about the most fruitful areas to target reform? Where should further energy be directed to optimise substantive, technical

and practical matters and do as much as possible to address tensions associated with the internal/external dialectic? How can such energy most effectively navigate the change dynamics noted above? From the above analyses, recommendations can be made for focusing future development.

Seamless assessment tasks and processes must be prepared which can jointly serve the needs of internal and external stakeholders. This might involve production of resources, which can be shared across boundaries perhaps via adaptation to different disciplinary, professional or cultural contexts, or it might involve embedding more generic materials within local assessments. Several of the initiatives reviewed above have progressed such options. They have identified ways for harmonising the production and delivery of materials drawn from different sources, for integrating processes and for using more professionally developed materials to seed change in local practice.

Further work should be invested into techniques that engineer validity into assessment development. Rather than defer to post hoc evaluation by assessment experts, the quality of assessment is most likely to be improved by intervening earlier in the development cycle to ensure that materials exceed minimally sufficient technical standards in the first place. A specific example includes larger use of principled assessment design frameworks that help scale up assessment, so that assessment creation can be better aligned with standards, connected with learning sciences, more efficiently implemented for scaling up technologically and with conceptual frameworks suited to tailoring to local needs within a broader framework (Mislevy et al. 2011; Luecht 2013). Any such development hinges obviously on a set of accepted standards and on an effective means for bringing such standards into play. Internationally, standards do exist (e.g. American Educational Research Association (AERA), American Psychological Association (APA) and National Council for Measurement in Education (NCME) 2014; International Testing Commission (ITC) 2015), and higher education institutions have a range of means for governing the incorporation of these into academic practice. While mention of the word ‘standards’ in higher education can provoke debates about standardisation and regression to the mean (Coates 2010), there would appear to be value in progressing such work if it places a transparent floor around the quality of assessment.

Technology-assisted collaboration and delivery has an important role to play in improving practice. As Bennett (2015) conveys, by affording rethinking of task design and delivery, it also provides a frame for advancing the substantive and technical frontiers. Technology-assisted collaboration is important as this is a major means for making assessment more transparent and enhancing quality and productivity. Peer and stakeholder review of tasks helps to iron out glitches and improve authenticity. Professional capacity is developed through feedback. Technical standards could be embedded in design architectures. Sharing design and development also reduces redundancy, duplication and expensive fixed costs associated with resource production (Coates 2015). As well, with appropriate security solutions now available (Richardson and Coates 2014), it is feasible to shift from paper to online administration and reap derivative efficiencies in time and money. Of course, many platforms exist and are available from collegial or commercial sources

(e.g. Dillon et al. 2002; National Board of Medical Education (NBME) 2008; Cisco Networking Academy 2012). The key is to marry these with enterprise-level learning systems, which have scaled to ubiquity over the last decade. Put simply, such technologies should distil insights from measurement science into systems that make good assessment easy for time-poor and non-expert academics.

There is much to be considered regarding the propagation of powerful technologies into higher education. As cautioned in 2005 in relation to the rapid expansion of enterprise-level learning management systems (Coates et al. 2005), such change should not be led by technozealots who see information systems as a panacea but rather by educational leaders who can shrewdly leverage technology for change. Among other matters, it is imperative to consider the influence of systems on teaching and learning, the uncertain effects on students' engagement, the new dynamics in academic work and the organisation of teaching and the possible corporatisation of academic knowledge. As Bennett (2015) contends, most value is to be had by exploiting the sophistication of 'third-generation' technology-based assessment. This involves not just transferring traditional paper-based batch processes to computer (first generation) or incremental improvement in quality and efficiency (second generation) but fundamental redesign which ensures that assessment serves both individual and institutional needs, is informed by cognitive principles, is a medium for enabling natural interaction with rich assessment tasks and is technically enabled to support and enhance curriculum and teaching.

1.5 Governance to Spur Progress²

The preceding sections set out a normative framework and used this to frame the past and chart future assessment initiatives. What are the most effective ways of translating this work into practice? This chapter's final section articulates governance options for spurring the kind of progress required to propel the field. In particular, it advocates for more open forms of assessment, even in the most confidential and secure fields of assessment.

Analysing how best to translate research into practice can be done in a variety of ways, and the most significant considerations go to governance. New ways of designing and managing academic work, including assessment, will almost certainly require new forms of governing academic activity, power and performance (Shattock 2012). While assessment is experienced mostly as a practical educational matter, it touches many facets of higher education leadership and management. Indeed, assessment goes right to the heart of important aspects of governance such as ownership, authority and power. The risks of poorly designed or conducted governance, and the need to get governance right, show up in sectoral or organisational failures. The governance of assessment is both critical and problematic.

²The following text builds on Canny, B. & Coates, H. (2014). *Governance Models for Collaborations Involving Assessment*. Sydney: Office for Learning and Teaching.

An important feature of higher education sector is the self-accrediting status of most institutions, particularly universities. Typically, some regulatory power—typically government—delegates institutions the authority over academic programmes of course including assessment. This continues tradition that stems from the origins of the university system in mediaeval Italy and England. It does, however, create a natural tension inasmuch that universities are funded by the public purse and produce graduates for the community's benefit, but the community may not have a direct involvement in assessing the relevance and standards of education. Managing public versus private authority is an important thread running through debates about assessment reform. What spectrum of governance arrangements might be considered?

The assessment of student learning is done in myriad ways within universities. Obviously, quite a lot of assessment is done by individual academics working alone within single institutions. Alternatively, assessment can be done by groups of academics within a single institution. In each of these cases, accreditation by a government or industry authority vests power in an institution's academic board which devolves power to individual academics. The situation in practice is far more complex than suggested by this straightforward chain of command, with academics drawing on all kinds of more or less indirect and informal networks. In key respects the quality and economics of this collegial fabric are hard to beat, but at the same time, its informal and elite nature falls short. Typically, there is loose institutional oversight, academics flying solo and deployment of non-validated materials using dated practical and technical approaches.

Alternatively, assessment can be enacted and governed by groups of academics across institutions, almost invariably but not necessarily within the same discipline or professional field. Academics collaborate in this way routinely in their research work—forming collaborations and networks to design, execute and publish work. It is reasonably common for academics across institutions to share teaching, perhaps to service particular knowledge needs or to diversify teacher perspectives and student experiences. It remains far less common, however, for such collaborations to spill over into assessment. A few reasons for this have been sketched above—such as security, confidentiality and privacy—and there are doubtless others that go to individual and institutional commercial and reputational factors. Operating between institutions also carries governance implications, inasmuch as the collaboration space lies strictly beyond the jurisdiction of any single institution's reach. These implications are addressed below via the proposed academic governance model.

Assessment may also be governed from outside institutions. This work may involve academics working with third-party organisations or third-party organisations working alone. This work may take place on university campuses, or it may be outsourced to collaborative academies or statutory bodies empowered to perform specific functions for community benefit like licencing and credentialing. Such external governance is reasonably common with admissions or licencing examinations but quite rare for in-course assessment even in highly regulated fields. The delegation of assessment in this way raises even more substantial governance considerations. For instance, what are the governance arrangements of these external

organisations? Why might oversight by a private testing firm or accreditation body be preferable to that provided by a university's academic board composed of potentially hundreds of experts? How are faculty engaged? How is the authenticity of materials assured? What controls are in place to minimise the duplication of effort between faculty and external agencies?

Figure 1.1 captures this spectrum of governance arrangements sketched above. These range from individuals working alone to academics collaborating among themselves and with other agencies to fully external arrangements. As with all models, this is an abstraction, but it is helpful in clarifying the main options at play.

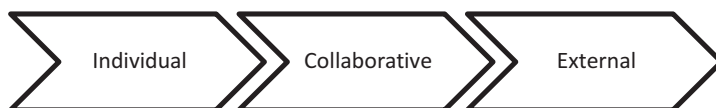


Fig. 1.1 Spectrum of assessment governance arrangements

These simplistic arrangements of course play out in an infinitely complex array of ways. Today's tertiary institutions are reforming in many areas, adopting a variety of approaches and spawning a proliferation of organisational forms of work (Coates 2017). What seems sure in all of this complexity is that the business/academic model of (in general terms) 'faculty working alone in isolation' is moving more towards 'faculty playing a role in a broader team'. Where new forms of academic infrastructure are emerging, these are typically affiliated or owned by existing institutions. In essence, renovating/improving rather than replacing existing academic arrangements seems to be the most common change underway. Shifting to more collaborative forms of assessment governance would do well to reflect such transition.

The 'sharing economy'—or perhaps in higher education the 'collaboration economy'—is reshaping many facets of economic and social life, and higher education is no exception. Rather than goods and services being created and used by individuals in isolation, teachers and learners are collaborating via advanced online systems to generate new ways of doing education. Teachers and institutions are collaborating on curriculum production, learners are collaborating on assignments and open admissions, provision and credit recognition are touching basic notions of the qualification. As signalled at the outset, the assessment component of education has been one of the most resistant areas to adapt to the changing environment. In many areas assessment is closely tied via content and implementation to local educational settings. It has obvious security, confidentiality and privacy aspects. As the tool for evaluating individual performance, it also helps measure the quality of programmes and institutions and through this carries reputational and commercial implications. For these and other reasons, assessment would appear to be one of the final frontiers in the contemporary unbundling of higher education.

By taking stock of recent signature developments and painting a picture of future practice, this chapter has advanced a framework for thinking through the growth of the field and productive new forms of governance. Technology almost certainly will

play a role in changing practice, but it is essential that effective governance architectures are in place. This chapter has advocated reform by strengthening and augmenting rather than replacing traditional collegial arrangements. Most particularly, the paper has advocated for the value of moving to more collaborative kinds of governance. Even—and perhaps especially—the most high-stakes assessment needs to become more open to improve.

References

- Allais, S., Young, M., & Raffe, D. (2009). Introduction. In S. Allais, D. Raffe, R. Strathdee, L. Wheelahan, & M. Young (Eds.), *Learning from the first qualifications frameworks*. Geneva: International Labour Office.
- American Educational Research Organisation (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for educational & psychological testing*. Washington, DC: American Psychological Association.
- Association of American Colleges and Universities (AAC&U), & State Higher Education Executive Offices (SHEEO). (2015). *MSC: A multi-state collaborative to advance learning outcomes assessment*. Accessed 20 Sept 2016 from: <http://www.sheeo.org/projects/msc-multi-state-collaborative-advance-learning-outcomes-assessment>
- Australian Learning, & Teaching Council (ALTC). (2010). *Learning and teaching academic standards project final report*. Sydney: Australian Learning and Teaching Council.
- Australian Qualifications Framework Council (AQFC). (2015). *Australian qualifications framework*. Accessed 1 Sept 2016 from: <http://www.aqf.edu.au>
- Bennett, R. (2015). The changing nature of educational assessment. *Review of Research in Education*, 39(1), 370–407.
- Braeken, J., & Blömeke, S. (2016). Comparison of beliefs across countries: Dealing with measurement invariance and local dependence using Bayesian elastics. *Assessment and Evaluation in Higher Education*, 41(5), 733.
- Canny, B., & Coates, H. (2014). *Governance models for collaborations involving assessment*. Sydney: Office for Learning and Teaching.
- Chakroun, B. (2010). National qualification frameworks: From policy borrowing to policy learning. *European Journal of Education*, 45(2), 199–216.
- Cisco Networking Academy. (2012). *Advancing assessment with technology*. Accessed 2 Dec 2016 from: http://www.cisco.com/c/dam/en_us/training-events/netacad/downloads/pdf/NetAcadPOV.pdf
- Coates, H. (2010). Defining and monitoring academic standards in Australian higher education. *Higher Education Management and Policy*, 22(1), 1–17.
- Coates, H. (2014). *Higher education learning outcomes assessment: International perspectives*. Frankfurt: Peter Lang.
- Coates, H. (2015). Assessment of learning outcomes. In R. Pricopie, P. Scott, J. Salmi, & A. Curaj (Eds.), *Future of higher education in Europe. Volume I and volume II*. Dordrecht: Springer.
- Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment and Evaluation in Higher Education*, 41(5), 662–676.
- Coates, H. (2017). *The market for learning: Leading transparent higher education*. Dordrecht: Springer.
- Coates, H., & Richardson, S. (2012). An international assessment of bachelor degree graduate's learning outcomes. *Higher Education Management and Policy*, 23(3), 51–69.

- Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary Education and Management, 11*, 19–36.
- Council for Aid to Education (CAE). (2016). *Collegiate Learning Assessment (CLA)*. New York: CAE.
- Dillon, G. F., Clyman, S. G., Clauser, B. E., & Margolis, M. J. (2002). The introduction of computer-based case simulations into the United States medical licensing examination. *Academic Medicine, 77*(10), 94–96.
- Educational Testing Service (ETS). (2014). *Proficiency profile*. Accessed 1 Sept 2014 from: www.ets.org/proficiencyprofile/about
- Edwards, D., Wilkinson, D., Coates, H., & Canny, B. (2012). *The Australian medical assessment collaboration: Developing the foundations for a national assessment of medical student learning outcomes*. Sydney: Office of Learning and Teaching.
- European Commission (EC). (2015). *European qualifications framework*. Accessed 1 Sept 2016 from: https://ec.europa.eu/ploteus/search/site?f%5B0%5D=im_field_entity_type%3A97
- González, J., & Wagenaar, R. (2008). *Universities' contribution to the Bologna process: An introduction*. Bilbao: Universidad de Deusto.
- Higher Education Funding Council for England (HEFCE). (2015). *£4 million awarded to 12 projects to pilot measures of learning gain*. Accessed 1 Sept 2016 from: <http://www.hefce.ac.uk/news/newsarchive/2015/Name,105306.en.html>
- International Testing Commission (ITC). (2015). *Guidelines*. Accessed 10 Sept 2016 from: <https://www.intestcom.org/page/5>
- Lennon, M. C., Frank, B., Lenton, R., Madsen, K., Omri, A., & Turner, R. (2014). *Tuning: Identifying and measuring sector-based learning outcomes in postsecondary education*. Toronto: Higher Education Quality Council of Ontario.
- Liu, O. L., Mao, L., Frankel, L., & Xu, J. (2016). *Assessing critical thinking in higher education: The HEIghten™ approach and preliminary validity evidence*. Princeton: Educational Testing Service.
- Loyalka, P. (2015). *Initial results from an international study of student learning in higher education: China, Russia, and the U.S.* Keynote paper presented at the 2nd symposium of learning science and online education, China, Beijing.
- Luecht, R. M. (2013). An introduction to assessment engineering for automatic item generation. In M. J. Gierl & T. M. Haladyna (Eds.), *Automatic item generation: Theory and practice*. New York: Routledge.
- Lumina Foundation. (2015). *The degree qualifications profile*. Accessed 1 Sept 2016 from: <http://www.luminafoundation.org/files/resources/dqp.pdf>
- Marshall, S., Henry, R., & Ramburuth, P. (2013). *Improving assessment in higher education: A whole of institution approach*. Sydney: New South Books.
- Melguizo, T. (2015). *Are students gaining general and subject area knowledge in university? Evidence from Brazil*, Higher Education.
- Mislevy, R., Haertel, G., Yarnall, L., & Wentland, E. (2011). Evidence centered task design in test development. In C. Secolsky (Ed.), *Measurement, assessment, and evaluation in higher education*. New York: Routledge.
- National Board of Medical Examiners (NBME). (2008). *Primum® Computer-based Case Simulations (CCS)*. Philadelphia: NBME.
- Ofqual. (2015). *Qualifications and credit framework*. Accessed 1 Sept 2016 from: <https://www.gov.uk/government/organisations/ofqual>
- Organisation for Economic Cooperation and Development (OECD). (2015). *Education at a glance*. Paris: OECD.
- Pearson. (2014). *MyEconLab*. Accessed 27 Aug 2014 from: www.pearsonmylabandmastering.com/northamerica/myeconlab
- Quality Assurance Agency (QAA). (2014). *The UK quality code for higher education*. Retrieved from: www.qaa.ac.uk

- Rhodes, T., & Finley, A. (2013). *Using the VALUE rubrics for improvement of learning and authentic assessment*. Washington: AAC&U.
- Richardson, S., & Coates, H. (2014). Essential foundations for establishing equivalence in cross-national higher education assessment. *Higher Education*, 68(6), 825–836.
- Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2012). The use of progress testing. *Perspect Medical Education*, 1(1), 24–30.
- Shattock, M. (2012). University governance. *Perspectives: Policy and practice in higher education*, 16(2), 56–61.
- Shavelson, R. J. (2007). *A brief history of student learning assessment: How we got where we are and a proposal for where to go next*. Washington, DC: Association of American Colleges and Universities.
- Shavelson, R. J., Domingue, B. W., Mariño, J. P., Molina-Mantilla, A., Morales, J. A., & Wiley, E. E. (2016). On the practices and challenges of measuring higher education value added: The case of Colombia. *Assessment and Evaluation in Higher Education*, 41(5), 695–720.
- Twigg, C. A. (2003, September/October). Improving learning and reducing costs: New models for online learning. *EDUCAUSE Review*, 38(5), 28–38.
- Wheelahan, L. (2009). From old to new – The Australian qualifications framework. In S. Allais, D. Raffe, R. Strathdee, L. Wheelahan, & M. Young (Eds.), *Learning from the first qualifications frameworks*. International Labour Office: Geneva.
- Zlatkin-Troitschanskaia, O., Kuhn, C., & Toepper, M. (2014). Modelling and assessing higher education learning outcomes in Germany. In H. Coates (Ed.), *Advancing higher education learning outcomes*. Frankfurt: Peter Lang.
- Zlatkin-Troitschanskaia, O., Pant, H. A., Toepper, M., Lautenbach, C., & Molerov, D. (2017). Valid competency assessment in higher education: Framework, results, and further perspectives of the German Research Program KoKoHs. *AERA Open*, 3(1), 1–22.

Chapter 2

Documenting and Improving Collegiate Learning in the USA



Timothy Reese Cain and James C. Hearn

Abstract This chapter examines the expansion and current features of learning outcomes assessment in the US higher education. It begins with an overview of the diverse and stratified higher education system, including the multiple layers of state and federal influence but lack of centralized tightly linked control mechanisms. It describes the changing context of, and increased interest in, learning outcomes in the late twentieth and early twenty-first centuries, including the roles of various external actors. In so doing it highlights the regional and national accrediting bodies that serve as gateways to federal financial aid funds and therefore serve as important means of affecting change. Although assessment is important for institutional accountability, its true value lies in its as-yet-unrealized potential to fundamentally improve the teaching and learning at the campus level. As such, the chapter concludes with considerations of on-campus actors at the heart of assessment processes and the evidence of changing practices that have been revealed through recent national surveys by research and voluntary membership groups.

2.1 Introduction

Over the past four and a half decades, higher education stakeholders in the USA have become increasingly concerned about what students know and what abilities they possess as a result of their college educations. Pushed by external demands for accountability and internal concerns about learning, outcomes assessment has become an important consideration at institutional, state, and federal levels—along with equity, safety, and completion, it is one of the most important concerns about college students in the second decade of the twenty-first century. At the same time, both the idea and practice of assessment have remained controversial. Many question whether it is even possible to measure the most important student outcomes and whether attempts to do so might change priorities and affect

T. R. Cain (✉) · J. C. Hearn
University of Georgia, Athens, GA, USA
e-mail: tcain@uga.edu; jhearn@uga.edu

activities in ways that are detrimental to student learning. Some link the practice to a growing corporatization of higher education, where faculty authority is being replaced by administrative oversight. Others are more confident that we can document learning but worry that we largely do so to meet accountability requirements, rather than using the results to effect change in student learning, thereby fulfilling assessment potential. Yet, even with these concerns, student learning outcomes remain vital to understanding and improving American colleges and universities.

In this chapter, we consider student learning outcomes assessment in the USA, emphasizing the key actors and the major issues with which they are engaging. We begin with a brief overview of the diversified system of American higher education, highlighting the different governance structures, institutional types, and current concerns that provide context for understanding how assessment has been shaped and operates. We then chart a history of the assessment movement in the USA, from its beginnings through the modern era, emphasizing the evolution of what was once seen as a fad to what has become a central facet of accountability mechanisms and key, if contested, campus concern. We turn next to three major external constituent groups that are influencing assessment: governmental actors; accrediting agencies that mandate that colleges and universities report on their students' learning; and the membership associations and foundations that are driving much of the conversation about assessment. The heart of assessment is what happens on college campuses, where students, faculty, administrators, and others interact—or in many cases do not—around learning outcomes and the ways to improve them, so we rely on large national surveys to consider on-the-ground campus practices. Finally, we conclude with key issues, challenges, and remaining questions.

2.2 Overview of American Higher Education

From an international perspective, US tertiary education is distinctive in several respects. Among its most prominent defining characteristics are its size, its differentiation, its marketized control and coordination, and the holistic missions and activities of its individual institutions.

Size

US higher education is approaching what sociologist Martin Trow (2007) termed “universal access.” Over 60% of all secondary school graduates enter postsecondary education directly after graduation (National Center for Higher Education Management Systems 2016), but substantial and growing numbers of students enroll many years after secondary school graduation. In fall 2014, 4665 institutions enrolled over 20 million students, representing approximately 1 out of every 12 US adults (*Chronicle of Higher Education Almanac* 2016; U.S. Census Bureau 2016). Some national leaders, however, express consternation over the USA losing its global lead in tertiary attendance and graduation rates (<https://www.whitehouse.gov/issues/education/higher-education>), and there are signs that the public's declining faith in the quality and value of college attendance may threaten further enrollment growth (Public Agenda 2016).