Learning Engineering for Online Education is a comprehensive overview of the emerging field of learning engineering, a form of educational optimization driven by analytics, design-based research, and fast-paced, large-scale experimentation. Chapters written by instructional design and distance learning innovators explore the theoretical context of learning engineering and provide design-based examples from top educational institutions. Concluding with an agenda for future research, this volume is essential for those interested in using data and high-quality outcome evidence to improve student engagement, instructional efficacy, and results in online and blended settings.

Chris Dede is Timothy E. Wirth Professor in Learning Technologies at Harvard Graduate School of Education, USA, where he served as Chair of the Learning and Teaching Department.

John Richards is Lecturer at Harvard Graduate School of Education, USA, and President of Consulting Services for Education (CS4Ed). He has served as President of the JASON Foundation for Education and Senior Vice President and General Manager of Turner Learning, Inc.

Bror Saxberg is Vice President of Learning Science at the Chan Zuckerberg Initiative, USA. He was previously the Chief Learning Officer at Kaplan, Inc.
This page intentionally left blank
LEARNING ENGINEERING FOR ONLINE EDUCATION

Theoretical Contexts and Design-Based Examples

Edited by Chris Dede, John Richards, and Bror Saxberg
CONTENTS

List of figures vii
List of tables xi
List of contributors xii
Foreword by Huntington Lambert xvii

1 Introduction: Improving Efficiency and Effectiveness Through Learning Engineering 1
Chris Dede

PART I
Dimensions of Learning Engineering 15

2 The Role of the Learning Engineer 17
Phillip D. Long

3 Tinkering Toward a Learning Utopia: Implementing Learning Engineering 36
Barbara Means

4 Learning Engineering Teams: Capabilities and Process 53
Craig D. Roberts and Shawn J. Miller
PART II

Cases of Practice

5 From Artisanship to Learning Engineering: Harvard Division of Continuing Education’s Framework for Improving Teaching and Learning

Henry Leitner, Rebecca Nesson, and Edward Walker

6 Personalization to Engage Differentiated Learners at Scale

Chris Jennings

7 Jill Watson: A Virtual Teaching Assistant for Online Education

Ashok K. Goel and Lalith Polepeddi

8 Creating Personalized Learning Using Aggregated Data from Students’ Online Conversational Interactions in Customized Groups

Scott M. Martin and Matthew L. Trang

9 Executing the Change to Learning Engineering at Scale

Bror Saxberg

10 Rethinking Learning in the 21st Century

Sanjay Sarma

PART III

Conclusions

11 Learning Engineering: Developing a Research Agenda to Bridge the Present to an Aspirational Future

John Richards, Chris Dede, and Bror Saxberg

Index
### FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Five Phases of Collaborative Data-intensive Improvement</td>
</tr>
<tr>
<td>5.1</td>
<td>High-level-goals view of the course with associated learning objectives</td>
</tr>
<tr>
<td>5.2</td>
<td>Mapping of the final project assignment of the course across the topics of</td>
</tr>
<tr>
<td></td>
<td>the course for which it is an aligned assessment</td>
</tr>
<tr>
<td>5.3</td>
<td>Properties of the course activities mapped on to the learning objectives of</td>
</tr>
<tr>
<td></td>
<td>the course (a)</td>
</tr>
<tr>
<td>5.4</td>
<td>Properties of the course activities mapped on to the learning objectives of</td>
</tr>
<tr>
<td></td>
<td>the course (b)</td>
</tr>
<tr>
<td>5.5</td>
<td>Properties of the course activities mapped on to the learning objectives of</td>
</tr>
<tr>
<td></td>
<td>the course (c)</td>
</tr>
<tr>
<td>5.6</td>
<td>Properties of the course activities mapped on to the learning objectives of</td>
</tr>
<tr>
<td></td>
<td>the course (d)</td>
</tr>
<tr>
<td>5.7</td>
<td>Infrastructure systems automatically generate data from student participation</td>
</tr>
<tr>
<td>5.8</td>
<td>Examples of simple user-activity dashboards</td>
</tr>
<tr>
<td>5.9</td>
<td>Views of 5-minute segments of a two-hour lecture by individual students</td>
</tr>
<tr>
<td>5.10</td>
<td>Word Cloud for one lecture from a course titled “Organizational Behavior”</td>
</tr>
<tr>
<td>5.11</td>
<td>Example of student data</td>
</tr>
<tr>
<td>5.12</td>
<td>Partnership course design</td>
</tr>
<tr>
<td>5.13</td>
<td>Examples of engineered interventions</td>
</tr>
<tr>
<td>5.14</td>
<td>A frame for research in learning engineering</td>
</tr>
</tbody>
</table>
6.1 Analytics Academy registrants segmented by business type  92
6.2 Analytics Academy registrants segmented by job role  93
6.3 Analytics Academy registrants segmented by gender  94
6.4 Analytics Academy registrants segmented by age  94
6.5 Analytics Academy registrants segmented by country  95
6.6 Analytics Academy registrants segmented by language  95
6.7 Google Analytics for Beginners course registrants who reported their subject-matter experience level as “beginner”  97
6.8 Advanced Google Analytics course registrants who reported their subject-matter experience level as “beginner”  98
6.9 GAB Unit 1 path (intended completers who selected ecommerce business type)  104
6.10 GAB Unit 1 path (intended completers who selected online support business type)  105
6.11 GAB Unit 1 path (intended completers who selected online community business type)  106
6.12 GAB Unit 1 path (intended completers who selected agency job role)  107
6.13 GAB Unit 1 path (intended completers who selected developer job role)  108
6.14 Student criteria responses mapped to suggested lessons and courses  113
7.1 While the video lessons in the OMSCS KBAI course are like a textbook, the class forum is like a virtual classroom where students ask questions, discuss ideas, and give feedback  123
7.2 In this question about a class project with a coding component, the student asks whether there is a limit to their program’s run time  128
7.3 In this question about a class assignment involving a writing component, the student asks whether there is a maximum word limit  129
7.4 In this introduction, the student expresses interest in learning more about artificial intelligence  131
7.5 In this question about a class assignment, the student asks whether there is a preferred way to name their submission  132
7.6 In this introduction, the student shares that they live in Atlanta  133
| 7.7 | In this introduction, the student shares that they are currently taking another OMS course called Computer Architecture in addition to KBAI. |
| 7.8 | In this question about a class assignment involving a written component, the student asks whether there is a preferred format for citations. |
| 7.9 | In this question about a class project involving a coding component, the student asks for more feedback after submitting their assignment to the automated grading system. |
| 7.10 | In this question about a class assignment involving a writing component, the student asks about whether there is a preferred format to name files. |
| 7.11 | In this question about the class midterm involving a written component, the student asks about the level of detail they should include in their responses. |
| 7.12 | Students react to our class post at the end of KBAI Spring 2016 class announcing the true identity of Jill Watson. |

| 8.1 | Archived Udemy Course UI with live learner interactive chat. |
| 8.2 | Screen shot of LiveEdu student interface with instructor live-stream. |
| 8.3 | Scriyb Instructor User Interface, with Group 2 chat viewable on the right-side column. |
| 8.4 | Example of grouping students (clustering) using r, d, z variables. |
| 8.5 | Sub-regrouping and dynamic regrouping of students. |
| 8.6 | Simplified DALI DNLN Deep-Learning Model. |
| 8.7 | Demonstration of an initial DALI recommendation and follow-up “helpful” solicitation with resultant student response. |
| 8.8 | Complete Scriyb Knowledge Acquisition System Model. |
| 9.1 | Working memory—long-term memory model. |
| 9.2 | Wheel of evidence-based learning. |
| 9.3 | Worked example vs. Kaplan video from KTP for LSAT logical reasoning problems. |
| 9.4 | Backward design. |
| 9.5 | Stages of mastery. |
| 9.6 | Knowledge Component Event Matrix—based on KLI framework from CMU. |
| 9.7 | Approach to validity and reliability, grounded in NCME. |
9.8 Additional project-based training to support evidence-based instructional design 179
9.9 Systematic approach to twice-yearly conversations about learning priorities, challenges, trade-offs 181
9.10 “Kaplan Way for Learning Checklist” components 182
9.11 Key Kaplan products evaluated with checklist 183
9.12 High volume technology-enhanced courses invite systematic controlled trials 183
9.13 Variety of areas to explore through RCTs at KU 185
9.14 Example of process used to track progress of pilot 186
9.15 Innovation portfolio at Kaplan University 187
10.1 Existing educational pathways. Is this the limit to our design language? 199
10.2 What happens when we add a third instrument to our orchestra? 203
11.1 The inner Assessment-Content Management loop provides personalized practice. The outer curriculum loop provides direct contact between teacher and students. 216
11.2 How does the student provide feedback to the teacher? 218
11.3 Classroom side of the DTP architecture 219
11.4 Multiple courses can provide data for the Analytics Engine and can foster a community of support for the instructor 220
6.1 General engagement metrics for beginner and advanced course registrants 96
6.2 Google Analytics for Beginners 99
6.3 Advanced Google Analytics 99
6.4 Google Analytics for Beginners (by business type sorted by completion rate) 100
6.5 Google Analytics for Beginners (by job role sorted by completion rate) 100
6.6 Advanced Google Analytics (by business type sorted by completion rate) 101
6.7 Advanced Google Analytics (by job role sorted by completion rate) 102
6.8 Google Analytics for Beginners (by developer job role sorted by completion rate) 102
6.9 Advanced Google Analytics (by developer job role sorted by completion rate) 103
7.1 The level of participation of online students in the OMSCS KBAI class on the digital forum is much higher than that of residential students 125
9.1 Changes to learning roles when learning engineering is more fully implemented 188
CONTRIBUTORS

Chris Dede is the Timothy E. Wirth Professor in Learning Technologies at Harvard's Graduate School of Education (HGSE). His fields of scholarship include emerging technologies, policy, and leadership. In 2013, he co-convened a NSF workshop on new technology-based models of postsecondary learning, and, in 2015, he led two NSF workshops on data-intensive research in the sciences, engineering, and education.

Ashok K. Goel is a Professor of Computer Science in the School of Interactive Computing at Georgia Institute of Technology in Atlanta, GA. Goel serves on Georgia Tech’s Commission on Next in Education and co-leads its task forces on Future of Pedagogy and Future Learning Systems. He is the Editor-in-Chief of AAAI’s AI Magazine.

Chris Jennings is a Senior Instructional Designer at Google in San Francisco, CA. He leads the Analytics Academy which offers free online instruction for Google measurement and advertising tools to over a million learners worldwide. Previously, Jennings directed the Center for Innovation in Teaching and Learning at New York University's Stern Business School and built online accredited degree programs for The University of Texas System.

Huntington D. Lambert (Hunt) is Dean of the Division of Continuing Education and University Extension at Harvard University. The Division serves 30,000 students annually. Previously, Lambert served as Associate Provost of OnlinePlus at Colorado State University (CSU). Prior to this, Lambert was a founder and interim CEO of CSU Global Campus and Associate Vice President of Economic Development at CSU. He is also the
former Director of the Colorado State University Entrepreneurship Center and a former member of the faculty at the CSU College of Business. During his business career, he was part of 25 startups in 12 countries, 22 of which are still operating. He teaches strategy, entrepreneurship, and business plan development.

As Associate Dean and Chief Innovation Officer for Harvard Division of Continuing Education (HDCE), **Henry Leitner** provides vision and leadership in the development and implementation of technology-based initiatives that serve both the administrative and academic areas, including online teaching and learning. He oversees the Master’s of Liberal Arts degrees in IT at Harvard Extension School—graduate programs of study in which adult learners are exposed to modern software engineering methodologies; theoretical and formal areas of computer science; advanced topics in cryptography, nanoscale science, digital media arts and sciences, and enterprise computing; as well as the latest trends in IT management. Since 1986, he has been a Senior Lecturer on Computer Science for the School of Engineering and Applied Sciences at Harvard, where he has taught a number of large courses.

**Phillip D. Long** was the CiNO for Project 2021, Associate Vice Provost, and clinical professor in Educational Psychology and the School of Information at the University of Texas, Austin. Dr. Long provided leadership to the university's strategy for technology enhanced learning. Phil contributed leadership to UTx, the local implementation of edX MOOCs, the UT learning analytics initiative and design of learning environments. Phil retired from UT Austin in May of 2018. He is now Senior Scholar at the Center for New Designs in Learning and Scholarship at Georgetown University and founder of RHHz Consulting, LLC. His current interests focus on technologies of trust, including the use of blockchains in credentialing, the design of creative environments, remote labs, emerging technologies and the role of big data, AI, and analytics in capturing and improving decision-making. He also serves as board member for Nexford University.

**Scott M. Martin** is an inventor, mentor, educator, and entrepreneur. He is founding Director of the Virginia Serious Game Institute (VSGI), previous Director of the Computer Game Design programs, and current professor at George Mason University in Fairfax, VA. Martin’s research focus has been on designing new machine-learning algorithms applied to teaching and learning, combining the fields of data science and the social sciences of human knowledge acquisition. He has founded multiple institutes, academics programs, research centers, schools, and colleges, and is the co-Founder of Scriyb LLC, a machine-learning education technology company.
Barbara Means studies the effectiveness of innovative education approaches supported by digital technology. Formerly the Director of the Center for Technology in Learning at SRI, she is now Executive Director of Learning Sciences Research at Digital Promise. Her recent work includes evaluating the implementation and impacts of newly developed, adaptive-learning courseware developed with funding from the Bill & Melinda Gates Foundation. A fellow of the American Educational Research Association, Means has served on many study committees of the National Academies of Science, Engineering, and Medicine, including the one currently producing a companion volume to the classic reference How People Learn. She has advised the US Department of Education on national education technology plans and authored or edited more than a half dozen books related to learning and technology. Her most recent book, Learning Analytics Goes to School, was co-authored with her Digital Promise colleague Dr Andrew Krumm. Means earned her undergraduate degree in psychology from Stanford University and her PhD in educational psychology from the University of California, Berkeley.

Shawn J. Miller leads Duke Learning Innovation in its mission to help Duke students learn more and to enable more people to learn from Duke. Learning Innovation promotes innovation in teaching, manages the direction and strategy for Duke’s learning technologies, and leads Duke’s online initiatives and production (including Duke’s partnership with Coursera). His interest in the concept of learning engineering stems from over 20 years of experience assembling, leading, and managing a wide variety of teams involved in exploring and developing digital learning innovations.

Rebecca Nesson is the Executive Director of Digital Teaching and Learning at Harvard Extension School. Together, she and her team of learning designers, instructional technologists, media producers, and faculty development specialists support the faculty in delivering over 600 online, credit-bearing courses each year to adult, part-time learners with the curiosity, ability, and drive to succeed at Harvard. At Harvard Division of Continuing Education and in her prior work, Nesson focuses on creating courses and technology that foster community and support meaningful interaction among students and teaching staff. Nesson also teaches in the Computer Science department at Harvard College and for the Harvard Summer School.

Lalith Polepeddi is a machine learning scientist at Georgia Institute of Technology in Atlanta, GA. His interests involve applying artificial intelligence to domains including education, energy, and climate change.

John Richards is Founder and President of Consulting Services for Education, Inc. (CS4Ed), providing strategy, marketing, and business development services to the education industry. He is a Lecturer on Education at the Harvard Graduate School of Education teaching courses in entrepreneurship in the education marketplace.
Craig D. Roberts is a Senior Program Officer at the Bill & Melinda Gates Foundation. There, he leads the evaluation of investments in postsecondary digital learning, advising, and developmental education. Prior to joining Gates, Roberts served as an executive at an education technology company and a university professor. As Chief Learning Officer of TechSmartKids, he enabled public school districts to teach computer science and coding in 2nd–12th grades. During his tenure at Duke University, he served as a neuroscience and education faculty member in the Duke Institute for Brain Sciences, as well as Director of Learning Innovation and Global Programs. Roberts’s background is in learning science with a focus on iterative improvement and the scaling of digital interventions. He holds a Bachelor of Science from Wofford College and a PhD in Neuroscience from the University of Miami.

Sanjay Sarma is the Fred Fort Flowers (1941) and Daniel Fort Flowers (1941) Professor of Mechanical Engineering at MIT. He is the Vice President for Open Learning at MIT. He co-founded the Auto-ID Center at MIT and developed many of the key technologies behind the EPC suite of RFID standards now used worldwide. He was also the founder and CTO of OATSystems, which was acquired by Checkpoint Systems (NYSE: CKP) in 2008. He serves on the boards of GS1US and Hochschild Mining and several start-up companies including Top Flight Technologies. Dr Sarma received his Bachelors from the Indian Institute of Technology, his Masters from Carnegie Mellon University, and his PhD from the University of California at Berkeley. Sarma also worked at Schlumberger Oilfield Services in Aberdeen, UK, and at the Lawrence Berkeley Laboratories in Berkeley, California. He has authored over 100 academic papers in computational geometry, sensing, RFID, automation, and CAD, and is the recipient of numerous awards for teaching and research including the MacVicar Fellowship, the Business Week eBiz Award, and Informationweek’s Innovators and Influencers Award. He advises several national governments and global companies.

Bror Saxberg is the Vice President of Learning Science at the Chan Zuckerberg Initiative. Previously, he served as the Chief Learning Officer for Kaplan, Inc., and, in addition to earlier employment at K12, Inc. and McKinsey and Co., he holds an MD from Harvard Medical School, a PhD from MIT, an MA from Oxford University, and bachelor degrees from the University of Washington.

Matthew L. Trang is currently a freshman engineering student and Pamplin Leadership Award recipient at Virginia Polytechnic Institute and State University (Virginia Tech). In 2018, he graduated as a National Merit Finalist and Valedictorian from Patriot High School and the Governor’s School at Innovation Park in Manassas, Virginia. During his high school years, he conducted research at the George Mason University Center for Applied Proteomics and Molecular Medicine and the Virginia Serious Game Institute. His contributions have led to
the development of multiple patents including a wearable knee rehabilitation device called the SmartSleeve and artificial cognitive learning maps. At the VSGI, his research responsibilities included developing algorithms, designing system diagrams, and exploring applications of neurological systems to improve online learning. In the future, he plans to continue researching and developing innovative technologies to advance the health care field.

Edward Walker advises senior managers of institutions, agencies, and foundations on the synthesis of innovative software technology and evolving business processes. His career includes executive experience as CEO of the IMS Global Learning Consortium and Vice President of BBN Systems and Technologies. He was a Principal Research Scientist at MIT, Member of Staff at Bell Laboratories, and staff member at the CMU Software Engineering Institute. He has taught courses at Harvard, MIT, Temple University, and the University of St. Andrews. He has a PhD in psycholinguistics and is a graduate of the Greater Boston Executive Program in Business Management of the Sloan School. He served as Co-Chair of the Advisory Board of the consortium for Multimedia Resources for Learning and Online Teaching (MERLOT) and as a member of the Advisory Boards of the Curriki Foundation and the Open Knowledge Initiative. He is Co-Chair of the Higher Education Board of the Software and Information Industry Association.
On June 9–11, 2017, an invitational research workshop titled “The Future of Technology-Supported Education: Improving Efficiency and Effectiveness Through ‘Learning Engineering’” was held at Harvard University. The Harvard Division of Continuing Education (DCE), with the Harvard Graduate School of Education (HGSE) and Bror Saxberg of Kaplan, Inc. (now of the Chan Zuckerberg Initiative), co-hosted this event. Top scholars across the United States shared their perspectives, strategies, findings, and insights about the current state and likely evolution of learning engineering. This book includes chapters extensively discussed at the workshop. It is meant to inform scholars, practitioners, policymakers, and funders about emerging practices and challenges in evidence-based learning engineering for improved learning outcomes.

This workshop is part of a series of applied R&D activities aimed at achieving global excellence by improving education outcomes and cost effectiveness within DCE, across Harvard, and throughout the field. The institutional partners hosting this workshop have substantial expertise and experience in on-campus, online, and blended learning. Harvard’s DCE, through its Extension School (one of Harvard’s 12 degree-granting entities), offers nearly 1,200 courses (600 of which are available online) taught by Harvard scholars, industry experts, and leading researchers. The Canvas Learning Management System and other digital tools are utilized by students and instructors engaged in online, on-campus, and blended courses that can lead to undergraduate and graduate degrees as well as numerous micro-credentials.
I sponsored this activity because of the contrast I saw between academic assessment approaches and the modern systems of Kaizan, or continuous improvement. I came late to higher education having spent 20 years in large companies and start-ups. I worked in and built environments where we toiled to find the right metrics to inform how we could improve our offerings and be sure we were earning loyal customers. I enjoyed and practiced Frederick Reichheld’s *Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value* (1996). I also ran a large research organization using older techniques—as well as the newest—to identify opportunities based on macro trends, statistics, deep thought, and action with metrics. Data and analysis drove me, and we were very successful. Imagine my surprise when I arrived in higher education and tried to improve our learning outcomes based on the university-approved assessment system. It did not take me long to realize that the assessment system was designed and administered to support accreditation and reporting requirements with little thought to its value to student learning outcomes or faculty teaching improvement.

One of my undergraduate areas of study was the neurophysiological and biochemical structure of memory. From this and continued readings in the neurological and brain science areas, I came to appreciate how much we did and did not know about the science of learning. Recently science has caught up with my hopes, and we know a great deal about learning from both the biological and experiential views. When I plunged into college and graduate school classrooms as a teacher at age 42, I also started to learn about teaching and the myriad of pedagogic, class management, and media tools that might make one effective in the classroom. The options were few, but they enabled a whole new niche of learners to participate: the part-time adult learner who could not control time and place. I quickly discovered the dearth of tools available for me to learn how to be a great teacher as well as the huge amount of time peers needed to invest in me to raise my effectiveness.

Years later, after winning professor-of-the-year awards, borrowing many active learning techniques and applying them in my innovation and entrepreneurship classes, helping lead a team that founded a new public online university (Colorado State University [CSU] Global Campus), growing a research-faculty-based online division of continuing education to success (CSU Online), and then being hired by Harvard University as Dean of Continuing Education and University Extension, I was lucky enough to meet Bror Saxberg. We met at an edX Global Forum event at MIT, and he used the words learning engineering to describe his work as Chief Learning Officer of Kaplan Inc.

I was inspired by Bror and Frederick Hess’s book *Breakthrough Leadership in the Digital Age: Using Learning Science to Reboot Schooling* (2013) and started building a learning engineering group at the Harvard Extension School. In designing this group, I returned to Nobel Laureate Herbert A. Simon’s 1967 article “The Job of the College President.” Simon defined the original learning engineering field all those years ago when he wrote:
There is no simple path that will take us immediately from the contemporary amateurism of the college to the professional design of learning environments and learning experiences. There are, however, some obvious first steps along the way.

The most important step is to find a place on the campus for a team of individuals who are professionals in the design of learning environments—learning engineers, if you will. Recalling our earlier budget equations, how will we pay for the new learning engineer? The increase of average class size is one possibility, but I won’t insist on that route if another, more appealing, suggests itself.

The learning engineers would have several responsibilities. The most important is that, working in collaboration with members of the faculty whose interest they can excite, they design and redesign learning experiences in particular disciplines. Like all staff experts—operations analysts in business, for example—their long-run effectiveness will hinge on their ability to transmit their expertness to the line organization, in this instance the teaching faculty. Because they are experts in designing learning experiences, an important part of their skill will be directed toward devising learning experiences for the faculty. In particular, concrete demonstrations of increased learning effectiveness, on however small a scale initially, will be the most powerful means of persuading a faculty that a professional approach to their students’ learning can be an exciting and challenging part of their lives (Simon 1967, p. 77)

I enjoy thinking about how Simon might have expanded his thinking if personal computers and the Internet and all the modern tools they enabled were added to his descriptions of the college campus learning environment of 1967. The opportunity for data collection, analysis, and teaching and learning improvements has grown exponentially because of technology. So why have so few of us leveraged those tools? Why do we not know the few best metrics that can help faculty be more effective? Why do so few colleges teach their graduate students how to teach or even their new faculty how to teach? And how long can we stay in the dark under the enormity of the pressure caused by the defunding of public higher education by the states, as well as the falling returns of endowments and research funding upon which private research universities depend?

My answer is we cannot wait much longer. The Harvard Extension School (HES) leads the university in teaching and learning innovations because its low tuition, open-access courses, and global-inclusion mission forces it to experiment with low-cost and extended-education models. To best serve Extension School learners, I am determined to bring forward the light of learning engineering. I believe that if Harvard can lead in being able to deliver and document success at low cost then our large public peers and others will be able to grow their
inclusive education units to the point where the United States again leads in education. The Harvard Extension School has done this brilliantly for over 100 years, as is so well illustrated in former Dean Michael Shinagel’s 2009 book *The Gates Unbarred*. While night classes expanded to include electric lights in the early 1900s, radio in the 1940s, TV in the 1950s, and even submarines in the 1960s, it was not until the Internet, learning management systems (LMSs), customer relationship management systems (CRMs), and learning analytic platforms that the data-driven, continuous improvement for tens of thousands of students each year emerged. Today, all of the 1,200 HES classes are open access, and students earn their way into admissions with honors performances in specified classes. This strange admissions approach leads to an 85% graduation rate, a huge win for students and the goal of low student debt. With no endowment support and a cultural legacy of serving the local community with courses priced at “the cost of two bushels of wheat” (as quoted in Shinagel, 2009, p. 5), HES has a mandate to serve both Harvard’s mission and the community’s access demands. Attending part-time and paying low tuitions, most HES students graduate with no debt. Because of this, HES has become a primary innovator for Harvard including offering Harvard’s first Internet-based, online class in Spring 1998. This year, HES will serve over 16,000 part-time students from 130 countries under its mission to “extend Harvard to part-time learners with the academic ability, curiosity and drive to succeed at Harvard rigor courses and programs.”

Today, our great universities have the tools to serve millions more who need education to be socially, civically, and economically relevant in the face of great global change. And yet we still have not mastered assessment, analytics, and learning engineering so we can accomplish this growth without risking quality or demanding unsustainable physical campus expansions. This book is an attempt by a few selected authors to bring together a subset of knowledge and start the process of defining the new jobs Simon called learning engineers.

Harvard’s 108 years of serving adult, part-time learners and its 20 years doing so online have taught me the following core lessons:

- Each of us must know what niche of the learner market we are serving, so we can engineer learning opportunities around their abilities and constraints. If we do not, in 20 years the United States and other Western countries will not have the workforce, civic, or social environment they value. There are many segments of learners, and the adult, part-time learner segment now represents 85% of all learners (Snyder, 2016). Segments in higher education are not only defined by demographics, academic ability, and program area, but they are also powerfully local. None of us can serve them all, and even Harvard does and should lose great students to trusted local schools.
- Each of us must choose what teaching and course delivery methods and blended varieties we can best deliver. Virtually all of the known teaching and delivery methods will continue and more will emerge, but only some will work for any particular learner niche or individual learner. To succeed, every
school needs to make a much more precise set of decisions about the **who** it is teaching and the **what** it is teaching, and then pick the **how**. The how may well vary for each faculty member, their learning objectives, their content, and their learners. At Harvard Extension, we no longer “go online”; we “extend our faculty” with technology. Thus, each of our online courses is unique, and we have developed the technology and techniques for small classes while generating budget surpluses. That fits for Harvard faculty and for our target part-time, adult learner who can do Harvard-level, rigorous courses. It also turns out to be the lower cost option for our class sizes as compared to building massive open online courses (MOOCs), fully online, active-mastery courses, or competency-based modules. It likely is not right for your institution; please determine and commit yourselves to what is.

- Each of us must state our value proposition for each segment of learners. Assuming almost every college and university goes online and hybrid over the next 20 years, then what makes your institution’s offerings special? If those offerings are not superior to the alternatives, determine where your institution can be the best. A key to the long-term success of Harvard’s Extension School is knowing whom we serve and whom we do not serve. By focusing, we can keep costs down and achieve a low-tuition, high-quality solution with no endowment support.

Learning engineering is about helping ourselves and our peers collect the right data, do the right analysis, and make decisions around the right metrics for each program, market segment, faculty style, learning objective, content type, and delivery method. Thank you for joining this journey. If the Lumina Foundation is even partly right (Lumina’s Goal, 2017), we are already way behind supporting the doubling of educated people the world needs in our complex civic, diverse social, and rapid job-change-driven economic communities.

**References**


This page intentionally left blank
INTRODUCTION

Improving Efficiency and Effectiveness Through Learning Engineering

Chris Dede

In the last two decades, digital media have greatly expanded the scope and impact of distance education, as well as widened the range of models used for teaching and learning. The rise of online learning as a major form of education for adults, coupled with the advent of MOOCs (massively open online courses), highlights the importance of enhancing digital education’s outcomes through improved instructional design. As a powerful method for accomplishing this goal, learning engineering applies a principled set of evidence-based strategies to the continual re-design of educational experiences to optimize their effectiveness and efficiency. This book centers on the use of learning engineering to improve online courses in higher education.

Both online teaching and blended instruction (online plus face-to-face) are good venues for learning engineering because the digital media used automatically generate rich, time-stamped log files documenting each student’s interactions with curricular materials, peers, and instructor. The evidence that guides constant, rapid cycles of improvement in learning engineering comes from this form of big data coupled with high quality outcome evidence (both near- and long-term). Doug Laney, an analyst with the META Group (now part of Gartner), described big data with a collection of “v” words (Laney, 2001), referring to (a) the increasing size of data (volume), (b) the increasing rate at which it is produced and analyzed (velocity), and (c) its increasing range of sources, formats, and representations (variety). To this, other authors have added veracity, to encompass the widely differing qualities of data sources, with significant differences in the coverage, accuracy, and timeliness of data (Dong & Srivasta, 2013).

Technological and methodological advances have enabled an unprecedented capability for decision making based on big data, and its use has become well established in business, entertainment, science, technology, and engineering (Dede, 2015). For example, online purchases are now guided by recommendation engines...
that analyze an individual’s shopping patterns and suggest products bought by others who have similar patterns of purchases. Big data is beginning to be utilized for decision making in higher education as well; one example is early identification of at-risk students based on analysis of their behavioral patterns. Thus far, these analytics focus on students’ macro-behaviors (e.g., adding or dropping courses) rather than their micro-behaviors (second-by-second activities in learning experiences coupled with evidence about learning outcomes).

Practical applications to analyzing students’ learning behaviors and outcomes in college and university instruction remain rare because of challenges unique to higher education (Dede, Ho, & Mitros, 2016). First, the sector lacks much of the computational infrastructure, tools, and human capacity required for effective collection, cleaning, analysis, and distribution of large datasets. Second, in collecting and analyzing student data, colleges and universities face privacy, safety, and security challenges not found in many scientific disciplines. Third, higher education should also be concerned with long-term goals—such as employability, critical thinking, and a healthy civic life—even though at many institutions these objectives are not apparent in their tactical decision making. Since it is difficult to measure these outcomes, particularly in short-term studies, researchers studying effectiveness in higher education often rely on theoretical and substantive arguments to justify imperfect, immediate proxies for these longitudinal objectives. This is made more difficult by a lack of training for or awareness among most faculty of the challenge of gathering valid and reliable evidence about learning outcomes, either near- or long-term.

Learning engineering cannot resolve these difficulties in measuring long- or short-term goals, but it can use big data to iteratively improve the design of learning experiences. Learning analytics and educational data mining are concerned with exploring the unique types of data that come from educational settings. Learning engineering combines methods from these fields with design-based research to better understand how students learn, what instructional strategies enable optimal learning (Baker & Yacef, 2009), and how to gather valid, reliable evidence about learners’ mastery of intended outcomes.

Learning engineering can improve online learning outcomes in higher education in a variety of ways. Through educational optimization, students’ engagement with courses can deepen, teachers can improve the efficiency of their instruction, and a broader range of learners can succeed because courses are tailored to their individual needs. The chapters in this book describe these and other types of improvements from learning engineering. Personalized learning is a conceptual framework that articulates the mechanisms that create many of these improvements.

**Personalized Learning**

The US Department of Education’s 2010 *National Education Technology Plan* provided an early, influential definition of personalization, stating:
Personalization refers to instruction that is paced to learning needs, tailored to learning preferences, and tailored to the specific interests of different learners. In an environment that is fully personalized, the learning objectives and content as well as the method and pace may all vary.

(US Department of Education, 2010, p. 12)

In the same time frame, the 2010 SIIA Symposium on [Re]Design for Personalized Learning articulated five essential elements of personalized learning: flexible, anytime/everywhere learning; redefine teacher role and expand “teacher”; project-based, authentic learning; student driven learning path; and mastery/competency-based progression/pace.

Since then, personalized learning has been clouded by many definitions, most of which weaken the concept to the point that everyone can claim they are already “personalizing” learning. (This bastardization of a rigorous educational innovation is unfortunately quite common.) For the purposes of contextualizing learning engineering, we define personalized learning as having four fundamental attributes:

1. Developing multimodal experiences and a differentiated curriculum based on universal design for learning principles;
2. Enabling each student’s agency in orchestrating the emphasis and process of his or her learning, in concert with the evidence about how learning works best and with mentoring about working toward long-term goals;
3. Providing community and collaboration to aid students in learning by fostering engagement, a growth mindset, self-efficacy, and academic tenacity; and
4. Guiding each student’s path through the curriculum based on diagnostic assessments embedded in each educational experience that are formative for further learning and instruction.

Substantial evidence exists that combining these four attributes leads to learning experiences that provide strong motivation and good educational outcomes for a broad spectrum of students (Dede & Richards, 2012).

As targets of opportunity for learning engineering in higher education, the strategies below for instructional interventions based on big data potentially provide several ways to improve learning through personalization (Dede, Ho, & Mitros, 2016):

- Individualizing a student’s path to content mastery, through adaptive, competency-based education. One example is using game-based environments for learning and assessment, where learning is situated in complex information and decision-making situations that adapt to each learner’s progress.
- Improving learning as a result of faster and more in-depth diagnosis of learning needs or course trouble spots, including assessment of skills such as systems thinking, collaboration, and problem solving in the context of deep, authentic subject-area knowledge assessments.
Increasing the efficiency of learning to reduce overall costs to students and institutions.

The chapters in this book provide examples of applying learning engineering to these and other targets of opportunity.

The Value of Personalization in Online Learning

Throughout its history, distance education has struggled with efficiency and effectiveness due to a lack of personalization. A brief history of distance education illustrates this problem (Dede, Brown–L’Bahy, Ketelhut, & Whitehouse, 2004). In the 19th century, distance education in the United States was shaped by new technologies that allowed educators to overcome barriers of distance and time—shifting understandings of the purpose of education—and social, political, and geographic forces. The development and implementation of the first correspondence courses were credited to Sir Isaac Pitman of England, the inventor of shorthand. In 1840, he used the postal service in England to reach learners at a distance. A more formal version of the early American correspondence course was created by Anna Ticknor of Boston in 1873. In order to increase educational opportunities for women, she originated the Society to Encourage Studies at Home. The society provided courses of study for women of all social classes and served over 10,000 women over its 24-year lifespan (Nasseh, 1997; Stevens-Long & Crowell, 2002).

In 1878, John H. Vincent, co-founder of the Chautauqua Movement, created the Chautauqua Literary and Scientific Circle. This Circle offered a four-year correspondence course of readings; students who successfully completed the course were awarded a diploma. This course was open to all adults, including women and senior citizens (Scott, 1999). By 1892, the 19th century version of the “Information Superhighway” (otherwise known as rural free delivery) paved the way for Penn State University to provide higher education to rural families (Banas & Emory, 1998). Other institutions of higher education, notably the University of Chicago and the University of Wisconsin, modeled their extension schools after the Chautauqua program (Scott, 1999).

At the start of the 20th century, distance education still relied on correspondence courses delivered primarily through the postal service. Many of these courses were delivered to their students by mail, as discussed earlier, but did not allow much interaction or individualization (Moran, 1993). Although rules for home study were established in 1926 to allow some form of governmental control, correspondence methods were not conducive to supporting learners nor were they standardized (PBS, 2002). One of the main goals of early distance education was to help inculcate immigrants into the “American way of life” (Sumner, 2000), but these learners needed substantial guidance and aid. As a result, poor curricular design and lack of support were particularly problematic, and the dropout rate was high (Shea & Boser, 2001).
While distance education is rooted in the 19th century, the field blossomed in the 20th century. Distance educators looked to technological innovations to provide new opportunities for their field, and the 20th century was rife with technological advances (Mood, 1995). During the 20th century, distance education embraced radio, television, computers, and ultimately the Internet. As the methods of delivery for distance education expanded, so did the diversity of learners seeking distance education and their reasons for enrolling in such courses. Individuals interested in learning cultural norms, becoming more capable in the workforce, or hoping to re-situate themselves in their social context after wartime service, became major consumers of distance education (Dymock, 1995).

In the 1920s, distance education started to utilize radio for delivery of lessons (Bourke Distance Education Centre, 2002; Nasseh, 1997). In a push to widen access, speed the interaction between student and professor, and personalize the delivery of distance education, the use of radio was seen as an exciting opportunity. In the mid-1930s, an American art history course was offered by radio broadcasts (Funk, 1998), and other courses supported forming listening groups to enhance learning (Mood, 1995).

However, despite the rapid rise of radio technology, distance education courses were rarely if ever offered for credit in higher education (Nasseh, 1997). The education community, along with society as a whole, regarded legitimate education as only possible in conventional locales, such as classrooms (Funk, 1998). To address concerns about a lack of teacher interaction in distance education, a modification of the correspondence course was designed in Soviet Russia in the 1930s, called the Consultation Model (Tait, 1994). As its name implies, this type of correspondence course included periodic face-to-face meetings with instructors; however, unlike its name, the consultations were mostly lecture-based meetings intended to spread communist dogma.

Television was the next big advance in distance education technology. As early as 1934, the State University of Iowa used television to deliver course content. Early research into learning via television indicated mixed results, with several studies showing that it was similar to conventional instruction. Gayle Childs referred to televized distance education as an “instrument of delivery, not a pedagogical method” (Jeffries, 2002, p. 6).

Prior to the introduction of computer technologies in the 1960s, correspondence-course and independent-study models of distance education posed challenges to the learning and teaching processes. This contributed to a persistent problem of credibility for the field. Tele-courses (Verduin & Clark, 1991), which developed in the 1970s, showed promise for minimizing some of these problems. Previously, television had primarily been used as an electronic blackboard and for the delivery of standardized content through lectures intended to reach wide audiences. The development of videotape allowed educators to customize the same content for different learning environments. This medium also allowed increased flexibility; course content could be stored, delivered, and repeated at
will. This minimized time-dependency, a drawback of previous televised courses. However, despite their advantages, the cost and complexity of producing telecourses made them impractical for teaching large numbers of students.

Around the same time, the open university concept was launched. The creation of universities open to all was driven by the need to provide alternative education for adults whose needs could not be met in the traditional classroom. The British Open University began in 1969 through video broadcasting of its weekly courses on the BBC. Over time and with the advent of new technologies, the British Open University’s model of distance learning evolved into a student-centered delivery system and administrative structure separate from a campus setting. More economically practical than tele-courses, this system envisioned each student as “a node in the network” (Granger, 1990, p. 189) that provided individualized instruction in a virtual classroom. The students had access to a virtual library—customizable based on their particular learning style—and to collaborative tools that encourage discourse and critical thinking (Prewitt, 1998). By encouraging a community of learners, this model overcame some of the problem of isolation.

During the 1970s, the capability of computers to automate tasks and deliver information made them invaluable tools for many companies, thereby increasing the need for technologically competent workers. This prompted the inception of corporate training programs focused on technology literacy. In schools, word processors, spreadsheets, and database applications enhanced the productivity of educators and students. In turn, the development of educational software offered interactive ways to deliver academic content. In the 1980s, due to the rapid evolution of information technology and our increasing dependence upon computers as a society, the use of educational technologies expanded considerably. However, during these first few decades, information technologies were used more to automate traditional models of educational delivery than to develop new forms of pedagogy or to enhance learning across distance. This issue of “old wine in new bottles” is a perennial problem in the history of all educational technologies (Cuban, 2013).

In contrast, during the 1990s, widespread usage of the Internet transformed the nature of distance education. The present-day Internet traces its origins to the ARPANET, a system developed in the late 1960s (Leiner et al., 1997). The ARPANET was initially used by instructors and researchers to share files of information. In 1972, however, email capability was added, transforming computers into a medium that facilitated direct, people-to-people interaction (WSU, 1997). In subsequent years, the advancement of networking technologies led to the eventual development of the Internet. Increasing use of personal computers in schools, businesses, and homes helped to establish this budding network of computers. In particular, with the development of the World Wide Web (WWW) as a means of representing and accessing information, Internet use expanded exponentially.
In 1992, it is estimated that the WWW contained a mere 50 webpages (Maddux, 2001). During the 1990s, decreasing prices for computer technologies led to increases in personal ownership, and self-publishing resulted in an explosion of web-based information. By 2000, the number of webpages rose to at least 1 billion (Maddux, 2001). With its capability to facilitate communication between people in various geographic locations and to disseminate information quickly and relatively inexpensively, the Internet appeared well matched for distance education; Long’s chapter in this book describes details of this evolution. However, change has been very slow in both face-to-face and distance-based educational practice (Cuban, 2013), with a perennial emphasis on presentational instruction and content coverage despite research-based findings emphasizing active learning and skill development (National Research Council, 2005, 2012; Clark & Mayer, 2016).

In online learning, while MOOCs were hailed as a breakthrough, their instructional practices have remained entrenched in outdated models of teaching (Dede, 2013). In particular, methods for personalization such as individualizing a student’s path to content mastery—through adaptive learning or competency-based education—have typically not been utilized (Dede, Ho, & Mitros, 2016). Similarly, more efficient and effective learning as a result of faster and more in-depth diagnosis of student needs or course trouble spots has languished, including assessment of skills such as systems thinking, collaboration, and problem solving in the context of deep, authentic subject-area knowledge assessments.

**Learning Engineering as a Means of Enhancing Educational Efficiency and Effectiveness**

As discussed earlier, learning engineers are professionals who understand theoretical and evidence-based research about learning and learning measurement, apply these findings to test their value in the crucible of specific situations of practice, and refine initial approaches to develop heuristics and models to make students’ learning more efficient and effective. Thus, learning engineering is characterized by evidence-based approaches, measurable and measured outcomes, and iterative processes (Hess & Saxberg, 2013; Saxberg, 2015). This chapter has framed the value of learning engineering for the goal of personalizing online and blended learning. However, learning engineering can take many forms in improving a spectrum of instructional approaches, ranging from data analysis of existing data gathered through students’ interaction with learning management systems to comparisons of targeted instructional variants (A/B testing “to”) to determine the effects of a particular intervention implemented across online courses (Rosen, 2016).

Contributing disciplines include cognitive science, computer science (human-computer interaction, machine learning, artificial intelligence), cognitive psychology, education (psychometrics, educational psychology, learning sciences), and statistics. In 2016, the MIT Online Education Policy Initiative called for
greater integration of insights about learning across these disciplines, using a coordinated research agenda (Willcox, Sarma, & Lippel, 2016). The concluding chapter of this book draws on insights from its chapter authors to offer suggestions about next steps in learning engineering research.

The 2016 MIT report also emphasized the importance of online learning in providing a dynamic digital scaffold for instrumented learning that can aid customization, remote collaboration, just-in-time scenarios, continuous assessment, and blended learning. In particular, among many other evidence-based ideas, the report highlighted spaced learning to improve retention, which allows students and teachers to focus on applying that learning to challenging problems, as well as game-based learning, which can contextualize abstract concepts and provide data on student challenges back to the teacher. Also, some types of online gaming can induce psychological immersion by students, which enables situated learning and transfer to real world applications (Dede, 2014).

Another recommendation of the 2016 MIT report calls for building capacity for educational effectiveness and efficiency through developing many more learning engineers. In addition to the characteristics described above, the importance of learning engineers being able to work with educators is stressed, both in creating new learning experiences and in integrating new technologies and approaches into existing courses. Finally, the 2016 MIT report emphasizes the importance of institutional and organizational changes to take full advantage of the opportunities learning engineers provide. This theme of change in organizational policies and culture is stressed in many of this book’s chapters, particularly Saxberg.

Overview of Chapters

The book begins with this introductory chapter introducing terms and conceptual frameworks, as well as providing a quick summary for the contents of each chapter, grouped into three types of discussions. Dimensions of Learning Engineering are delineated in chapters by Long, Means, and Roberts and Miller. Cases of Practice that illustrate learning engineering knowledge, processes, and pragmatics in action are provided by Leitner, Nesson, and Walker; Jennings; Martin; Goel and Polepeddi; and Saxberg. The book concludes with a chapter by Richards, Saxberg, and Dede summarizing cross-cutting themes from the chapters, describing Digital Teaching Platforms as an infrastructure learning engineers are creating, and delineating important next steps in research and practice for the learning engineering field.

Dimensions of Learning Engineering

Long’s chapter, “The Role of the Learning Engineer,” explores the background for and implications of what learning engineer means in the context of course design, development, and support. The growth of technology development, an
exponential increase in our scientific understanding of the world, and the resulting specialization of faculty, particularly at research intensive institutions, challenge the model of the renaissance person commanding all they need to design and build courses in today’s university environment. Most faculties in disciplines other than cognitive or educational psychology find themselves in the classroom with no background in learning sciences. Designing engaging learning experiences demands a motivational context grounded in the identity of the learner and a grasp of technical, psychological, disciplinary, and practical group management knowledge that exceeds what can reasonably be expected of the academic training of today’s instructor or tenure track faculty. Four themes shape this analysis of learning engineering as a response to this challenge: the history of learning technologies; recognition of what an engineering skill set brings to course development; the role of learning sciences in the design of learning activities; and how the art of design deals with the complexity of real-world situations and their resistance to experimental control.

Means’s chapter, “Tinkering Toward a Learning Utopia: Implementing Learning Engineering,” articulates how the capability of digital learning systems to collect data as people use them, coupled with advances in data science, offers exciting opportunities to improve learning. The learning data available for research has expanded dramatically in terms of the amount of detail around individuals’ learning processes, the number of people learning with systems that generate such data, and the ability to run a series of rapid experiments online. Learning system data can be used in feedback loops that inform the design of learning technology products, students’ future learning approaches, instructor practices, and the knowledge base around human learning.

Bringing together teams of collaborators with different kinds of expertise—teaching, subject matter knowledge, instructional design, and data analysis—is a prerequisite for realizing the full potential of learning system data. They characterize such work as collaborative data-intensive improvement research (CDIR) to signal their twin commitments to doing research with, rather than on, educators and to using data to help education systems improve. This chapter describes several of our CDIR collaborations, including the successes and the challenges encountered in their execution. These include the rush to premature impact evaluation, difficulty synchronizing research activities with product development cycles, and a lack of evaluation capacity within educational institutions. These challenges and possibilities for ameliorating them are discussed from an organizational change perspective. Finally, the chapter makes a case for the importance of obtaining consensus around key learner outcomes and of having valid and reliable methods for assessing those outcomes.

Roberts and Miller’s chapter, “Learning Engineering Teams: Capabilities and Process,” discusses how effective improvement of learning outcomes through the implementation of evidence-based interventions requires the coordination of many stakeholders from across an organization. This chapter addresses the
capabilities a learning engineering team should contain and the processes for engaging stakeholders in an iterative process to improve learning outcomes. They recommend a learning engineering team should have the ability to: (a) assess learner and faculty needs and advocate for programs, resources, and services to meet those needs by working with appropriate individuals and groups across the university and collaborating with other universities and technology vendors; (b) apply instructional design theory and utilize established, curriculum-development methodology to help instructors and programs achieve their learning goals; (c) lead production, management, deployment, and quality assurance of digital assets; (d) conduct research that generates opportunities for intervention, assesses usability, optimizes outcomes, measures outcomes longitudinally, and synthesizes contemporary approaches; and (e) understand the behavior of faculty, students, and other stakeholders as a key variable in the sustainable uptake, adoption, and implementation of evidence-based interventions. In terms of process, they propose a learning engineering team engage contemporary models (e.g., ADDIE, design-thinking) and approaches to developing innovations but remain flexible with regards to the particular innovation’s stakeholders as well as the particular phase of development. For example, early collaboration with faculty stakeholders might rely on principles of learner-centered design thinking, while prototyping and production phases involving broader technical teams and stakeholders might adopt agile or agile-like frameworks for rapid iteration.

**Cases of Practice**

Leitner, Nesson, and Walker’s chapter, “From Artisanship to Learning Engineering: Harvard DCE’s Framework for Improving Teaching and Learning,” discusses how Harvard’s Division of Continuing Education (DCE) maintains a commitment to learner-centric design and delivery of courses while applying insights that link evidence and structure to improved course quality, student success, and administrative efficiency. An approach to learning engineering is described that enables asking big-picture questions about the design and delivery of courses and applying answers that amplify the impact of faculty leadership, knowledge, and effort. DCE’s classrooms and instructional technologies are instrumented to provide a rich cascade of clickstream and other data, which provides the basis for continuous improvement of course design and enables flexible, timely adaptation during course delivery. A partnership model between faculty and learning engineers enables integrating complex, higher-level relationships between course structure and instructional methodology with practice based on evidence from the dynamic interaction that occurs during teaching and learning.

Jennings’s chapter, “Personalization to Engage Differentiated Learners at Scale,” discusses how personalization can make learning more relevant and meaningful but has yet to be embraced by mainstream online instruction. With the arrival of MOOCs, a one-size-fits-all curriculum serves to disengage all but
the most highly motivated learners. Google’s Analytics Academy includes users from different business types, job roles, and experience levels from multiple countries speaking different languages. Data shows that some segments of users are less engaged and that most students “prospect” the course for relevant information. To increase engagement, completion, and satisfaction metrics, the Academy utilizes a personalization model that suggests specific learning paths through courses and customizes lesson content for users.

Martin and Trang’s chapter, “Creating Personalized Learning Using Aggregated Data from Students’ Online Conversational Interactions in Customized Groups,” describes several platforms that move beyond the limitations of conventional, online teaching by using live and archived streaming instruction coupled with interactive communication chat channels. In addition to providing knowledge from instructors and peers, these platforms are configured to generate teaching and learning assessment datasets. Deep machine learning using unique datasets derived from each learner’s communication tendencies and social interactivity may help optimize the learning environment for that individual, provide guidance and support, as well as enable the use of cognitive memory maps to store and retrieve a student’s academic journey.

Goel and Polepeddi’s chapter, “Jill Watson: A Virtual Teaching Assistant for Online Education,” presents a case study from the Georgia Tech Online Master of Science in Computer Science (OMSCS), a low cost, easily accessible, accredited degree program. One recommendation for improving learning and retention in MOOCs is to enhance the interaction between the teacher and the students. However, the number of teachers required to provide learning assistance to all students enrolled in all MOOCs is prohibitively high. One strategy for improving interactivity in MOOCs is to use virtual teaching assistants to augment and amplify interaction with human teachers. This chapter describes the use of a virtual teaching assistant called Jill Watson (JW) for the Georgia Tech OMSCS 7637 class, Knowledge-Based Artificial Intelligence (KBAI). JW has been operating in the online discussion forums of different offerings of the KBAI class since Spring 2016. By now some 750 students have interacted with different versions of JW. In the Spring 2017 offering of the KBAI class, JW autonomously responded to student introductions, posted weekly announcements, and answered routine, frequently asked questions. This chapter describes the motivations, background, and evolution of the virtual question-answering teaching assistant.

Saxberg’s chapter, “Executing the Change to Learning Engineering at Scale,” discusses how recognizing the potential value of applying learning science, and even having examples of the value of applying evidence-based approaches (learning engineering), still does not address how to make such practices stick at scale. What is known about the increasing value (and rate of change) of better skills for almost all people over time is that this definitely has to happen at scale to not leave people behind (economically, socially, emotionally) for decades. A way to move a large organization from varied, traditional approaches to learning to a
more evidence-based learning engineering approach involves a series of phases. First, provide exposure to key learning development people to the possibilities and promise of evidence-based approaches—things they did not know mattered—and get a few examples done with early adopters. Second, as interest begins to build, manage and deliver education to these same learning development folks across the whole organization in a consistent way to share language ideas and even evidence-based processes. Third, with these key professionals ready to help, get the whole organization to mobilize effort behind the changes-over-time required, especially including the general managers who allocate effort and resources across all the activities of the organization. Finally, establish strong evaluation practices and cycles to make this new way of focusing on evidence-based learning measurement and success “the way we do things around here.”

Sarma’s chapter, “Rethinking Learning in the 21st Century,” discusses the evolution of learning over time. In the last century, the modern lecture-based classroom and teaching have become more pervasive, displacing earlier educational approaches of apprenticeships, mutual instruction, and hands-on work. Yet we know that those historic models are closer to our theories and evidence about how people gain deep understandings than are instructional strategies involving teaching by telling and learning by listening. Sarma describes online learning as a new “instrument in the orchestra,” an emerging, novel model that enables entirely new combinations of educational experiences.

Following these two sections on Dimensions of Learning Engineering and Cases of Practice, a concluding chapter by Richards, Saxberg, and Dede summarizes cross-cutting themes, describes Digital Teaching Platforms as an infrastructure learning engineers are creating, and delineates important next steps in research and practice for the learning engineering field.

References


Bourke Distance Education Centre. (2002). Bourke School of Distance Education. Bourke Distance Education Centre.


Jeffries, M. (n.d.). *IPSE: Research in Distance Education.* IHETS.


PART I

Dimensions of Learning Engineering
This page intentionally left blank
THE ROLE OF THE LEARNING ENGINEER

Phillip D. Long

We have a long and complex history with technology in support of learning in academe. Ambivalence about technology in education goes back at least to Socrates. In Plato’s *Phaedrus*, Socrates tells the story of the ancient Egyptian god Theuth, inventor of letters, and what the god and king Ammon (Thamus in Greek) said to Theuth about his invention:

This discovery of yours [writing] will create forgetfulness in the learners’ souls, because they will not use their memories; they will trust to the external written characters and not remember of themselves. The specific which you have discovered is an aid not to memory, but to reminiscence, and you give your disciples not truth, but only the semblance of truth; they will be hearers of many things and will have learned nothing; they will appear to be omniscient and will generally know nothing; they will be tiresome company, having the show of wisdom without the reality.

(Dialogues of Plato, *trans.* 1892, pp. 275–277)

Since the invention of writing, from the Gutenberg printing press to digital representations of words, sounds, and images, humans have struggled with the role, use, design, and relevance of technology in learning. This collection of chapters posits a description of a new role in the academy in support of that goal. It is a role that has been proposed off and on since it was introduced in 1967 when Herbert Simon wrote a mischievously titled article, “The Job of a College President.” Perhaps it was intended for college presidents, but I am sure he was writing as much for his fellow faculty as for presidents. There he stated plainly the thesis taken up by the learned authors in this volume (and also included in Huntington Lambert’s preface to this volume):
The most important step is to find a place on the campus for a team of individuals who are professionals in the design of learning environments—learning engineers, if you will...

The learning engineers would have several responsibilities. The most important is that, working in collaboration with members of the faculty whose interest they can excite, they design and redesign learning experiences in particular disciplines. Like all staff experts—operations analysts in business, for example—their long-run effectiveness will hinge on their ability to transmit their expertise to the line organization, in this instance the teaching faculty. Because they are experts in designing learning experiences, an important part of their skill will be directed toward devising learning experiences for the faculty. In particular, concrete demonstrations of increased learning effectiveness, on however small a scale initially, will be the most powerful means of persuading a faculty that a professional approach to their students’ learning can be an exciting and challenging part of their lives (Simon 1967, p. 77)

The factors that led to Simon to declare the way forward must necessarily result in the creation of the new role of learning engineer emerging from many parallel threads. One develops through the emergence of instructional media and design, initially referred to as visual education. A second comes from the digital world of computer aided instruction (CAI). A third evolves from design thinking. And, most recently, data-based analytics has emerged to shine a new and different light on learning. What follows is not a comprehensive history of these areas, nor is it a consideration of others that could rightfully be included. For that, more knowledgeable scientists and practitioners should be consulted. Rather, here we are looking at selected historical antecedents that might inform our thinking about learning engineers or at particular recurring ideas or patterns that might inform our consideration of this new species in the environment of the academy.

I am not unaware that my remarks have been unkind to several varieties of sacred cows…I know of no way to introduce professionalism into the learning process on the college campus without disturbing highly venerated practices and strongly held myths. (Simon, 1967, p. 77)

**Instructional Media and Technology**

In the US, the institutional use of media for instruction began at the turn of the 20th century with the founding of school museums (Saettler, 1968). Starting in St. Louis in 1905, then shortly thereafter in Reading, PA, and Cleveland, OH, these were the equivalent of district-wide media centers. In these museums, instructional technologies, then referred to as visual education, were viewed as supplemental curricula not intended to displace the teacher or the textbook.
Motion picture projectors entered the educational sphere in 1910 with the Rochester, NY, school district the first to adopt motion pictures for instruction. Thomas Edison gave one of the first of many infamous quotes on the impact of technology on learning in 1913 when he said, “Books will soon be obsolete in the schools… It is possible to teach every branch of human knowledge with the motion picture. Our school system will be completely changed in the next ten years” (as cited in Saettler, 1968, p. 98).

With the development of sound recordings and their introduction into film, the field rebranded itself as the audio-visual instruction movement (Finn, 1972). It is worth a brief note here that the lure of technology, even in the 1920s and 30s, seemed immune to financial losses. By 1930, nearly $50 million had been lost by investors in educational media companies, which according to Finn was only partially due to the crash of 1929. This is a recurring theme in the history of learning technologies. Another major thread in the development of instructional technology was the creation by Sydney Pressey of the teaching machine (Pressey, 1927). Pressey’s machine, a typewriter-like device, presented a multiple choice question with four answers. On the other side of the machine were four keys for the learner to register their answer. The user response was recorded, and the next question presented. The responses were then scored and presented back to the learner.

This was the realization of an idea first expressed by Edward Thorndike at the start of the century. Thorndike described his wish for a machine that, “if, by a miracle of mechanical ingenuity, a book could be so arranged that only to him who had done what was directed on page one would page two become visible, and so on, much that now requires personal instruction could be managed by print” (Thorndike, 1912, p. 165). Here were the seeds of not only programmed instruction but also competency- or mastery-based learning. Pressey’s machine implemented this idea with a lever that forced the learner to respond to a question before moving on to the next one and failed to advance until the correct answer was given, combining immediate feedback with advancement only on correct responses.

In 1932, Pressey wrote, “Education was the one major activity in this country which has thus far not systematically applied ingenuity to the solution of its problems” (p. 668). He was confident that the machine he developed would lead to an “industrial revolution in education” (p. 672). But, as with educational films, the Great Depression, followed by the Second World War, diverted attention away from advancing this notion further.

Media itself was thought to be the primary contributor to learning, an idea consolidated in the 1930s and the result of the supposed value of realism in the presentation of visual information. A hierarchy of media types was introduced, from abstract representations to photorealistic images, their value ascending in that order (Heinich, Molenda, Russell, & Smaldino, 1999). This is a notable precursor to the learning hierarchy developed by Edgar Dale in 1946 called the
Cone of Experience (Dale, 1946)\(^2\) and one of the more resilient “vampire” learning theories.\(^3\)

Radio also developed in the 1930s as the next shiny bauble of learning technology. It was to become, along with film and television, “as common as the book and powerful in [its] effect on learning and teaching” (Morgan, 1932, p. ix). However, its impact on learning and schools was negligible. What was not negligible was the recognition that clear objectives were needed to facilitate learning whatever the mode of instruction. Robert Tyler (1933) suggested each objective must be defined in terms that clarify the kind of behavior the course should help to develop. Tyler directed an eight-year study of this process in a set of public schools. He found teachers had great difficulty articulating the objectives of their instruction. Without behavioral terms to describe desired learning outcomes, it was hard to evaluate the effects of teaching.

World War II ushered in extensive investment in, and use, of training films for US military personnel. While they were never formally assessed, there was street credibility from instructors and soldiers alike that these were among the more effective tools, relative to other methods at the time, to prepare them for what lay ahead. More significantly, training films scaled well given the daunting task of training tens of thousands of soldiers, airmen, and sailors for war. Similarly, while men went off to fight, industries in the US producing war material were in high gear and trained workers in demand. The US government produced 457 training films from 1941–1945 (Reisler, 2001) to meet the needs of war industries.

The perceived success of audio-visual training materials reinvigorated educational interest in these tools, and for the first time this included research. Much of this research compared receiving the same lesson through different media types. These media comparison studies (Clark, 1983) came to the same conclusions that online learning demonstrated decades later (US Department of Education, 2009). Based on meta-analysis and other studies, Richard Clark, an educational psychologist from USC, found the following:

> Consistent evidence is found for the generalization that there are no learning benefits to be gained from employing any specific medium to deliver instruction. Research showing performance or time-saving gains from one or another medium are shown to be vulnerable to compelling rival hypotheses concerning the uncontrolled effects of instructional method and novelty. (Clark 1983, p. 445)

Clark is saying that the type of media delivery is not a significant factor in the performance of the learner. Rather the learning design seems to be significant, not the use of a video or an animation or a set of images with text.

The successes of scientists in their contributions to the war effort were profound. A reflective observation on the pace of scientific knowledge challenging the capabilities of the researcher was written by Vannever Bush at the war’s end
and published in *The Atlantic Monthly* in 1945. He described the challenge presented by this exploding quantity of information:

> The difficulty seems to be, not so much that we publish unduly in view of the extent and variety of present day interests, but rather that publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships.

*(Bush, 1945, p. 2)*

He was influenced at the time by the advancements in photography, cameras, and photocells, as well as the miniaturization of these complex components made reliably and cheaply as a result of advances in manufacturing. Bush envisioned an associative selection system to record and retrieve information, breaking the existing hierarchical models of information storage to mimic the thinking he described as associative trails.

Specialization brought with it increasing difficulty for scientists to stay abreast of their narrowing fields of study. The memex, as he termed the device that afforded electromechanical augmentation of information storage and retrieval, could offload to technology what he considered mechanical tasks and free scientists to pursue the creative work that only they could achieve. But his thinking was still anchored in the mold of the individual researcher now technologically supported in his solo pursuit of discovery. Another several decades would pass before these ideas began to change.

In the 1950s, new communication models emerged (Shannon & Weaver, 1963). These models argued that all the elements of the communications process, not just the media, are key to understanding the behavior of the system. While this groundbreaking work set the stage for information theory, cryptography, and, ultimately, digital computers, the impact on the instructional design community was minor. It was in fact overwhelmed by the rise of instructional television.

In the mid-1950s, B. F. Skinner published “The Science of Learning and the Art of Teaching” (Skinner, 1954). This introduced programmed instructional materials. Data identified as measuring the effectiveness of the programmed instruction were collected, instructional weaknesses were identified, and the materials were revised accordingly. Formative assessment (a term to arise years later for this activity) was coupled with rapid revision of content. Closely on the heels of this, Tyler’s idea that educators should write out their learning objectives resurfaced (Mager, 1962). These included a description of desired learner behaviors, the conditions under which the behaviors are to be performed, and the standards (criteria) by which the behaviors are to be judged (Mager, 1997). This time the idea took hold. This triple construct—the desired learning
behaviors + the conditions under which they are performed + the criteria by which they will be judged—infuses modern-day learning design practices.

Complementing this, Bloom, Engelhart, Furst, Hill, and Krathwohl (1956) asserted that various learning outcomes could be classified by the type of learner behavior sought on multiple dimensions and the instructional methods that were best to achieve them. Extending work done earlier in the US Air Force, he developed a classification system for three domains: the cognitive, the affective, and the psychomotor. Statements of educational objectives were arranged in a hierarchy from less complex to more complex. What a student should know was associated with a level of knowing using sets of verbs and illustrated with a behavioral statement. For example, level two was “comprehension,” with a definition of “student translates, comprehends, or interprets information based on prior learning.” Verbs associated with this level of understanding included illustrate, generalize, explain, summarize, paraphrase, describe, and interpret. A behavioral description might say, “The student will explain the purpose of Bloom’s taxonomy of the cognitive domain.” The practical driver for the work was to more efficiently create tests using a common language around domains of knowledge, starting with the cognitive. It continues to be widely used today, updated from time to time, for example by Lorin Anderson and David Krathwohl in 2001, who added interactions between cognitive processes (verbs) and knowledge content (nouns). They also expanded the knowledge domain into four dimensions: factual, conceptual, procedural, and metacognitive.

How does one determine if a learner has acquired the behaviors that programmed instruction was intending to teach them? Until the early 1960s, most testing was done comparing an individual’s performance against a large population of learners. These norm-referenced-based tests were designed to place an individual’s performance in the context of a normalized reference group. An alternative approach developed by Robert Glaser (1963), criterion-referenced measures, looked at how well a student performed without respect to others. Criterion-referenced measures had the value of assessing where a student was at the start of programmed instruction and where the same student ended up afterwards.

Contemporaneously, Robert Gagne published The Conditions of Learning (Gagne, 1965), describing five domains of learning outcomes: verbal information, intellectual skills, psychomotor skills, attitudes, and cognitive strategies. What is significant is each was contextualized by specific sets of conditions that promoted these learning outcomes. Teachers created “events of instruction” that promoted specific learning outcomes. Gagne considered the intellectual skills as hierarchical—you had to learn the subordinate skill before you could understand the superordinate skill to which it was foundational. This provided an intellectual justification for faculty to structure their courses with the intention by which disciplinary concepts must be introduced and student mastery of them demonstrated before higher order concepts could hope to be understood. Learning or instructional task analysis was developed to identify key subordinate skills and continues
to be a significant instructional design methodology. In this was the harbinger of
backwards design represented today by Wiggens and McTighe (2005).

The focus on the medium of television as an educational technology grew
during the 1960s, fueled by enormous investments by the Ford Foundation. After
the Ford Foundation spent $160 million with little to show for it, Ford shifted
their focus to public television. The failed experiments in educational technology
mirror the old-wine-in-new-bottles pattern of technology inventions in the past
by using television to deliver the same learning activities, dominated by lectures,
that has characterized the classroom for centuries. Like putting the initial motion
picture cameras in the front row of theaters to film stage plays, the potential for a
grammar and syntax unique to the new technologies for learning failed to
develop and interest waned as other applications and the next technological
revolution to transform learning stole the spotlight.

The tools of learning began to shift from television to computers starting in the
1960s. Details of these developments are well covered by many authors and will
not be chronicled here (Campbell-Kelly, Aspray, Ensmenger, & Jost, 2014; Cer-
designing instruction sprouted during this period (summarized in Gustafson and
Branch, 1997). The US military adopted an instructional design strategy to
coordinate the development of training materials. Many universities established or
refocused their centers for teaching and learning to help faculty use media and
instructional design methodologies to enhance the quality of their teaching.

There was a brief period of growth in the application of instructional design in
higher education during the 1970s and early 1980s. But this quickly died down,
and many university teaching and learning centers were already being cut back or
shut down by the mid-1980s.

Online Learning

As discussed in Chapter 1 of this book, online learning is that subset of distance
education that is primarily digitally mediated through a network. Distance educa-
tion may encompass a variety of delivery methods, such as print-based, video/
audio conferencing, and Internet-based. I am not addressing the variety of differ-
ences among fully online programs, where all instruction is delivered via the net-
work, or hybrid or blended models, where some of the education is face-to-face.
Two examples illustrate the creativity and pedagogical design that led to an
explosion of growth in the late 1960s through the early 1980s: Programmed Logic
for Automatic Teaching Operations (PLATO) and the Electronic Information
Exchange System (EIES). Things then shifted toward rapid diversification of
technology, drawing attention away from its application to learning engineering.

PLATO is generally considered the first such system (Coordinated Science
Laboratory, 1960). Developed by the University of Illinois in 1960 and running
on an ILLIAC 1 computer, it used a TV for a display and a keyboard with
function keys designed for navigating the system menu. While novel, it was not entirely practical since only one user could work with the computer at a time. A two-user version was introduced a year later, but its wider impact was not felt until PLATO III in 1963.

PLATO introduced a range of innovations including a display that incorporated memory and bitmapped graphics, user created character fonts, and a touch sensitive screen for user interaction. It presaged the now familiar concepts of peripherals and applications that could be connected to the system augmenting the learner experience. A music synthesizer, the Gooch Cybernetic Synthesizer, is an example; the software applications supporting it included the music composition language OPAL (Octave, Pitch, Accent, Length), two music text editors, a filing system for music binaries, applications to play compositions in real-time and print musical scores, plus debugging tools. All of this could be attached via the video terminal. Instructional applications could be built using TUTOR, a programming language developed by a biology graduate student. An application designed to augment TUTOR called VORTAX provided text-to-speech translation and could respond to commands selecting the language of choice (using the “saylang” command).

PLATO provided an outline of things to come. The advent of the Intel 8080 processor allowed the terminal to execute programs locally, like today’s Java applets and ActiveX controls, allowing small applications to be downloaded into the terminal that augmented PLATO courseware with rich animation and other sophisticated capabilities not available using a traditional terminal (Coordinated Science Laboratory, 1960).

Rapid progress was made in developing the initial concepts of computer-based instruction along with the infrastructure that enabled a wider community to augment it. But it remained a captive of its time, both in expense (in 1972 dollars, terminals cost $12,000 each) and network limitations.

PLATO introduced one of the first online community applications, the message board. It added to this in rapid succession Personal Notes (e-mail), Talkomatic (chat rooms), Term-Talk (instant messaging), monitor mode (remote screen sharing), and emoticons. By the late 1970s, thousands of people were in online discussions.

The University of Illinois continued development of PLATO, establishing the online service Nova NET, which was bought by National Computer Systems (NCS), and which, in turn, was bought by Pearson, now Pearson Digital Learning.

As a second example, EIES was developed by Dr Murray Turoff, Starr Roxanne Hiltz, and their colleagues at New Jersey Institute of Technology in 1976. It represents the realization of computer-mediated communications (CMC) but was extended to real-time teaching, team collaboration, and even distributed conferences. This is the turn toward the social in the online environment, paving the way toward a host of issue-driven, community activity such as the Whole Earth 'Lectronic Link—better known as The Well in San Francisco—CoSy at the
University of Guelph, and Cyclops at the Open University (Turner, 2005; Banks, 2008; McConnell & Sharples, 1983).

The pattern here is initial, rapid technological development, followed by standardization and enterprise deployment. This is part of commercialization extending on to the network to engage individuals in social learning. The next phase led to learning management systems that took the emergence of content, organized it, and presented it. For some time, developments around this phenomenon focused on moving content around and providing interoperability of the content across systems. The relentless impact of the social also made feedback possible. The emergence of the social graph (people as notes in networks) and a desire to better understand how learners engaged with content and each other coalesced to regard content, networks, and objects as data. It is that data that gives the raw material on which learning engineers operate.

**Design Science/Design Thinking**

Design science was introduced by Buckminster Fuller in 1963 (Maher & Gero, 2012), but it was Herbert Simon who is most closely associated with it and has established how we think of it today. “The Sciences of the Artificial” (Simon, 1967) distinguished the artificial, or practical sciences, from the natural sciences. Simon described design as an ill-structured problem, much like the learning environment, which involves man-made responses to the world. Design science is influenced by the limitations of human cognition unlike mathematical models. Human decision-making is further constrained by practical attributes of limited time and available information. This bounded rationality makes us prone to seek adequate as opposed to optimal solutions to problems. That is, we engage in satisficing not optimizing. Design is central to the artificial sciences: “Everyone designs who devises courses of action aimed at changing existing situations into desired ones” (p. 129). Natural sciences are concerned with understanding what is; design science instead asks about what should be. This distinction separates the study of the science of learning from the design of learning. Learning scientists are interested in how humans learn. Learning engineers are part of a team focused on how students ought to learn (Mor & Winters, 2007).

Simon recognized that sciences as a rule proceed in solving complex problems through various forms of decomposition by breaking them into simpler ones. Simon described the dimensions by which decomposition was applied to design problems. Solution components of design are iteratively generated in the design process and are then tested against a set of functional requirements.

Design thinking represents those processes that designers use to create new designs, possible approaches to problem solution spaces where none existed before. A problem-solving method has been derived from this and applied to human social interactions iteratively taking the designer and/or co-design participants from inspiration to ideation and then to implementation. The designer
and design team may have a mental model of the solution to a proposed problem, but it is essential to externalize this representation in terms of a sketch, a description of a learning design sequence, or by actual prototyping of the activities in which the learner is asked to engage. All involved can see attributes of the proposed design solution that were not apparent in the conceptualization of it. This process of externalizing and prototyping design solutions allows it to be situated in larger and different contexts, what Donald Schon (1983) called reframing the design, situating it in contexts other than originally considered.

It is important when considering the learning engineer to recognize that the engineering side of the equation values and often dictates that solutions meet externally imposed design requirements. This forces us to consider the nature of those requirements—are they really what is driving the need for the solution sought or is there something else underlying or peripheral to the articulation of the problem that is being expressed as the need?

Donald Norman introduced user-centered design (Norman, 2002) to give individuals who will work with and use the design solution voice in the ultimate configuration of design processes. In a real sense, this enables the cultural context of the institution and program where the design is to be instantiated to influence the outcome. For the design to engage the user—a learning activity in our case—the learner has to be able to build a qualitative understanding of the behavior of the design. That is, it must offer affordances (Gibson, 1979), ways for the user/learner to interact with the design solution. Course designers need to consider learners and engage with them in their role as students.

As learning environments are intentionally designed in digital contexts, the opportunity to instrument the learning environment emerges. Learners benefit in terms of feedback or suggested possible actions. Evaluators can assess how the course performed on a number of dimensions. The faculty and others in the learning-design team can get data through the instrumented learner behaviors, which may provide insight into how the design is working, for whom it is working, and in what context. Machine learning, artificial intelligence, collaboration technologies, learner interaction technologies, and social computing can be called upon to provide data to refine and revise the system.

**Learning Analytics**

The most recent trend relevant to the learning engineer is the rise of learning analytics (Long & Siemens, 2011). Data collected on student progress through the college experience, from recruitment through admissions and on to progress in degree programs, has always been a part of the business of higher education. Attention has increased with the growing cost of higher education. Questions about the return on this investment have added further to the social responsibility of the academy to deliver value for money to students. The difficulty is in identifying what the value is or, politically, reaching agreement on that point.
Much has been written about the marriage of large data, statistical techniques, and predictive modeling in the context of contemporary higher education (Campbell, DeBlois, & Oblinger, 2007; Macfadyen, Dawson, Pardo, & Gasevie, 2014; Norris, Baer, Leonard, Pugliese, & Lefrere, 2008; Pardo & Teasley, 2014). Learning analytics is informed by the work of those looking at student performance in relation to progress and retention, characterized by Tinto’s theory of student departure (Tinto, 1975, 1993), and Astin’s theory of student involvement (Astin, 1984, 1993). For the learning engineer, the interactions students engage in through digitally mediated coursework and in-class activities provide the potential for interventions and data-informed actions students can take to clarify and acquire understanding. It also informs the faculty and learning engineers on learner progress, problems, and the acquisition of knowledge.

Faculty and professional staff were collecting and analyzing data from their courses and programs well before digital data-collection opportunities existed. With the transition of learning interactions mediated by digital environments, new possibilities have emerged. But even today, the instrumentation of the learning environment remains uneven.

Practical difficulties exist in formatting the data so it can be easily used by faculty in statistical analyses. A deeper issue stems from the fact that much of the currently available data are artifacts generated by learning management systems. Rather than the result of orchestrated learning interactions, the administrative actions from distributing course information, events, and other course-management tasks often dominate events collected. These data are therefore often artifacts of the managed digital environment, and their relation to how students learn may be tangential.

Recently, standards have emerged such as Caliper from the IMS Global Learning Consortium that provide interoperable data structures that can be used across digital learning environments. The number of sensors currently are few, and the applied-learning research community needs to step forward and define the kinds of sensors and their associated contextual information that learning tools and the vendors developing them need to implement.

The context of the data collection, and the policies under which they are carried out and disclosed, are only starting to get the attention that they require to explore their contribution to learning design. The primary goal is to provide better, more timely, and contextually actionable feedback to the learner. Better tracking of student progress or difficulty is of value to the learner but also to those providing the instruction, including the learning designer if different from the instructor. But there are cautions to be raised at this point that appear in the rise of predictive analytics, the algorithms they use, and the rise of “black-box” machine learning and artificial intelligence systems that guide learner activities (Long, 2017). This was described by Cathy O’Neal in her 2016 book Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. It is the tale of how algorithms developed by mathematicians, data scientists, and
statisticians have become opaque to other, equally intelligent co-participants in the creation of predictive tools, thereby generating outcomes that can have devastating consequences for their subjects.

It is also a reminder of the powerful influence that training data have in shaping the “learning” of these machine-learning algorithms. The incorporation of biases and value judgements that are often less than clear can nevertheless have significant impact. Algorithmic bias is one consequence of skewed training data used in the development of machine-learning algorithms (O’Neal, 2016; Buolamwini, 2016). It is also present in natural language processing of text from published documents and corpora (Bolukbasi, Chang, Zou, Saligrama, & Kalai, 2016; Rutkin, 2016).

Learning analytics has brought powerful statistical tools to the analysis of learning interactions. It leverages the concurrent trends of instrumented learning environment and the growth of large, online classes coming from experiments in Massive Open Online Courses (MOOCs), which captured public attention with the Stanford Machine Learning course (Ng & Widom, 2014). Big data from large online, usually distance-learning courses opened new avenues toward understanding student behavior in digital learning environments. This is both useful and requires caution, as the expectations of experimental design often go critically unexamined when developing models, feedback systems, and predictive analyses from big educational datasets. Nevertheless, it is an essential skill set in the toolbox of the learning engineer.

**Where Are We Now?**

Several themes have emerged from this history of the antecedents of contemporary thinking in learning design. First, some patterns have developed. With the exception of online learning, technology solutions often fail their initial optimistic expectations as transformational learning inventions and only briefly give rise to broader considerations of context. The role of investment capital spurs technological trends but just as quickly leaves, allowing promising work to starve. Returns are then sought on investments elsewhere after initial enthusiasms fail to reap the rapid financial payback sought. Research takes time that exceeds the patience of many investors. Interestingly, the role of foundations starts to emerge in the later part of the 1950s and 60s, and they have driven research and application development but with similar problems of attention span. At the time of this writing, with the decline in federal funding on the horizon and anticipated to accelerate, the influence of foundations is of growing significance (Hess & Henig, 2015).

We have seen one aspect of the influence of learning sciences in terms of behavioral psychology on instructional design. Cognitive science is another key influence and is contributing important insights based on research in memory, recall, and functional areas of the brain involved in emotion, executive function,
and attention (discussed further in Chapter 3, “Tinkering Toward a Learning Utopia: Implementing Learning Engineering”).

Learning analytics has raised valuable questions about the design of learning activities and their orchestration. Increasingly, we are confronting the discrepancy between effects measured in laboratory-controlled experiments and their real-world applications. Should the study of learning activities in digital environments use research efficacy as a paradigm for learning design? Should edtech decision-making affect classroom learning environments?

**Effect Sizes and Replication**

John Hattie (2008) has used meta-analyses to summarize the effect sizes of different educational and teaching practices. Yet, when trying to determine the effects of different teaching practices, assisted or not with various technologies, we are often unsure of what we are comparing. Indeed, an effect size of $\leq 0.4$ is essentially considered “noise” in these analyses (technically, this is the progress a student would make in one year in an average school in a developed country). And while there is a certain logical consistency to the larger effect sizes reported relative to our general perceptions of what is and is not effective, it is clear that if we are not careful we will fall into the trap of confusing logically consistent but unsupported relationships that often plagues the instructional design community.

To avoid this, we need more careful descriptions of the remarkable diversity and creativity of innovations going on in teaching. It is tedious to carefully document what is actually done in experimental research; we assume the shorthand of our craft is sufficient. It developed from a combination of professional convenience, the carving out of a group identity, and as a practical response to the reality that the means by which research was distributed was costly. If we could summarize methods succinctly, journal publication costs were reduced. But the cost of failing to assiduously document is painfully being revealed to us. Without more careful description and particular emphasis on the context and methods used, we face the design version of the reproducibility problem that has arisen elsewhere in the sciences (Ioannidis, 2005; Open Science Collaborative, 2015). Of course, this is not the sole way out of the reproducibility problem, but it is a significant contributor (Gilbert, King, Petti-grew, & Wilson, 2016).

Care in documenting what we are doing in classroom learning experiments must be coupled with a cultural shift toward sharing our results—not just among disciplinary researchers reading pedagogical journals but also among those who step into the classroom to teach. Without a culture of building on the work of others, innovating, testing, improving, and sharing, there are no shoulders on which to stand. The view remains obscure and horizon ever distant.
Efficacy Research versus Clinical Effectiveness

We must examine and consider adopting the pathway set by the Institute of Medicine to develop innovative treatments and therapies and their productive application to improve human health. In terms of the learning engineer, this means the development of that bridge between randomized controlled trial (RCT) laboratory experiments and clinical- or classroom-applied experiments with students. As Olsen and McGinnis (2010) argued, “Currently, the rapid expansion in scientific knowledge is inefficiently translated from scientific lab to clinical practice” (p. 1).

This is not just a problem of lack of resources. It is a problem of the paradigm we use to think about how we approach doing this research in the first place and, importantly, what is missing so that practical guidance can be drawn from it. Olsen and McGinnis explain:

Information relevant to guiding decision making in clinical practice requires the assessment of a broad range of research questions (e.g., how, when, for whom, and in what settings are treatments best used?), yet the current research paradigm, based on a hierarchical arrangement of study designs, assigns greater weight or strength to evidence produced from methods higher in the hierarchy, without necessarily considering the appropriateness of the design for the particular question under investigation.

(Olsen and McGinnis, 2010, pp. 1–2)

We are approaching a crossroads, a choice of paradigms that has the potential to recast the learning enterprise, putting the application of learning sciences, pedagogy, and design on to an accelerated pathway of effective learning practices. Until recently we had a paucity of instruments with which to gather data about the learning process at sufficient scale to do applied research that can distinguish signal from noise. We define the outputs of a system (intended learning outcomes) and the means by which they are to be achieved (instructional methods and learning sequences); we use tests and other assessment techniques to judge how well we have approximated our learning outcomes; and we use this to assess the student’s and the instructor’s performance in a course (Laurillard, 2012). But this leaves gaps. We must take advantage of the digital learning environment and measure the correlates of what we care about associated with the learner and learning process. If the idea of learning engineer is to be more than a smart job title, this must be addressed.

A Possible Path Forward

Learning engineers represent the integration of instructional (learning) design, design thinking, learning analytics, and the pedagogical practices influenced by
cognitive and educational psychology. This is itself a tall order. But it will be old wine in a new bottle if some of the underlying paradigms on which the design practice of creating learning experiences are not rethought.

We have addressed two already: (a) the need to document teaching practices carefully to mitigate the reproducibility problem that bedevils psychology and much of the natural and medical sciences; (b) the introduction of a bridging paradigm to move from efficacy to clinical, effective educational research.

The third paradigm shift, and the most problematic, is a cultural change to the collaborative development of courses. This idea was presented by Simon in his description the learning engineer 50 years ago. Initially, Simon introduced the role of learning engineering in terms of professional staff supporting faculty “whose interest they can excite” in applying new ideas to the design of learning activities. This assumes the traditional faculty role as guild members crafting and delivering their own courses with assistance when sought.

Later in his career Simon gave a lecture at Carnegie Mellon entitled “Need Teaching be a Loner’s Sport?” for the CMU Center for Innovation in Learning Distinguished Lecture Series in which he asserted, “Improvement in post secondary education will require converting teaching from a solo sport to a community based research activity” (Simon, 1996a). It is hard to tell if he really meant the implied paradigm transition that this statement suggests when he uttered these words. But the fact is teaching has remained a solo sport for the vast majority of practitioners of it. Only a few have approached the task of course development and design differently, and most of those have been in the context of for-profit institutions delivering online distance learning. One of the few well-known, non-profit institutions that has fully embraced this team-based, course-development practice is the Open University of the UK (OU). Their course design teams consist of subject matter experts, assessment specialists, learning designers and analysts, developers, and others that work to develop courses delivered both online and by tutors around the UK and elsewhere. The reality of the OU is that the diversity of delivery environments and locations makes it impractical for a faculty member to collaboratively develop courses and deliver them. However, it is time to consider that the complexity of discipline content, learning design options, delivery mechanisms, interactivity possibilities, and assessment strategies require a team approach. Whether the faculty member is left to teach using the learning environment of their choice or not is now a separate consideration. Course development and design is truly no longer a solo sport.

The Challenge

There is an echo here of the learning-technology adoption pattern recognized by David Tyack and Larry Cuban (1995). They described the “grammar of school” and its role in thwarting creative use of personal computers by teachers in K–12 classrooms. They used the simile of the immune system in the human body repelling
foreign objects dangerous to its equilibrium. Learning engineers and the paradigm shifts that accompany them are like invading viruses in the body of the academy. The academy’s immune response seeks to block their subversive and destabilizing effects and restore homeostasis to the system. One method to do this in our context is to treat learning engineers as staff in the campus center for teaching and learning (CTL). In so doing, the effects of the learning engineer can be isolated from the mainstream departmental faculty who retain individual responsibility for course design and development. After the antibodies of the academy isolate these disruptive entities and they are sequestered away, stability is restored. Faculty know how to deal with suggestions from the CTL. In short, before learning engineers can change the academy, the academy changes the role of learning engineers and all is calm.

Whether or not learning engineers can be a part of the discussion and changes to core paradigms of the university remains to be seen. It is an idea 50 years in the making. Its time may have come. The rest of this volume will help you consider that proposition.

Notes

1 This was Simon’s perception of the role at the time and the general view of most faculty. Today, there is reason to question the expectation that, compared to research support, designing instruction and effective teaching have not evolved commensurately.

2 The Cone of Experience is still cited today because of its intuitive hierarchy, and its embellished percentages of retention have morphed into the Learning Pyramid. Let me state clearly and unambiguously, this is entirely unsubstantiated bunk (Benjes-Small & Archer, 2014; Willingham, 2013).

3 “Vampire” learning theories refer to those ideas that are found to be unsubstantiated by rigorous testing yet persist in the lexicon of the general public or others whose interests are served by them.

4 The University of Michigan was the first to form a teaching and learning center, the Center for Research in Learning and Teaching (CRLT), in 1962, followed shortly after by the University of Texas at Austin.

5 Design Science has two meanings, which can be confusing. The first is a scientific understanding of design. The second is design as science, the practice of design made better by scientific methods.

6 For example, high effect-size teaching practices include providing formative teacher evaluations (0.68), spaced vs. mass practice (0.60) and metacognitive strategies (0.53). Whereas, methods below the 0.4 cut-off that Hattie uses to distinguish signal from noise include games and simulations (0.37), inquiry-based teaching (0.35), and mentoring (0.009) (Hattie, 2016).

References


In their 1995 classic book, *Tinkering Toward Utopia*, Stanford professors David Tyack and Larry Cuban reviewed 100 years of American school reform efforts (Tyack and Cuban, 1995). Despite widespread embrace of the goal of making education a potent force for social mobility and personal improvement, the authors argued a century’s worth of widely heralded education reforms that were expected to dramatically change schooling and improve educational outcomes resulted in little significant change. Their analysis of the forces within education systems that resist radical change helps us understand why small, incremental changes are more common than break-the-mold transformations within public institutions such as schools and colleges.

Today, nearly a quarter of a century later, we are still debating the potential—and the wisdom—of new visions for transforming education, many of which rely on the technological advances and infrastructure of the last quarter century. The complexity of these efforts and the very modest successes of past efforts suggest that educational transformation will be rare. But there is something new in the mix: converging trends are now motivating higher education institutions to undertake improvement efforts informed by the massive amounts of learning-process data being generated as students interact with digital systems.

First, institutions in the higher education sector are feeling a sense of urgency in response to the rising costs and reduced public investment that threaten the very survival of many colleges (Bowen, 2012). The acuteness of this situation suggests that higher education institutions may be at a juncture where striving to achieve a utopian vision of technology-enhanced access, effectiveness, and efficiency is their best hope for the future (Bowen, 2012).

Certainly, the use of digital learning in higher education is at an all-time high. The Babson Survey, which has been tracking enrollment in online college courses
for more than a decade, reports that 5.8 million American college students were taking one or more online courses in the fall of 2014 (Allen & Seaman, 2016). In that survey, 63% of college chief academic officers reported that online learning is critical to their institution’s long-term strategy. Moreover, purely online courses are only the tip of the iceberg. Many more college courses are delivered through a combination of online and face-to-face learning activities (often called blended or hybrid learning). In SRI’s research on the Next Generation Courseware Challenge sponsored by the Gates Foundation, for example, even though the learning-technology products were designed as complete “course solutions,” we found that only 14% of the instructors using these products with their students were doing so within a course offered completely online. The remaining 86% used the courseware in combination with classroom-based learning activities.

The number of courses delivered in such blended formats is not captured in federal databases such as the Integrated Postsecondary Education Data System (IPEDS) nor in the Babson Survey. In addition, learning-behavior data are available from learning management systems (LMSs) that facilitate instruction by organizing and distributing course materials, assignments, and assessments. LMSs are used in the vast majority of colleges and universities, an estimated 96% according to a 2013 survey by EDUCAUSE (Dahlstrom, Brooks, & Bichsel, 2014).

An important implication of this widespread adoption of digital learning resources and learning management systems is the potential they open up for leveraging the system logfile data produced as students interact with digital systems (Dede, 2015; DiCerbo & Behrens, 2014). We now have the opportunity to examine learning behaviors in detail for individual students and for clusters of students with similar behavior patterns to study learning in vivo, extract information that will help instructors (and the learning systems themselves) adapt to individual learner needs, and identify learners in need of additional supports early in a course so that instructors can offer that support before it is too late. All of these endeavors build on the rapidly advancing field of data science, which is now moving beyond academic research on new methods and models to provide actionable insights to higher education institutions (Civitas Learning, 2015, 2016).

These capabilities for leveraging digital learning system data to improve education are coming at an opportune time for higher education institutions. Dramatic drops in state funding for higher education and rising costs have been accompanied by federal and state calls for data transparency and for meeting specific targets in terms of degree completion and subsequent employment for students who enroll in higher education institutions (Greer, 2013). In this context, higher education institutions are more motivated than ever before to be able to predict likely completion rates for students in their various programs and to understand the points at which students fail or otherwise fall through the cracks so that they can design strategies to improve these outcomes.

At the same time, nearly two decades of research on the basic mechanisms of different types of learning and on the multiple psychological, social, and cultural
factors that affect student learning have provided insights that can be leveraged by efforts to improve learning outcomes (Sawyer, 2006, 2014). Learning researchers have pounced on digital learning system logfile data as a new resource for building and testing theories about how people learn (DiCerbo & Behrens, 2014; Gobert et al., 2012; Koedinger, Stamper, McLaughlin, & Nixon, 2013).

Finally, another development that I will suggest is providing a conducive environment for using data to improve higher education outcomes is the growing application of improvement science methods. Such methods have a long history of application in manufacturing and services industries, including healthcare, and are now gaining traction in education (Bryk, Gomez, Grunow, & LeMahieu, 2015). The remainder of this chapter will provide a description of how the sense of urgency on the part of educational institutions around improving outcomes for all kinds of students; opportunities opened up by the increased use of digital learning systems and advances in data science; and the application of improvement science processes can converge to support tinkering toward a better education system. I use the term tinkering here not in a pejorative sense suggesting a lack of seriousness, but rather in the sense of trial-and-error processes informed by experience and executed repeatedly and deliberately over time to yield improvements that may be relatively small for each iteration but accumulate over time.

It Takes a Team to Leverage Data for Student Success

Over the last five years, researchers at SRI’s Center for Technology in Learning (CTL) have been leveraging digital learning system logfile data for purposes of understanding and improving education outcomes (Krumm, Means, & Bienkowski, 2018). We began this work in the context of evaluations of the implementation of various digital learning applications. Logfile data could tell us how much time each student used a digital learning system, the consistency of use over time, when use occurred (in some systems), and which resources were accessed (again, in some systems but not others). Such measures could be triangulated with reports from teachers and students about how and how often they used the digital products and could give us early indications of the fidelity with which an intended implementation model (e.g., 30 minutes of daily use of a learning app for eight weeks) was being followed. The data often revealed patterns that prompted further inquiry. For example, in a study of how software was used in teaching writing, we found that many classes did use the software’s peer review feature, but the average peer review was just four words long—suggesting that students need explicit instruction on how to construct informative feedback (Means & Cassidy, 2017).

At the same time, our research team was becoming increasingly interested in new ways of doing education research that involve working much more collaboratively with education partners and going beyond evaluation per se to help our partners improve their educational outcomes (Penuel, Fishman, Cheng, &
Sabelli, 2011). Putting the pieces together, we have been developing a framework for education research that is:

- **Collaborative**, studying and designing *with* rather than *for* education practitioners;
- **Data-intensive**, capitalizing on the immediacy and level of detail in learning system data; and
- **Improvement-oriented**, seeking not only to understand but also to enhance outcomes.

By definition, this kind of work involves individuals in different kinds of organizations and with different types of expertise. One of the biggest lessons we learned from engaging in this type of work is just how deeply the different types of participants depend on each other to accomplish significant improvements. It will come as no surprise that most colleges and universities, yet alone K-12 school districts, do not employ data scientists immersed in the latest techniques for extracting and manipulating huge data sets. What is perhaps less intuitively obvious is data scientists’ need for insights from practitioners and from learning-science researchers in order to (a) have a good sense of the particular combinations of user “clicks” that will constitute a variable with educational meaning, and (b) make sense of the data patterns emerging from their analyses.

Digital learning systems are used within particular contexts that have explicit and implicit norms for using the learning system, rewards attached for certain system uses (e.g., completing quizzes or earning competency points) but not others, and constraints on how the software can be used. Without knowledge of these contextual features, someone looking at system logfile data may draw wildly inappropriate inferences. To take a simple example, students sometimes forget their user-name and password for a learning system, and some classes let them borrow those of classmates, with the result that usage estimates and learning behavior sequences for the account being used by multiple students cannot be interpreted.

At the same time, educators stand to benefit significantly from kinds of expertise most of them have not acquired. Those who are responsible for teaching do not have the time or the training to engage in scraping logfiles or constructing variables that can be built from logfile data elements. Similarly, while they may have very sound instincts about the kinds of changes in instruction that could enhance student engagement and learning outcomes, few educators have any formal background in instructional design, learning science, or measurement. They may have a sense of the kinds of learning outcomes they hope their students will achieve but do not often design assessments or set up experiments that would allow them to say with any certainty whether or not a change in their practice has led to improvements in terms of the type of learning they want to enhance.

In part, this is because they have not been trained in assessment design and validation or in research design, but more fundamentally, educators’ time is limited, and these activities usually are not considered central to their profession. This is not to say that every instructor should become an expert in research and
assessment, but rather to reinforce my point that educators can benefit from being part of a team that gives them access to these kinds of expertise. At an organizational level, the vast majority of higher education institutions lack processes for systematically evaluating and improving the learning outcomes of their courses and programs.

Comparing outcomes for students who received different versions of a course or program is fundamental to identifying which is superior. Being able to attribute any change in outcomes to a particular change in curriculum, instructional practices, or learning-software features requires having a comparison group and being able to rule out plausible alternative explanations for any observed differences in outcomes.

As discussed in Chapter 9, “Executing the Change to Learning Engineering at Scale,” the conceptually simplest way to do this is through random assignment of students to version A or version B treatments (for example, the old and the new version of a course or learning application), so that students in the two groups are equivalent before experiencing one or the other treatment—but it is still wise to check this assumption, especially if the number of students involved is fairly small. Such random-assignment or “true” experiments are relatively rare in education institutions, however, and it is more common for students to select the particular course section they enroll in, which complicates comparing learning outcomes for sections with and without the new approach because any observed differences in outcomes might stem from pre-existing differences between the two groups of students signing up for different course sections.

Risks that the influence of the new approach is confounded with other factors, such as students’ prior achievement levels, are even larger when historical data from students taking the course in prior academic terms are used as the comparison. Such designs without random assignment are called quasi-experimental or observational approaches, and when using them it is important to have data on student characteristics prior to the treatment to be able to equate statistically the students experiencing the different treatments under study. If students in the new version of a course were less likely to have taken and failed the course previously, had higher college admission test scores, or were less likely to be full-time students, for example, it is impossible to know whether any observed differences in course outcomes should be attributed to changes in the course or to differences in the students. (Similar arguments can be made about the course instructors if the same instructor is not teaching both versions of the course.)

Differences between student groups can be controlled for statistically if we have data for individual students and the differences are not too large, but in my experience, most colleges and faculty members making decisions about whether or not some change was successful do not incorporate this kind of data in their analyses.

This situation suggests that colleges could do a better job of evaluating the effectiveness of their course offerings if they worked with researchers to develop unbiased estimates of the impact of the different approaches they are trying over time. Making changes without measuring their impacts is more like taking a
random walk than like following a pathway toward utopia. The need to combine the understanding of intended goals and a particular educational context with research and data-analysis expertise calls for new kinds of partnerships.

**An Early Application of the Collaborative Data-intensive Improvement Model**

One of the first partnerships in which SRI researchers moved intentionally from analyzing system data for purposes of research and evaluation to analyzing it for purposes of educational improvement was a collaboration with the Carnegie Foundation for the Advancement of Teaching (Krumm et al., 2015, 2016). Under the leadership of Tony Bryk, the Carnegie Foundation has become a leading voice in the call for applying improvement science concepts and tools to education (Bryk, Gomez, & Grunow, 2010; Bryk et al., 2015). Viewing education as a system, improvement research is organized around three questions:

- What is the specific problem we are trying to address?
- What change could be made to lessen that problem?
- How can we tell whether the change we have tried is an improvement?

Improvement science stresses the importance of multiple cycles of change implementation, measurement of processes and outcomes, analysis of those data, and planning of further revisions (Bryk et al., 2015). Carnegie advocates for small tests of change that can be executed rapidly and also argues for a sequence of such tests so that organizations can keep learning and improving over time. In theory at least, each new implementation brings some degree of improvement; the more cycles implemented, the more positive the results.

Another important aspect of improvement science as practiced by Carnegie is the idea of working concurrently with multiple educational organizations all focused on the same problem and committed to the same target improvement outcomes and to sharing their data and insights within a networked improvement community (Bryk et al., 2010). Carnegie’s first networked improvement community, and the one that SRI engaged with, is tackling the problem of low success rates in the developmental mathematics courses required before students can earn credit for a college-level math course. Only a small portion of the students required to take developmental mathematics before enrolling in a credit-bearing math course ever earn the math credit they need for graduation (Bailey, Jeong, & Cho, 2010).

Starting in 2008, Carnegie brought together educators, researchers, and improvement science experts interested in working together on this problem. The collaborators designed two different course sequences (or pathways) representing alternative, intensified approaches. Both courses involved learning software originally built on the platform developed by Carnegie Mellon University’s Open Learning Initiative. In the Statway pathway, students encounter the basic mathematics they
need as they are learning college-level statistics. At the end of two semesters, they will have fulfilled their developmental math requirement and earned a college credit for statistics. (In the alternative Quantway pathway, students complete an accelerated developmental math experience in their first semester and then complete a credit-earning quantitative reasoning course the second semester.) In the 2012–13 academic year, 1,439 students in 58 course sections started Statway. Before adopting the Pathways approach, the colleges implementing Statway saw only 6% of their entering students requiring developmental mathematics earn a college-level math credit within 12 months of continuous enrollment. Of students taking Statway in 2012–13, 68% successfully completed the first semester, and 52% successfully completed both semesters, earning a college math credit.

Despite this dramatic improvement in the proportion of students earning a college math credit within 12 months of college enrollment, the Pathways improvement community was concerned about the 48% of students who still were not succeeding. Reports from instructors working with these students and from researchers who had been observing them highlighted the importance of factors other than students’ math skills per se. Some students appeared to push through the difficulties they encountered with certain portions of the math curriculum while others appeared to give up and disengage. The collaborators turned their focus to what they called “productive persistence,” that is, academic tenacity and the use of effective learning strategies.

Various team members began experimenting with interventions to boost students’ belief that they could learn math, their sense that they belonged in their math class, and the way that they set goals for themselves and monitored progress toward these goals. Since students in the blended Statway classes were using an online learning system for practice and assessment, there was the potential to use system data to obtain behavioral measures of productive persistence and to investigate the online behaviors that are associated with higher course success rates. The hope was that online behaviors early in the course could detect a student who was starting to disengage in time for the instructor to take corrective action.

The basic implementation model for Statway involved group work in face-to-face class meetings and extensive use of the learning software outside of class sessions. The curriculum was organized into modules with the idea that each module would introduce new concepts and an element of challenge (calling for productive persistence) and that the instructor would make explicit connections between the online content and the learning taking place within the classroom. In early Statway implementations, there was little guidance for the course instructors or the students in terms of the best way to use the online learning system.

The basic organization of the Statway online learning modules was to have the student read a page of material and then engage in reflection and self-explanation as he or she responded to a set of practice questions (called try these). At the end of each topic and the end of a module, the software had checkpoints, which were online quizzes. The online system captured each page viewed, when it was
viewed, each *try these* item that was attempted, when it was attempted, and whether the item was answered correctly or not. Along with pages viewed and *try these* attempts, the OLI system also captured data on the correctness of responses to the checkpoint assessments.

When SRI began engaging with the Pathways effort, our role was to analyze the data collected by the OLI platform to define online learning behaviors that (a) predict learning and successful course completion, and (b) reveal students who were not persisting productively. My colleague Andy Krumm began this work by engineering a set of features related to concepts of productive persistence (Krumm et al., 2015, 2016). One of the earliest features SRI looked at, for example, was the date at which students completed each of the six module checkpoints. One of the early discoveries was the large amount of variation between Statway course sections in the average date on which each checkpoint was completed and in the way the various checkpoints were sequenced. Discussions with the instructors themselves revealed rationales for these variations, but instructors reassessed their thinking after SRI analysts provided data visualizations showing that class sections where the checkpoints were completed out of order (i.e., not following the sequence intended by the curriculum designers) had a lower proportion of students complete the course with a grade of C or higher. This simple piece of information influenced instructors and instructor trainers in their next iteration of the course.

SRI analysts also conducted more complex data analyses examining strings of student behaviors, classified as page views (P), try these inputs (T), and checkpoint responses (C). The logic of the courseware design would suggest that student sessions would consist of P-T-C chunks, with students going back to view a page and reflect again if they did poorly on a Checkpoint quiz (producing a P-T-C-P-T-C pattern). There were student logfile data that followed this pattern, but there were also sessions that consisted entirely of a sequence of C attempts; and cases where a low checkpoint score was followed by moving to the next topic. Analysts looked at the relationship between these sequences and students’ incoming mathematics knowledge and their course success rates and at variables reflecting a student’s time on the system and use of each kind of resource relative to other students in the same class (Krumm et al., 2015, 2016). In this way, data analysts were able to provide a measure of productive persistence based on learning behaviors with the software rather than students’ answers to questionnaires about their beliefs or usual behaviors.

Starting in the fall of 2015 SRI data analysts, Carnegie researchers, and Statway instructors participated in a series of workshops for the purpose of collaboratively designing data products, based on system logfile data, that faculty could use while teaching the course (rather than retrospectively, as SRI’s original analyses were done). The goal was to jointly produce a combination of data products and instructor practices that would result in a higher proportion of students working productively through the Statway learning modules. Faculty at the workshops
developed paper-and-pencil representations of data product prototypes that Carnegie and SRI researchers later implemented within the learning software.

A Five-Phase Process for Collaborative Data-intensive Improvement

The Pathways example illustrates the potential for leveraging fine-grained learning-system data for educational improvement and the many different kinds of players (data scientists, education researchers, instructional designers, faculty developers, instructors, and college leaders) involved in this kind of work. As we have engaged in using data analytics to support improvement in the context of multiple partnerships, we have learned some lessons about the conditions for successful collaborations and the kinds of activities that need to precede other activities. We have formulated our approach into a sequence of phases, as shown in Figure 3.1. These phases do not always occur in this exact order, and sometimes a particular finding or a major change in one of the partner organizations will necessitate moving back to an early phase, but we have found that this sequence is helpful for setting out plans and expectations at the start of a collaboration.

Phase I, the Prepare phase, is actually one of the most difficult to execute well and also one of the most critical. Collaborators come together because of a shared interest in some educational challenge, but often their sense of that challenge is heartfelt but underspecified. A group might coalesce around a desire to improve the educational experiences of first-generation college-goers, for example, but some participants might attribute the problem to poor student preparation for gateway courses while others feel their campus is deficient in promoting a sense

![FIGURE 3.1 Five Phases of Collaborative Data-intensive Improvement](image-url)
of belonging and psychological safety for these students. Researchers coming to the team will have their own set of motivations, which may include external funding that sets parameters for what they can spend their time studying. Essentially there is a conundrum: a successful improvement effort needs to have a specific focus, but (a) it is risky to get specific about the problem before you understand it well, and (b) the more specific the problem statement, the fewer people will agree that they want to tackle it. Initial meetings during Phase I seek to take a group from their existing consensus around a general area of concern to an agreed-upon specific focus for their joint activities. This entails specifying an agreed-upon measurable goal.

In the case of the Pathways initiative described above, the aim was to double the proportion of college students in need of developmental math who earned a college math credit within 12 months of college entry. Other activities during this phase include making sure that the team includes all of the needed types of expertise and preparing memoranda of understanding for the partner organizations that set forth expectations as to roles and responsibilities, especially with respect to data sharing, data handling, and privacy protections.

We think of the second phase of our process as the Understand phase. At this point, the collaborators have agreed on a more specific problem to focus on (say performance of students from groups under-represented in STEM majors in gateway science courses) but need to understand the problem better in order to generate good ideas for how to improve the situation. Phase II involves using data to confirm or disconfirm the team’s understanding of the problem as well as exploring the research literature and possibly engaging in rapid data collections to explore ways the problem might be addressed.

For a group concerned with the success rates of under-represented minority (URM) students in gateway science courses, for example, it would be helpful to look at grades earned in these courses, disaggregated for URM versus other students, for multiple student cohorts (i.e., multiple years of data). Looking at data to see if the gap in course success rates for URM versus other students is similar for the different introductory science courses and for different course instructors might reveal variability across courses or across instructors that could trigger ideas about possible root causes of the problem (for example, some instructional approaches might yield better success rates than others for URM students). It would be instructive also to compare URM students’ course outcomes to those of students who are from other ethnic backgrounds who come to the college with similar levels of math and reading skills. If there was still a success gap after controlling for prior achievement, it would suggest that something other than level of prior preparation is at play. Concurrently, researchers can assemble a rapid, concise review of the research literature on factors predicting the college success of under-represented minority students in college science courses. The group might also want to gather some data from students or instructors through a short survey, interviews, or confidential in-class poll.
Phase III is the Explore stage in which collaborators dig more deeply into the kinds of data available to them, develop new data measures, and start to formulate ideas for how to improve the system that is producing the current set of disparate outcomes. If instruction is making use of digital learning products, logfile data are a major resource for this activity. Typically, this will require matching the logfile data to student records maintained by the college to get data on student characteristics and academic outcomes (such as course grades). In the example of URM student success in gateway science courses, linking logfile data to academic records would be necessary to identify which logfiles belong to URMs and to be able to append variables such as prior achievement, gender, and age, that can be useful as covariates in data analyses.

Logfile data often reveal differences in student-learning behaviors that predict student success. For example, some students access course materials online in a regular pattern, working with course materials multiple days each week. Others tend to work in spurts right before a major event, such as the mid-term or final. The collaborators in an improvement effort might choose to construct variables, such as the average number of days a week a student logs into the learning system, and look to see whether those variables discriminate between students who are more and less successful in the course.

During this stage, it is important for researchers to develop representations of the data that educators find easy to work with and to be representing data that have implications for making educational improvements. This means that the collaborators are seeking malleable variables associated with positive outcomes. In our example of trying to improve the success of students from under-represented minorities in gateway science courses, data representations showing that URM students get lower course grades are not particularly useful at this stage since you cannot change a student’s ethnicity. Representations showing relationships between study behaviors and success, on the other hand, are useful because these behaviors can be changed.

Suppose that one of the patterns the group examines in the logfile data is the actions that students take after receiving a low score on a quiz embedded in the course learning system. Some students might respond to a low quiz score by going back to the material covered in the quiz and reviewing it, while other students might just move on to the next topic. It is also possible that after getting a low score some students might spend more time with the learning system before attempting the next quiz, while other students would decrease their learning system time, perhaps avoiding engaging with subject matter they fear they will not understand. With the logfiles linked to student-record data, the team can determine whether URM students are disproportionately likely to exhibit learning strategies associated with lower likelihoods of course success. Digging into different patterns of student behavior and trying to understand the reasons behind them is a focus for Phase III.

Phase IV of our process, the Co-Develop phase, is where we depart from the practices of many learning-analytics and educational-data-mining researchers. We
work within a multi-organizational team that seeks not just to understand or predict outcomes but also to improve them. This requires coming up with one or more ideas for changes or interventions that manipulate what the team has come to believe are the root causes for the problem the team is trying to address. Returning to the example of differential success rates for gateway science courses, the team might have seen data patterns suggesting that a low grade on an interim assessment has a differential negative impact on students from groups under-represented in science fields. Experiencing a disappointing grade may lead these students to question their ability to succeed in a course and might produce anxiety that interferes with their studying. A change idea might be as simple as sending students who get a poor grade on an interim quiz a message of encouragement, perhaps reminding them that effort produces competence, that there are tutoring services available on campus, and that they can still do well in the course. In Phase IV, it is critical that practitioners share their knowledge of the context in which these courses are given and the specifics of campus resources and instructor practices with the team designing the interventions to try out.

Phase V, Test, marks the culmination of the initial cycle of a data-intensive improvement process. This is where the intervention is tried out on a small scale and data are collected to see if there are indications that it is having the intended effect. For example, half of the students in an introductory chemistry course might be assigned to a condition in which they receive messages of encouragement after poor performance on an online quiz while the other students experience the course as usually taught. Logfile data for students who had a poor quiz performance can be examined to see if those receiving the messages of encouragement spent more time online the next week than those who did not get the messages, and these data can be disaggregated for URM and other students. The collaborators would want to review the data to see whether the messages appear to be helpful generally and, specifically, whether they appear helpful for URMs.

Other kinds of data can be brought to bear also. For example, some of the students who had a poor performance on a quiz somewhere along the line might be interviewed concerning their perceptions of the course and how they are doing in it. The subset of these students who received the online messages might be asked what they thought of the messages. Both the learning-system data and the interview data may suggest that the intervention could be improved; for example, some students might interpret the mention of the availability of tutorial services as confirmation that they are not “smart enough” for the course. Or students might think that an invitation from their instructor to come in during office hours would show more caring than referral to the campus learning center. The next step for the improvement team is to think about how to refine the intervention (or perhaps to think of an entirely new approach if it was totally ineffective) returning to Phase IV (Co-Develop) and launching another cycle of testing and interpretation.
Challenges in Doing This Kind of Work

The above, abbreviated description of our model of collaborative, data-intensive improvement processes (see Krumm et al., 2018, for a much more extensive treatment) should not be construed as an argument that such work is easy. Fundamentally, such collaborations around using data to improve student outcomes are outside the bounds of the usual practices and incentives within colleges, research units, and technology companies.

In my experience, a fundamental barrier in many cases is that most educational institutions are not organized in a way that promotes the use of data to evaluate and refine their practices. Institutional research offices (IROs) were established to provide the data required for federal and state reporting purposes. These data concern student characteristics, enrollments, and credits and degrees earned, but they do not address the particular learning experiences of students. Moreover, IROs are often thinly staffed, and they rarely work with faculty who are trying to improve their instructional practices. (External researchers sometimes find themselves in the ironic position of introducing the faculty and research office staff with whom they work to each other.) Developers of technology products too are usually not set up to conduct research on the effectiveness of their products. Many learning technology systems and applications are designed largely based on the developers’ intuitions with feedback from a handful of early customers. Usability, customer satisfaction, and the sheer number of users tend to be the metrics that developers attend to. Many technology-development organizations do not have researchers on the staff who could help them implement tests of the learning impacts of their system or application. Moreover, even those few generally larger and more established technology developers who do have a research capacity struggle to synchronize conducting efficacy research with their product development cycles.

Another challenge we have encountered is the rush to premature large-scale evaluation of the impacts of a digital learning system. Small-scale studies looking at the way in which the learning technology is used as well as variations in student outcomes for different kinds of students can help both the technology developer and the educational institutions implementing an early-stage learning-technology product figure out areas where they can improve. Investing in large-scale studies to establish evidence of the impacts of a technology-supported intervention makes sense only after several cycles of small-scale pilot testing and improvement have resulted in specification of an implementation model that includes not only the use of the digital learning system but also a description of how it is to be used and the other instructor and student behaviors needed for a successful outcome.

Finally, a recurring concern in studies of implementations of instructional technology is the lack of clarity and alignment with respect to the outcomes that the technology is intended to produce. Learning technology product selection
does not always start with a consideration of learning goals. When adopters do not have a clear learning goal in mind, they lack a basis for specifying criteria for success. This situation can easily lead to misguided evaluation efforts. For example, I have observed many cases in elementary and secondary education where a teacher or school adopts a learning application designed to foster deeper learning through inquiry or critical thinking and then tries to evaluate its success using an assessment that does not measure inquiry or critical thinking. Under these circumstances, there is no reason to expect higher scores for students who experienced the technology-based learning. Higher education differs from K–12 in that there are no externally mandated assessments that may be misaligned to a learning application’s goals, but assessment in higher education is typically ad hoc and dependent on individual faculty members’ judgment without any input from individuals trained in measurement. Moreover, when every faculty member devises his or her own assessments of course learning, it is not possible to get an apples-to-apples comparison of learning outcomes for students who do and do not experience a particular digital learning product.

**Prerequisites for Making the Organizational Transformation**

Earlier in this chapter, I made the argument that motivation to change on the part of higher education institutions, advances in the learning sciences, unprecedented availability of fine-grained learning-process data from digital learning systems, and improvement science routines make the time ripe for harnessing data to improve educational outcomes. Doing so, however, requires addressing the challenges described above. I do not want to suggest that overcoming these challenges will be easy. In fact, it will require fundamental organizational change.

In many educational institutions, collaborative data-intensive improvement activity focused on a particular problem of practice and specific instructional changes will need to be accompanied by a broader effort to change the institutional culture. A supportive culture for this work values and incentivizes enhancements of student learning to such a degree that goals for enhancing learning drive decisions about academic organization, rewards, and curriculum and instruction. This means a commitment to measuring learning impacts and to sharing and discussing the resulting data so that the data can inform future actions. The prerequisite organizational culture for this kind of work is marked by (a) mutual trust, (b) commitment to enhancing equitable learning, and (c) agreement on improvement goals and metrics.

Collaborative data-intensive improvement efforts pose challenges for the various types of experts involved in the collaboration. The data generated through collaborative data-intensive research efforts will not always put the institution, a particular department, or a particular instructor in a favorable light. Trust based in mutual respect is a prerequisite for the kinds of frank conversations around such data that lead to individual and institutional learning (Bryk & Schneider, 2002).
Even earlier in the process, trust and mutual respect will be important as different actors within the educational institution exchange ideas around the nature of the problems that are holding some of their students back and the strategies that should be tried in order to address them.

Leadership around a commitment to achieving better and more equitable learning outcomes for students is a critical factor in successful institutional change efforts. Clear articulation of this mission is necessary to mobilize the collective efforts of the many actors needed to change the way an educational institution operates and justify the expenditure of time and resources in collecting, analyzing, and reflecting upon learning data. It is difficult to obtain this kind of broad commitment to improvement absent an energetic and articulate champion at a high level of the organizational hierarchy.

Once mutual trust and a collective commitment to improving student learning outcomes have been established, the third needed cultural change is a willingness to transcend the silo represented by the individual instructor’s classroom door. If every faculty member is totally free to set the learning objectives for his or her course and to measure student learning in an idiosyncratic manner, it will be very difficult to be able to say whether any instructional innovation is an improvement because there will not be any contrasting version of the course with the same learning outcome. Some academic departments do work together to establish learning objectives and common assessments used across different sections of a course, particularly in gateway and other lower-division courses. But this is far from standard practice and, as discussed in Chapter 2, the lack of consistent learning measures in higher education greatly impedes the ability of colleges to learn how to become better at enhancing student learning.

Conclusion

Tyack and Cuban’s (1995) historical perspective on education reform efforts warns us against overpromising that the next new advance—in this case, big data from digital learning systems—will produce dramatic change. Educational change is a slow, complex process, and this observation is as accurate today as it was in the early 1990s when Tyack and Cuban wrote their book.

But improvement science recognizes the complexity of education systems and the fact that early implementations of something new within such systems are likely to encounter difficulties and unlikely to be as good as they could be. Improvement science offers a set of practices for successive refinement of new approaches that is compatible with the idea of changing an organization through a series of small but incremental improvements. (See Kurzweil & Wu, 2015, and the YouTube presentation by Tim Renick (2016) for an inspiring example of the cumulative impact of such incremental improvements at Georgia State University.) The virtue of big data lies in its potential to be turned into feedback for the individuals and organizations executing these incremental changes so that the organization learns and
improves with each new iteration. Over time—and multiple iterations—the feedback that data on learning processes and outcomes provides can enable purposive tinkering along a path toward better student outcomes.

References


LEARNING ENGINEERING TEAMS

Capabilities and Process

Craig D. Roberts and Shawn J. Miller

Introduction

Though the recent resurgence of the concept of the “learning engineer” (Rosen, 2016) might lead some to consider simply generating a job description and hiring for a single person or people designated as such, we propose that the term can also apply to an entire team focused on the continual application of data about learning and experimental design processes to improve learning. This chapter addresses the considerations required to build such a team within a higher education institution. Particularly, we situate the learning engineering concept within a university team dedicated to design, development, and overall production of educational technology projects and look at how various traditional roles (instructional designers, project managers, learning management support personnel, etc.) might also fit into a new paradigm driven by a continuous stream of new data and opportunities for new design processes informed by that data and new learning science research.

Technology Planning

Where Is the Team Situated in the Institution?

A learning engineering team should assess learner and faculty needs and then advocate for programs, resources, and services to meet those needs, working with appropriate individuals and groups across the university and collaborating with other universities and technology vendors. In order to achieve this, the team must be aware of their place in their particular institution’s organizational infrastructure. Some learning-engineering teams will find themselves placed
within a central IT structure where the focus is on enterprise-wide responsibilities and solutions. Other teams might be situated under a particular school, department, or service group (e.g., the library or a teaching and learning center), making their alignment with the institution’s mission and goals perhaps more narrow in scope. By developing a team-wide awareness of this situatedness, a team can better set and understand internal working goals through the lens of the entire university mission.

This situatedness also impacts the overall size of the responsibility of the learning-engineering team. A team located under the central IT organization might be tasked with more ongoing support issues than a team focused on an individual school’s online program. We strongly suggest that one way a learning engineer might differ from a standard instructional designer is the learning engineer’s continued concern for alignment with overall university priorities and goals despite their particular team’s situatedness in the institution. For example, for a learning engineer, it is essential to get access to several different points of data, all of which might not be automatically provided for the learning engineer or the larger team. As a result, it is essential for the learning engineer to be aware of the situatedness of their team in order to form the right partnerships. This requires a continual analysis of other university stakeholders.

Who are the stakeholders for a learning engineer? The easy answer might be: “the students of course.” Others might say: “the faculty who ultimately want to impact the students.” As student-centered learning continues to take hold at even the most traditional of institutions, an initial focus on learning and learning design might indeed create the greatest agency, but, again, this often depends on the perspectives, assumptions, and background of many key stakeholders (Rogers, 2003). Central IT organizations, university leadership, library leadership, and even department and school leadership, all involve stakeholders that can impact the overall support and success of many IT, educational technology, and more traditional teaching and learning efforts.

A learning-engineering team needs to be as agile and as free to experiment as possible. This is why it is essential for the team to develop a strong self-awareness of their situatedness within the university. What are the limits for experimentation? What are the expectations? With whom should relevant new research be shared? With whom should the team partner to conduct research in the first place? How will new ideas and innovations be implemented or stabilized if the team wants to remain focused on driving new innovations? Finally, which stakeholders can partner with the learning engineering team to provide incentives to potential faculty partners?

**Reviewing the Infrastructure**

A learning-engineering team must collaborate with groups throughout the university to identify, evaluate, and plan for educational use of new hardware and
software tools related to major instructional technology initiatives (e.g., online courses, learning management systems, multimedia and graphics tools, streaming video, data visualization software, etc.). Therefore, the team needs to understand the current infrastructure and strategy so they can help drive new and/or better technology choices by influencing the overall strategy.

A common staple at most institutions, the learning management system (LMS) is also the most widely available and widely visible teaching and learning technology. To many, it represents the lowest common denominator of such technologies, but—with the advent of learning tools interoperability (LTI) and other standards—many modern LMSs have also become linking hubs to other, more well-defined or well-focused learning technologies. Regardless of the particular LMS, the learning engineering team needs to gain some access to both the strategy setting for the system and, more importantly, any learning data that can be retrieved or shared. Some institutions also have more than one LMS, making this retrieval or sharing all the more difficult. But herein lies another value in the existence of a learning-engineering team. This team should advocate for the centralization of any learning analytics as well as begin a conversation at the institution about how such data can be shared with faculty and students to drive change and begin to develop a culture of continual assessment.

Often a large portion of the budget a university is willing to spend on educational technologies is wrapped up in the LMS licensing and support costs. Understanding this can be critical to the flexibility of the learning engineering team to suggest to the institution new or experimental technologies that might not have the same widespread adoption or impact as the LMS. Many institutions also have an entire ecosystem of connected tools and technologies that either link out directly from the LMS or run parallel to it, impacting the overall learning experience. For some institutions, these are open-source or local developments that help glue together some process or meet some non-university-wide need. In addition, individual faculty or departments might have licensing deals with publishers or educational technology start-ups, or even be using other free web-based tools, as part of their teaching and learning toolkit.

A learning-engineering team should contribute to the best practices and approaches for selecting new technologies for teaching and learning. These approaches need to go beyond simply dividing different tools’ functions into two spreadsheet columns (Feldstein, 2014) and instead use the learning engineers’ skillsets in analysis and research to suggest the guiding set of optimal functions, including modern user interface (UI) and user experience (UX) elements. A learning-engineering team can generate value for the institution through this sort of deep analysis. In addition, the team can:

- Make use of connections with peer institutions through continued collaboration or communication—be it via Slack group, email listserv, or informal annual meetings—to continue to align university strategy with the broader community.
- Participate in and share benchmarks and standards from meta-institutional sources (e.g., EDUCAUSE surveys).
- Drive institutional philosophies around technology by generating agreed-upon standards. (For example, does the university prefer open-source? Should other educational technology always connect with the LMS? What is the tolerance for privacy vs. openness?)

**Understanding the Future: The Next Generation Digital Learning Environments and Beyond**

A learning-engineering team should monitor the higher education environment for trends in how learning and outcomes are being measured in online and technology supported environments. This is one way a learning-engineering team drives university strategy and influences new technology adoption. In particular, the notion of the Next Generation Digital Learning Environment (NGDLE) (Brown, Dehoney, & Millichap, 2015) will continue to evolve as a driver of new digital education practices and technologies. Therefore, understanding the future trends as well as the university’s current learning technologies ecosystem is an essential part of the learning engineering team’s capacity. As Siemens, Gasevic, and Dawson (2015, p. 230) tell us, the “technologies selected will determine the quality of learning, the scope of teaching practices, and ultimately how well learners are equipped for both employment and engagement in democratic and equitable models of modern global society.” As the traditional LMS continues to give way to the disaggregation and multiplicity of learning technologies (be it those chosen by faculty or students, provided by publishers, or injected by start-ups or LTI), developing ways to aggregate any learning data, as well as to push for connectivity between these tools when possible, will become increasingly important.

A learning-engineering team should either have direct access to or be the unit maintaining the back-end data generated by the LMS and associated tools. In most cases, this requires a somewhat messy amalgamation of tools and expertise. New means of cloud-based data warehousing (e.g., Amazon RedShift, etc.) might make the aggregation of some data more readily accessible, but it is highly likely that many LTI tools will either lack a robust application program interface (API) or not share data in a common standard. On top of this, the learning-engineering team should also consider and build expertise in tools that can visually display and share data with staff and other stakeholders (e.g., Tableau, Looker, etc.). The learning-engineering team will also need to navigate the particular institution’s student-data privacy and sharing policies. For this reason, we strongly recommend a team form a close relationship with the institution’s data security officer and the person or staff in charge of the institution’s institutional review board (IRB).
Curriculum Development

What Is the Service the Learning Engineering Team Provides?

The learning-engineering team must add value to an institution and needs to be able to clearly and quickly demonstrate and communicate value to faculty and administrative stakeholders. What are the services the team provides?

- Consulting: The team provides expert technical and pedagogical consulting and training to help instructors and programs find innovative and effective ways to achieve their teaching goals for campus-based and online courses.
- Design and Development: The team applies instructional design theory and techniques for adult learners and utilizes established curriculum-development methodology in the design and development of appropriate, effective curriculum.
- Project Management: The team develops timelines, creates organizational strategies, and manages communication between stakeholders and other team members.
- Evaluation and Quality Assurance: The team evaluates the effectiveness of their designs and conducts testing and revision work as necessary.
- Research: The team conducts research that generates opportunities for intervention, assesses usability, optimizes outcomes, measures outcomes longitudinally, and synthesizes contemporary approaches.

Being a Partner Versus Support Unit

The services above might also reflect a more traditional online or educational technology production unit. We propose that a learning-engineering team differs from such teams by placing a higher strategic value on research and emphasizing an ongoing relationship between research, data analysis, and iterative production. Doing this might mean placing less emphasis on merely solving problems or creating products and more emphasis on engaging faculty and other stakeholders in an ongoing relationship. Therefore, a learning-engineering team should continue to develop and cultivate close relationships with faculty and administrators to identify course and program learning requirements. This requires being proactive and forging an ongoing relationship with groups of faculty and stakeholders that goes beyond them contacting the team simply when they need help. Such relationships might be developed by:

- communicating current research and best practices in effective ways (newsletters, brown-bags, shared articles, hallway conversations);
- connecting design decisions to established research when possible such as how often teams forget to connect the dots between recent research and design decisions (e.g., shorter video in online courses);
• sharing outcomes of recent experiments (why did we do this? What did we learn?);
• sharing trends and new information with stakeholders as it applies to them and their needs; and
• providing a connective hub for faculty. A learning-engineering team primarily creates new and meaningful connections between faculty and students through the design, development, and evaluation of new learning materials or technologies. But as the team develops a better understanding of faculty practices and goals, the team might also become a connective force between different faculty across the disciplines. Faculty sharing stories about new innovations with each other can be essential to any new innovation taking hold. Rogers’s work on innovation tells us that “most people depend mainly upon a subjective evaluation of an innovation that is conveyed to them from other individuals like themselves who have already adopted the innovation” (Rogers, 2003, pp. 18–19). Furthermore, a learning-engineering team might also offer additional connections between faculty and administration, as “diffusion is a very social process that involves interpersonal communication relationships” (Rogers, 2003, pp. 18–19).

Content Development

A learning engineering team might choose to engage contemporary models (e.g., ADDIE; user-centered design) and approaches to developing innovations. We would suggest that these models are useful and relevant but also recommend that the team remain flexible with regards to the particular innovation’s stakeholders as well as the particular phase of development. For example, early collaboration with faculty stakeholders might rely on principles of learner-centered design, while prototyping and production phases involving broader technical teams and stakeholders might adopt agile or agile-like frameworks for rapid iteration.

Collaborating with University Faculty to Develop and Execute Content Plans

University faculty are an essential part of any learning-engineering team, yet they generally are not part of the team’s organization administratively. This can dramatically affect the motivation of the faculty to work with the team, especially within the team’s suggested timeframe. Digital learning innovations often offer incentives to entice faculty involvement although, as Rogers points out, “if individuals adopt an innovation partly in order to obtain an incentive, there is relatively less motivation to continue using the innovation (if it can be discontinued), and so the innovation’s sustainability may be lessened” (Rogers, 2003, pp. 238–239). Depending on the institution, a faculty
member might be motivated to see a change in student learning but might also come to see the learning innovation project as diametrically opposed to their time spent on other commitments. Therefore, the learning-engineering team should be as flexible and adaptable as possible; otherwise, the team will find that faculty tend to not neatly fit into the team’s exquisitely designed project management schedule.

Teams can mitigate potential faculty bottlenecks by learning more about their faculty partners in initial meetings to better inform the overall content development plan. For example, the team might try to find out:

- What other commitments do the participating faculty have during the project’s time frame? Do they have extensive research commitments? Are they principal investigators (PIs) on any grants? What writing projects are they working on? What does their administration expect from them? When and where are they traveling during the project timeframe?
- Why are the participating faculty interested in this project in the first place? Will this get them professionally? Personally?
- How do the participating faculty typically like to work? Some faculty are very collaborative and might enjoy team meetings and discussions, while others might prefer time to cogitate on new ideas and come back to the team with reasoned responses. Some faculty prefer information in email, while others would appreciate a phone call.
- When do participating faculty typically like to work? For example, do not schedule meetings in the morning if that is when faculty members do most of their teaching.

Ultimately, faculty time can often become the unknown quantity; no amount of planning can accurately predict how long some of the faculty contributions might take.

**Leading the Scheduling, Production, Organization, and Management of Digital Assets**

Learning-engineering teams should take the lead on the scheduling, production, organization, and management of digital assets. Teams producing digital assets might be involved in one or more of the following:

- storyboarding or scripting video,
- developing animations,
- filming and editing video,
- formatting, editing, and designing web-based text and graphics,
- creating assessments, and
- testing any or all of the above.
Strong project management supports the above efforts, while also taking undue burden off of the faculty partners. Faculty can focus on content without needing to keep up with various tools and technologies to produce digital assets.

**Tracking and Communicating the Development of Course Materials Through Production and Deployment**

Learning-engineering-team project managers can make use of many technologies to enable the tracking of course production.

- Shared storage for artifacts: Cloud storage tools like Box, Dropbox, or Microsoft OneDrive are commonplace at many institutions and can provide a workable solution for easy sharing and central storage of documents, slide decks, images, and even video (if necessary).\(^3\)
- Internal and external team communication tools: While email might continue to be a default communication tool of many academic institutions, we would recommend learning-engineering teams consider using instant communication technologies like Slack, at least for the internal team. Keeping a channel open that works well on mobile devices to capture conversations or questions and even to informally share documents or content can save the team time and create new efficiencies. If the faculty partner does not want to participate in that intensive a communication flow, we recommend setting up a team email address that the faculty partner and/or the rest of the team can easily use to “ping” everyone without having to remember adding everyone else to the message.
- Task manager: Content development and deployment can be complex enough to warrant a task management tool that allows project managers to describe and assign tasks with deadlines. Some of these tools (e.g., Basecamp) also allow project managers to create templates for commonly created content.
- Development instruments: Teams might also develop other ad hoc tools that become essential to production. For example, a simple copyright tracker that helps clear content copyright might be developed in a Google Spreadsheet.
- Quality assurance (QA) instruments: A survey tool like Qualtrics provides a fairly robust and flexible way to deploy, manage, and share QA processes (e.g., surveys, rating forms, etc.).
- Asset storage: A semi-permanent home for the final digital assets is needed outside of the final delivery system (often an LMS or other platform). Platforms come and go, but the content developed should have a stable repository.

**Monitoring Deployed Course Materials for Quality and Consistency**

Adopting and implementing an evaluation strategy as part of a team’s QA process demonstrates the desire of the learning-engineering team to remain both
connected to and informed by the research. In essence, this is the piece that separates a mere educational technology-content-production team from a learning engineering team. A learning-engineering team thrives in a culture of evaluation and iteration.

Project managers and other project stakeholders should develop a strategy at the project’s outset to maintain and inform quality assurance. This strategy can be most effective when driven by agreed-upon project goals and outcomes at the outset of the project. Using these goals, project managers and leads can compare new changes or project development decisions with the original goals. On a micro-level, instructional designers should also adhere to an established set of quality metrics. For example, some instructional design teams have a strong affinity for Quality Matters as a framework for quality control of all aspects of an online course. However, at times the team might be simply developing a set of modules within the framework of a larger course or program and should be prepared to scale QA processes accordingly. For example, the team might adopt a short, five-point rubric describing what makes a single five-to-six-minute video effective and then apply that paired down framework to each video in progress. Likewise, assessments and other elements can be pre-tested by staff or willing students before launching.

Once a product, course, or set of modules is launched, the work does not simply end. Data-collection needs to immediately inform the quality aspects of the course. If all of the students are getting question number five wrong, for example, that might be the beginning of an investigation pointing to either poor question design or the lack of clear content connection to the assessment. We will discuss the various paths and levels of this type of applied research in the following section.

Research

While not intended to be all-encompassing, this section covers several modes of inquiry that your team will find useful in the iterative improvement of outcomes found to be most valuable amongst your stakeholders. Helpful resources on how to conduct applied research for online users are available through the 18F website (a digital services agency within the U.S. federal government) and in the 18F Guides.

At the outset, it is important for a learning-engineering team to assess the motivations behind the research it will conduct. Many institutional stakeholders may be familiar and comfortable with basic research, which generates findings broadly applicable across multiple contexts. We suggest the team makes some modifications in approach to engage in applied research, which generates solution-specific knowledge to inform their decisions. While action research is a defined set of approaches, its ethos may help the learning engineering team evaluate the types of research that may best facilitate their needs (Best, 1989).
More specifically, we recommend the team’s research goals be focused on immediate application, not on the development of theory on general application.

Understanding the differences between basic and applied research are important to bear in mind when hiring and managing researchers, who are often trained to formulate research questions that advance knowledge in a field (e.g., does metacognitive scaffolding increase knowledge retention and transfer?). These are important questions that advance educational practice, but they may not be the primary outcomes valued by your institution’s students, faculty, and administrators (e.g., does an online course save time/resources while maintaining or improving education quality standards?). Ideally, the who, what, and how of your team’s research questions will answer whether your interventions are alleviating the pain points being experienced by your stakeholders (e.g., our online learning intervention is improving student retention and reducing course textbook fees).

The phase of the learning-engineering project will inform the type of inquiry conducted. At the outset, the team should define what problems may be blocking the success of current learning programs. Contextual inquiry combined with institutional data can help the team better understand these challenges and the workarounds stakeholders may be currently employing to solve them. User experience research can enable the team to prototype components of a solution and collect data on whether and to what extent nascent interventions are addressing stakeholder challenges. Once deployed in pilot form, the intervention can be refined through quality assurance and psychometric approaches. These methods will qualify the intervention for application at scale, and longitudinal studies will enable the team to evaluate outcomes beyond the experimental module as well as potential interactions with other modules or learning experiences.

**Generative Research Capabilities**

We recommend that the learning-engineering team begin their research, and indeed even their work as a team, by first observing their stakeholders in the learning or instructional environments. We often observe learning technology teams seeking to apply a technology to a problem instead of the inverse. Conducting contextual inquiry at the outset enables the team to observe their learners and instructors in their standard environment (e.g., online from a dormitory or domicile, a library, a classroom, or faculty office) and then conduct follow-up interviews to clarify observations (Wixon, 1990). This provides the team a deeper understanding of the problem space around which they will develop their interventions. Otherwise, the tendency is to make possibly erroneous assumptions about the needs of learners based on prior experience and to design off-target solutions.

We also observe teams starting with institutional or learning platform data sets (e.g., LMS logs) to define learner problems. These types of data can inform what is happening (e.g., low scores on formative assessments), but they provide limited insight into why it is happening (e.g., students report misalignment between
reading objectives and assessment items). The combination of these two data types will improve the likelihood the learning-engineering team achieves the educational outcomes that matter most to its stakeholders.

Merging the aforementioned data sets from contextual inquiry and institutional or LMS sources can enable the team to develop archetypes termed “user personas” (Jenkinson, 1994; Goodwin, 2009). Originally employed in market segmentation and software development, personas are a set of fictional characters representing the motivations, goals, and challenges faced by subsets of our learners. This approach will help the team identify and consolidate the needs of learner subgroupings. Particularly relevant are the challenges being faced by your personas, often referred to as “pain points.” The learning-engineering team can begin to develop potential solutions that directly address these learner challenges.

Once these learner pain points are well defined, the team may work with students and faculty to generate potential solutions and collect feedback using low fidelity prototypes. It is helpful to begin by storyboarding the student and faculty experience learning from your intervention. This is typically a set of illustrations that narrate the actions and experiences in which your stakeholders will engage. Depending on the nature of your intervention, these documents can inform the subsequent drafting of wireframes, experiment design, content production, and software requirements.

If the team’s intervention is software-based, it would likely hand over production to an engineering group at this point in the process. The learning engineering team will want to be aware of accessibility standards that may exist for your institution or system, and it may be helpful to conduct usability testing as versions become available and time permits.

The team member conducting generative research above may have experience in the fields of human-centered design, anthropology, or user experience (UX) research. While there is an increasing presence of these fields within academic institutions, much of the work done in human-centered design and UX research takes place outside of academia. The evaluation of these capabilities may also differ from the publication and funding records common in academia. These individuals may have played a role in the research phase or iterative design of a digital product. These fields can also differ from academic approaches in the agile and iterative processes they engage, which can provide advantages to a team conducting learning engineering.

**Optimization and Longitudinal Research Capabilities**

Once the team’s intervention is ready to go live, they should structure the implementation to test hypothesis/es regarding effectiveness to improve outcomes. These research designs will align with those of more traditional education and social science research methods, which may include
randomized controlled trials and quasi-experimental designs. Individuals with this experience may reside in university departments of education, psychology, or economics.

The learning-engineering team may engage members of these respective departments to contribute to the design, administration, or analysis of this research, keeping in mind the motivations of research-track faculty discussed earlier. We also recommend the learning engineering team create a memorandum of understanding (MOU) outlining the statement of work, timeline, and output products; involving the administration overseeing the individuals engaged in this process is also important. This can help ensure there is sufficient alignment between the goals of the learning-engineering team and those of the partnering department. This type of involvement may also involve a synthesis of the literature relevant to your intervention, resulting in a set of principles or design standards to guide your generative work. Given the extensive number of publications on the use of these methods, we will not cover the topic in this chapter and instead refer the reader to Johnson and Christensen (2008).

The type of designs employed by the team will depend upon the intervention, questions, and cohort; however, we do suggest keeping two key factors in mind during the design process. First, the data generated by your design will directly inform future instructional and/or programmatic decisions. Second, the timing of experimental measurements should be sufficiently frequent to facilitate an iterative improvement process.

Implementation

Faculty and Student Buy-In

The behaviors of faculty, students, and other stakeholders are key variables in the sustainable uptake, adoption, and implementation of evidence-based interventions. Learning institutions are often highly decentralized in their decision-making processes. As a result, having funding, administrative support, and even a few faculty members willing to pilot may not be enough for your intervention to succeed and scale. Ultimately, the executors and/or subjects of your intervention (often students or faculty) must buy-in to the value of what you are proposing. Understanding how your faculty operate will be an essential component to your success and can vary widely between institution types. At a large research university, faculty may have limited time to prepare for or modify the one or two courses they teach because of competing research priorities. Conversely, faculty at the community college level may have limited time to modify their practice because they are teaching five or six courses or sections per semester. Your faculty’s existing practices for course prep, teaching, and assessment should shape the nature of your intervention, how it is applied, and your analysis of project
risk. We also caution against making assumptions about your faculty’s workflow or relying solely upon their description of instructional practices. Conducting the generative research described earlier will provide the opportunity to directly observe these processes.

Student buy-in is an increasingly important factor in the success of interventions. If students believe they are part of an experiment, they may express dissatisfaction with the course or instruction through a decrease in course evaluations despite an increase in learning outcomes. We can see examples of this on a course-wide scale from the learner-centered intervention of flipped course design and in laboratory experiments where interventions that increase retention and transfer are often perceived as more difficult and even less effective (Van Sickle, 2016; Marsh & Butler, 2014; Bjork, 1994). Student evaluations often serve as the sole metric for the evaluation of faculty teaching, and thus lowered evaluation scores (or the fear of lower scores) can drive disengagement by faculty or hesitancy to buy-in to an intervention.

The generative research a team conducts at the outset of a project can provide a tool for driving stakeholder engagement in addition to the approaches mentioned earlier in this chapter. If your faculty members and students feel they have contributed to the design of the intervention (through contextual inquiry observations, interviews, and focus groups), they are more likely to identify as contributors to the process of improving learning outcomes and less likely to instead feel they are recipients of a treatment of unknown origin or value. You can also follow generative research with a design workshop focused on faculty and students. This can serve dual purposes of generating additional solution ideas and increasing stakeholder buy-in.

**Institutional Considerations**

Many of the interventions in which the learning-engineering team engages are likely to be approaches that have proven successful at another institution. When evaluating these interventions for application at your institution, we recommend consideration of the evidentiary context, readiness for replication, and alignment with institutional needs, fit, resource availability and capacity. Blase, Kiser, and Van Dyke (2013) have developed a hexagonal tool to aid in this consideration. Many dimensions will not be captured in traditional publication formats, so we recommend contacting the individuals who conducted the study to learn more. There may be financial supports or institutional incentives (e.g., corporate partnerships, course buy-outs) that were provided to offset costs associated with developing or deploying the learning intervention. The team must carefully consider these when planning the development, implementation, and sustainability of the project. Successful implementation and scaling of your intervention should engage and synergize with the motivations of your institutional stakeholders.
Summary

We propose that the successful learning-engineering team will integrate an array of skills and processes to incrementally improve the outcomes of the population it serves. It also has a deep understanding of how its place within the organizational structure influences the ability to facilitate these outcomes. The nature and processes that drive engagement with stakeholders are key to the success of its mission. The assimilation of technology into learning environments continues to accelerate across the educational spectrum. This places particular importance on the rapid selection, validation, and scaling of digital learning interventions. We also project the organizational structure and key objectives of learning institutions will continue to evolve. The high level of institutional engagement we propose in this chapter will enable the learning engineering team to continue improving outcomes as organizations grow.

Notes

1 Interestingly, the ADDIE model’s origin has been called into question. Simply invoking the “ADDIE model” does not necessarily entail any specific method. According to Molenda (2003, p. 36), “What is emerging in the recent literature is a tendency to accept the ADDIE term as an umbrella term, and then to go on to elaborate more fully fleshed-out models and narrative descriptions. However, it should be recognized that authors who do this are essentially creating and disseminating their own models, as there does not appear to be an original, authoritative version of the ADDIE model to be revealed and interpreted. Unfortunately for the sake of academic rigor, there is no real or authentic meaning for the term.”

2 Consult Herman (2013) for examples of faculty incentives that are focused on online course delivery.

3 Taking this even further, video production teams might want to invest in tools designed to facilitate better video tracking, sharing, or commenting.

4 Quality matters is a fairly elaborate and popular set of standards for producing quality online courses (Quality Matters).

5 Information about 18F and the 18F Guide can be found, respectively, at https://18f.gsa.gov/ and https://guides.18f.gov/.

6 The User Experience Professionals’ Association has information and resources about usability testing (Usability Body of Knowledge).

7 Additional resources for how to facilitate these types of engagements are available through Stanford’s Design School and IDEO.org (Stanford Design School; Design Kit).

8 Additional resources for implementation can be found at the National Implementation Research Network (Blase, Van Dyke, & Fixsen, 2013).

References


PART II

Cases of Practice
5

FROM ARTISANSHIP TO LEARNING ENGINEERING

Harvard Division of Continuing Education’s Framework for Improving Teaching and Learning

Henry Leitner, Rebecca Nesson, and Edward Walker

Extending Harvard to part-time learners with the academic ability, curiosity, and drive to succeed.

Huntington Lambert, Dean of Harvard Division of Continuing Education, preface to this volume.

The Role of the Division of Continuing Education at Harvard

At 108 years, the Division of Continuing Education (DCE) is older than most Harvard divisions, and it occupies a unique position at Harvard. Its students are drawn from a global community of mostly part-time, adult learners. An open-enrollment policy provides those students with access to Harvard resources for a very affordable fee. The combined offerings of DCE, primarily through the Harvard Extension School (HES) and the Harvard Summer School (HSS), constitute some 1,200 courses and serve more than 22,000 unique students annually. HES is currently the third biggest degree-granting school within Harvard; approximately 4,000 students are currently matriculated toward one of its degrees. Its courses draw from the Faculty of Arts and Sciences, the School of Engineering, the Graduate School of Education, the School of Design, and the School of Law, as well as from HBX, HarvardX, and EdX.

DCE has operated profitably since its inception and carries out active experimentation and innovation in course design and teaching methodology in order to enhance the learning of non-traditional and traditional students. In a virtuous cycle, DCE’s intellectual and monetary profits are returned to the Faculty of Arts and Sciences in particular, and to Harvard at large. That return is applied to improve teaching and learning throughout Harvard and results in DCE itself being able to offer more and even better courses to its students.
Achieving success at this scale and degree of diversity depends on providing an ever-increasing level of support for faculty developing their models of courses and teaching styles, as well as on a robust technical infrastructure for creating content and for teaching that is consistent with underlying models of subject matter content and pedagogy. Thus, DCE’s organizational focus is teaching and learning rather than the subject matter of courses or research on education, per se.

**Extending the Scope of Learning Engineering**

[To achieve] the professional design of learning environments and learning experiences, the most important step is to find a place on the campus for a team of individuals who are professionals in the design of learning environments—learning engineers.

*(Simon, 1967, p. 77)*

Online learning began at DCE in 1997 with a few technical courses and a cohort of faculty and students who were prepared to cope with an immature, online-learning infrastructure. DCE now offers more than 600 online courses to a worldwide student body. Sophisticated variations of _interactive lecture, seminar, case method and problem-based_, and _laboratory_ styles for teaching and learning have evolved from rudimentary lecture capture, conference, and simulated practice platforms.

Courses taught at DCE are supported by highly interactive, network-oriented teaching and learning environments as well as classroom-based courses. Students may choose among several modalities for participating: in-person on-campus, live online, or asynchronously through a recording; fully-online courses taught live in a web-conference; and fully-online courses taught asynchronously using high-production-value, active-mastery materials. Many DCE courses combine these teaching environments to implement a flipped-classroom methodology or to optimize the combination of live interactive sessions with flexible asynchronous materials. A popular category of blended courses combines online formats with an intensive on-campus weekend.

For example, Professor V. G. Narayanan of Harvard Business School offers a credit-bearing variant of his popular HBX Financial Accounting course through DCE. In his course, students spend about half of their course hours working in a gated, self-paced, online environment with interactive course materials and receiving didactic content, opportunity for practice, automated feedback, and asynchronous interaction with peers. The other half of their course time is spent preparing for and participating in case discussions in which real-time, online student participation (via web conference) is both required and assessed. Prior to the live sessions, students submit case write-ups on which they receive feedback and grades from the teaching staff.

Larry Bouthillier and David Heitmeyer use a variation of this format for their introductory courses in HTML, web development, and JavaScript. Students watch short video lectures, do readings, and perform many small practice
problems in a gated, self-paced, asynchronous online format. Students may attend live online help sections with teaching fellows weekly and receive expert feedback and assessment on their problem sets and programming assignments.

Daniel Spratt’s online courses in biology exemplify a third major teaching methodology. Spratt teaches live in a classroom each week. Many local students choose to attend by coming to campus each week. Others from around the world attend live by joining a live web conference session. Using in-house-developed technology called HELIX Classroom, these students appear on several large screens in the room and can be seen and heard as clearly and as easily as any student in the room. They can see other students, the teacher, and course content from in the room as well as the other students who are live online. This allows Spratt to teach a highly interactive class using his well-developed, classroom-based teaching style to a real and virtual classroom of students from around the world. A third group of students who place a high value on flexibility choose to time-shift the class session and watch it later via a recording. The recordings are provided via Opencast, an open source lecture capture and video distribution platform that DCE’s software development team has substantially customized for their students. Students watching via Opencast have high definition video of the classroom and course content with automatically generated, searchable transcripts of the content and an ability to adjust view, quality, and speed. In addition, through Opencast, students are able to collectively annotate the video so that they can respond to questions from the instructor and contribute to the class conversation while watching on their own time.

Students who matriculate in an HES degree program must take a demanding course in expository writing, which is taught in an online format that is arguably the most similar to a traditional classroom experience. Expository writing courses are taught in class sizes of no more than 15 students. The courses meet live online in a web conference for several hours each week for a highly interactive, discussion-based session at which student attendance is required.

DCE’s embrace of online technology, combined with its focus on supporting the faculty and student teaching and learning experience, provides unique opportunities for developing the practice of learning engineering. Like MOOC-style courses and purpose-built online schools and programs, these processes and infrastructure generate the kinds and amounts of data that are necessary for surfacing rare events, for validating subtle differences, and for meticulously exploring parameters that affect performance and outcomes. On the other hand, like traditional schools, DCE’s faculty retain autonomy over the model and content of their courses and the intent that underlies their teaching objectives and style. They have the support required to react flexibly and creatively in response to feedback or opportunity during course delivery. They iteratively improve and alter their courses from one semester to the next based on the successes and challenges of the prior semester and the growth and change in their field of study. For purposes of research, experimentation, and measurement, it would be desirable to have fixed, unchanging learning objectives, course structure, and
assessments for DCE courses. However, DCE is unwilling to sacrifice the faculty-driven and flexible design of courses to simplify learning engineering.

To address this challenge, we have begun to elicit structured models of courses from faculty members by using Coursetune. Coursetune is a course design tool that provides faculty with an easy way to elaborate their high-level course goals, course topics and learning objectives, and the mapping of those topics and objectives onto course activities and assessments. The result is a highly structured description of the course that gives the information needed to measure the effectiveness of methodologies and interventions. It also makes it easier to identify opportunities to support faculty with course design and instructional technology services, to identify potential gaps or overlaps in the curriculum, and to establish and maintain standards for course workloads and instructor-student interaction.

At DCE, we are still in the early stages of implementing Coursetune. Over 100 courses from the Sustainability program have gone through the initial stages of modeling in Coursetune, and a handful of early-adopter faculty have done a complete model of their courses in Coursetune. The examples below, from the DCE course in Applied Online Course Design taught by Adrienne Phelps-Coco, give an idea of the types of structured data gathered from faculty about their courses.

FIGURE 5.1 High-level-goals view of the course with associated learning objectives
Goals and learning objectives that map to the DCE institutional goal of providing education that promotes career advancement are highlighted.

In the descriptions section, the properties of the project are also mapped, including the modality in which students complete the work, the type of engagement or interactivity involved, the motivation or way in which the assignment is assessed, and the estimated time-to-complete for students.

Figures 5.3 and 5.4 show which objectives are accomplished with activities that permit an online, asynchronous method of participation and those that permit real-time, online participation. Figure 5.5 shows which objectives are accomplished by activities that include students interacting directly with the faculty member or teaching staff. Figure 5.6 shows which objectives are assessed by activities that have moderate stakes.

Sources of Data

As with other online programs, the technical infrastructure at DCE automatically generates millions of data points daily from a range of primary sources. Because

![FIGURE 5.2 Mapping of the final project assignment of the course across the topics of the course for which it is an aligned assessment](image-url)
our students register for courses through an online student information system, information such as demographic data and prior academic history also is available (see Figure 5.7.). The automatically collected data is combined with curated data such as that from course evaluations and Coursetune models to create a rich dataset for research. Intelligently combining data from various sources to enable applied research while protecting student privacy is a continuing challenge. Specific sources of data include:

- OpenCast video capture system running on an Amazon Web Services platform;
- Zoom web conference system;
- Canvas learning management system and associated LTI compliant tools;
- Banner student information system;
- Salesforce relationship management system;
- Student evaluations;
- Automatic or manual transcripts of videos;
- Word clouds;
- Activity heat maps; and
- Archives of courses and student records.
Example: OpenCast Data

Useraction events are the base units of data that get generated when students view recorded lectures in OpenCast. So long as the video player window is open, a regular “heartbeat” event is logged every 30 seconds. In addition to basic video control actions (play, pause, seek, etc.), the player also logs when, for example, the user switches the video layout, adjusts the volume or playback speed, or turns on the captions. All of these events are recorded by the video player and sent back to an OpenCast user-tracking service API (application programming interface). This service identifies the user based on HTTP cookies and stores the event in the OpenCast database.

OpenCast itself does not have any functionality for processing or displaying user-action event data. So, the events are extracted, transformed, and loaded into a separate storage solution that is more amenable to analysis using a relatively simple Python script that runs at two-minute intervals and harvests the most recent entries. For each event, the script pulls out and augments the event record with metadata about the lecture being watched. During typical peak-usage time, roughly 80,000 events are harvested every hour.
The harvester stores the events in an Amazon Web Services (AWS) SQS queue where they are picked up by Logstash, an open-source system built specifically for pipelines from which data from an arbitrary source is ingested, transformed, and sent to another service or storage endpoint. In our case, Logstash is configured to do geoip lookups based on the user’s IP address to associate location data with events. Logstash also parses various parts of the user-agent string (e.g., Firefox, Chrome, Windows, MacOS, version numbers).

The last stage of the pipeline is Elasticsearch, an open source search and analytics product from which data can be accessed, queried, and visualized in a variety of ways such as those shown in Figure 5.8.

A single viewing session can be comprised of a thousand or more of individual useraction events, from which we can reconstruct who watched what and how they watched. The reconstruction can be done across an entire term’s worth of videos or at the granularity of a single student watching a single lecture. Figure 5.9 shows the number of times each student (represented by a color) viewed each five-minute segment (represented by a vertical bar) of a two-hour video. In this example, the average number of student views per segment is six.
From transcripts and viewer annotations, we also can generate word clouds to visualize the topics and relationships that faculty and students actually discussed during class (see Figure 5.10).

Enrollment data from the student information system can also be visualized in order to advise students on paths taken through various degree programs. For example, for each of the HES degree programs, the courses students enrolled in at various points in their matriculation can be identified. Figure 5.11 shows an example from the environmental management program that illustrates that more than one third of students enrolled in the Environmental Management I course as their first course:

**Strategic Objectives for Learning Engineering**

Many online learning programs have been built from the ground up or implemented as entirely separate programs. In contrast, online courses at DCE grew organically from an essentially traditional rootstock. In addition, online students at DCE remain first-class citizens. They are held to the same academic and other
standards as face-to-face students, and they receive the same transcripts and credentials, pay the same prices, and are taught by the same DCE faculty.

Expert DCE staff members do actively support faculty in the design and delivery of online courses. However, the objective of this support is to enable and contribute to iterative improvement of faculty-centered design for the course and faculty-centered capability to react flexibly and dynamically to the ongoing needs of particular students in a group of students, rather than to hold faculty to standard course designs or inculcate standard teaching methodologies. The ultimate goal of this support is to provide an actionable mapping between the structural characteristics of the course and mode of delivery, and the faculty and student

FIGURE 5.7 Infrastructure systems automatically generate data from student participation

FIGURE 5.8 Examples of simple user-activity dashboards
evaluations of their teaching and learning experience, plus evidence of effect from click-stream and outcome data.

How We Think About Learning Engineering

Where are our talented, creative, user-centric ‘learning engineers’—professionals who understand the research about learning, test it, and apply it to help more students learn more effectively.

(Saxburg, 2015)

Validating an approach to learning engineering that is based on articulating essentially artisanal models and strategies, as well as on design practices based on evidence from outcome data, raises the ante for nominal performance by

FIGURE 5.9 Views of 5-minute segments of a two-hour lecture by individual students

FIGURE 5.10 Word Cloud for one lecture from a course titled “Organizational Behavior”
learning engineers. It requires answers to questions about enabling and supporting course creation and about the awareness of faculty and the engagement of students during teaching and learning. We intend to pass between the Scylla of theories without useful consequences and the Charybdis of data without meaningful implications by identifying, understanding, and applying interventions that affect measures such as significant learning, community formation, meaningful faculty-student relationships, and more. The cascade of click-stream events that are generated by our IT department and process infrastructure as well as data from deeper sources serve as evidence of predictive reliability rather than as proof of efficacy or causation. Thus, our courses test predictions about how the faculty member’s teaching goals and pedagogical methods, paired with evidence-based engineering support from the staff and infrastructure of DCE, will affect quality, outcomes, and satisfaction.

In this approach to learning engineering, courses are designed and delivered by a partnership between a faculty member acting as a composer and teacher of the course and a learning designer acting as technical consultant and amanuensis. Targeted intervention may be applied at both composition time and performance time,
The support provided by DCE emphasizes interaction and relationships. The analysis of results and the application of design principles is primarily at the level of the full course. Their benefit for DCE is evidenced by improved organizational efficiency in the form of:

- Lower drop rates;
- Improved faculty satisfaction and/or student evaluation; and
- Higher quality feedback.

We do not yet have the integrated data system built out sufficiently for our learning designers to bring the full set of available data to bear on the evaluation of
the interventions. Our clearest successes to date include several courses in which we have scaled the size of the course to more than 150% of its original size while maintaining or slightly improving overall course evaluation ratings and maintaining or slightly improving course drop rates.

For instance, in Graeme Bird’s introductory mathematics course, which serves as one of the pre-admission courses, our design team implemented automated grading for the vast majority of student problems. This allowed the number of students to triple with only a small increase in teaching staff, improved the speed of feedback to students, and allowed a greater amount of teaching staff time to be spent in teaching interactions with students rather than in grading. We assessed the effectiveness of this change by measuring the course evaluations and drop rate, both of which showed a slight improvement.

Another example of success is David Heitmeyer’s course in HTML and web development. Our team worked with Heitmeyer to completely redesign the course by transitioning it from a classroom-based course to an asynchronous online course with online sections. Due to the density of the course material, the classroom-based setting made it difficult for students to absorb everything they needed in a single, long sitting. By switching the delivery format to short segments of presentation of materials interleaved with practice in a gated, self-paced environment, the course was able to scale to a larger size with improved student satisfaction (as evidenced by course evaluations) and no change in the drop rate. The design team hypothesizes that students are watching more of the lecture content now than when it was delivered in the long format. That prediction remains difficult to measure and evaluate because we cannot easily measure the attendance or attention patterns for those students who attended the classroom-based version of the course on campus.

Some very simple, common sense examples that illustrate how the partnership model works to improve the impact of faculty and DCE’s organizational capability and performance include the following:

- Providing students (and faculty) in a course that has real or perceived lack of organization with an explicit organization, orientation, and a guide to help.
- Suggesting and supporting the creation and use of such devices as minute papers to counteract diminished or unavailable face-to-face contact.
- Developing automation or providing training to lower faculty workload or reduce delay in feedback interactions.

The goal of interventions like these is to improve the precision and impact of faculty knowledge, experience, and intuition on student success and satisfaction not only within, but also across, courses and programs. By pushing towards the ability to design more universally applicable, model-based interventions and to measure their impact on a deeper set of outcomes, we are beginning to discover the specific benefits of the partnership model for working with faculty on courses.
These interventions add value to faculty capabilities at the expense to DCE of providing structure, improving access, or increasing efficiency. This distribution of benefit for cost reflects a general commitment by DCE to avoid adding faculty or student effort or restricting flexibility on their parts in the design and delivery of courses. In other words, we assess the benefit/cost ratio from both DCE’s and the faculty and students’ perspectives, and we seek to maximize the net benefit for faculty and students. Understanding the relationship of cost and benefit in the partnership model is the key expertise of DCE learning designers and engineers.

Bror Saxberg’s chapter in this volume, “Executing the Change to Learning Engineering at Scale,” describes four key conditions to make an organizational change to learning engineering at scale:

1. Expose the organization to possibilities.
2. Educate the key action takers.
3. Accelerate systematic effort.
4. Embed cyclic evaluation and improvement.

He describes how at Kaplan they began work with early adopters across the organization who saw the potential for the benefits in applying learning engineering to their areas of concern. Then, with energy built from awareness of external work showing the promise of evidence-based approaches to learning and learning measurement, together with the success of key initial internal projects consistent with this external promise, the stage was set for wider organizational use of these methods. Kaplan used backward design. They started from the detailed description of what they intended learners to be able to decide and do at the end of a segment of instruction and worked backwards to design what evidence/assessment data they wanted to collect to demonstrate mastery, and what practice, feedback, demonstrations, information, and overviews were required. According to Saxberg, with blueprints and training in place for evidence-based approaches to improving learning environments, the hardest part of the work begins: actually using these approaches to make a difference to learning in the operations at scale. This is the stage at which DCE’s learning engineering efforts currently center and provide a foundation for systematically improving learning over time.

The Art of Learning Engineering

Online learning technology now allows learning engineers to move beyond an initial approach that relied on reducing artisanal variety in order to simplify measurement and evaluation to reach one that preserves flexibility and facilitates dynamic control of complex activities that have complex inputs and outputs and require sophisticated forms of evaluation. The development of a learning engineering that is based on strategic objectives, as well as on tactical impact, aims to include specific features, but must take place in the context of a number of constraints.
Features

- Rich and complex data
- Faculty flexibility in leading course design and delivery decisions
- Practices and interventions that address problems of common interest to the faculty member, the learning designer, and the institution
- Design principles or decisions that map directly between underlying science and methods of evaluation
- Evaluation that combines automated and manual data collection and processing
- Cost of engineering interventions that is born by infrastructure and staff

Constraints

- Learning objectives that are implicit or ambiguous
- Program paths that are not well defined
- Sample sizes that are relatively small
- Comparisons that are mostly at the course level and likely are limited to previous and current semesters
- Sources of variance, such as changes in the course content, in the curriculum, in the student cohort, or in the larger socio-economic context that cannot be controlled
- Experimental conditions that render randomized trials infeasible

Faculty members tend to resist or ignore prescriptive methods for developing courses or inflexible pedagogical strategies. For example, a faculty member typically is not able or is disinclined to provide detailed learning objectives or to annotate statements in a syllabus with explicit motivation or guidance to students. However, faculty members do have tacit models of their courses, and they do know instinctively or from experience what kind of materials and teaching styles works for them. The key to revealing the faculty member’s model of a course is to elaborate detailed course objectives. As we described earlier, DCE learning designers are using Coursetune as an engineering tool to help faculty express learning objectives and create a more meaningful and detailed course structure than the typical flat pdf or webpage that often serves as a syllabus for courses. The data about the faculty model for a course and teaching style that we seek to elaborate include:

- Course topics and subtopics;
- Learning objectives;
- Alignment between the course and program topics and objectives; and
- Course activities.

An experimental framework for describing the variables in faculty teaching style that affect student engagement and faculty awareness reflects such features of the teaching and learning experience as the:
• Kind and degree of interactivity;
• Relevance and expression of feedback;
• Strengths and weaknesses of modalities;
• Drivers of motivation;
• Expectations and limits of time, effort, and sequence; and
• Type and importance of objectives.

These features affect both the design of courses and their delivery. Figure 5.14 illustrates a concept model for a delivery-time dashboard that enables real-time learning engineering by providing a faculty member the means to monitor the engagement of online students, detect and diagnose problems or opportunities, and intervene as appropriate.

With structure built at a level that can be applied to more than one course and more than one faculty member and evidence that maps to classes of learning objectives and teaching methods, the effect of differences in approach across many courses at once or simultaneously within individual courses can be studied. With an understanding of the relationships between course structure and delivery styles, we can better support nuances and refinements in course design or improve tools for faculty members to enhance their teaching style, as well as to detect and repair inconsistencies among the model, the course design, and the teaching method.

**Where We Think We Want To Go**

Theory, technology, and analytics have created the means for learning designers to collaborate with faculty to make better engineering decisions and

![Figure 5.14 A frame for research in learning engineering](image)
to discover deeper and more precise insights about teaching and learning. An increasing variety of data sources is being used to address sophisticated questions about the psychology and physiology of human learning, as well as to determine how best to evaluate evidence of learning in order to improve teaching and learning. The kinds of questions which we hope to discover using the framework of learning engineering that we have described include the following:

- How does real-time interaction with the teaching staff and with their peers affect student learning and the persistence of attainment?
- How does student readiness, prior attainment, or other characteristics of the student cohort affect faculty and support staff in designing and delivering courses?
- How does the autonomy of the faculty member as designer and teacher affect student learning and success?
- How does motivation imparted by social conditions and social consequences compare to motivation imparted by grades or structured interaction?

Many aspects of DCE’s approach to learning engineering are aspirational, and some are cautionary. Making progress will require us to seek advice and build on results from other researchers, and we will have to share our hypotheses and experimental successes and failures across the community of learning engineers. Among the initial questions for which we seek answers and advice are the following:

- What capabilities do we need to have available in order to enable learning engineering research?
- How do we make progress on interesting applied learning questions using our approach to learning engineering?
- What features or capabilities would make our experimental frame interesting/appealing to other researchers?
- What do other researchers need in order to make our context, data, and results usable for their work?
- What are the risks and constraints of respecting the values that are embedded in our approach?

Applied research requires reliable measures and experimental paradigms that can be used in multiple contexts. Many researchers assume that quantitative measures and the randomized controlled paradigm are standard requirements. Questions about designing and evaluating the results of experiments include the following:

- Which measures of variables such as readiness, motivation, intent, intuitive knowledge, or experience that we can use will meet requirements?
Is there a form of randomized controlled trial that is applicable in the DCE context?

Do other experimental paradigms and methods of evaluation have sufficient power and meaningful applicability in learning engineering?

Course designs that are prescriptively structured and standardized, and scalable, data-rich learning contexts have focused attention on measures of accountability for course designs, teaching methods, and evaluations of outcomes. Cheap storage and processing have reduced the time needed to gather and analyze data that is sufficiently fine-grained to enable real-time feedback control and intervention during the content creation process and during the teaching and learning process. As in baseball, so too in education: the benefits to be derived solely from better measurement and increased processing power are soon exhausted. The ultimate goal of data collection and analysis is to test the validity of scientific insight, improve the predictive force and reliability of engineering practice, and explore the generality of both across authentic contexts.

Our values compel DCE to respect the fundamental leadership role of the faculty member as course designer and the teacher and student as arbiters of educational experiences. We have extended the goals and practices of learning engineering in order to support a higher level of structure for courses and to relate the strategic intent of the faculty to specific characteristics of courses and teaching modes. We see exciting opportunities to illuminate, enable, and refine tools and methods that will increase faculty members’ ability to design and teach courses, as well as to augment their capacity to notice and react to events and trends as they teach.

References
PERSONALIZATION TO ENGAGE DIFFERENTIATED LEARNERS AT SCALE

Chris Jennings

Introduction

Google Analytics (GA) is a data-collection platform used to measure user behavior on websites and apps. Online businesses can get audience insights, better understand online user behavior, target advertising, and measure the overall effectiveness of marketing campaigns. To train users how to strategically collect and analyze data, Google launched an open online-learning platform in 2013 called Analytics Academy. The Academy teaches students how to collect and analyze website data using GA and associated measurement tools, as well as how to evaluate online marketing efforts. Academy courses are completely online, self-directed, asynchronous, and run continuously. They are free of charge and have been translated into 19 languages. Registrants may receive a certificate by passing each assessment in a course with a score of 80% or better. The course instructors are experts in the digital analytics industry who appear in videos, as well as in a live-streamed, question-and-answer session available on YouTube after course launch. If students have additional questions, they are directed to a community discussion board. To date, the Academy has launched seven courses, which range in size from 90,000 to over 800,000 registrants, depending on the specificity of the subject matter. As of August 1, 2017, Academy pages have been viewed by over 5.9 million users and have 226,806 active users per month.

In March 2017, the Academy launched a new learning application that saves user progress and certification data. Rather than the traditional xMOOC (extended massive open online courses)—a university model of courses that run within specific windows of time—the Academy made self-directed certification available year-round. The Academy also launched two flagship courses: Google Analytics for Beginners (GAB) and Advanced Google Analytics (AGA). Google Analytics
for Beginners is intended for users who are new to GA and acquaints them with data collection and set up; how to navigate GA; how to read and analyze reports; how to create and share dashboards; and basic campaign and conversion tracking. Advanced Google Analytics is the next course in the sequence and teaches more sophisticated analysis concepts such as data processing, custom implementations, segmentation, multi-channel funnels, benchmarking, and remarketing. Both GAB and AGA included new learning formats to situate users in real-world analytics problem solving. Guided tours walk students through the interface of a GA demo account with live data from the Google Merchandise Store. The courses also introduced multiple-choice quiz activities that let students apply learned concepts to the demo account, providing an opportunity to practice real-world analysis with business data.

**Academy Implementation and Data Collection**

The Academy uses GA in conjunction with Google Tag Manager for the majority of its data collection and analysis. This approach has inherent advantages and limitations compared to other data-collection solutions or learning management systems with built-in analytics. All data collected in GA is anonymous and collected in aggregate per GA product policy. Therefore, the Academy cannot see any personally identifiable information or associate individual behavior back to a particular user. The Academy also collects course metrics in a single GA account, meaning that if a student visits multiple courses within a single session, the Academy is unable to attribute behavior metrics to a specific course. Despite inconsistently exact metrics, the main advantage of using GA for learning analysis is the ability to quickly understand general user behavior and trends, compare data with previous courses, and derive actionable insights based on changes in the data.

In launching its learning application, the Academy made a number of modifications to its data collection and implementation to better understand user needs and evaluate courses. In order to save users’ course progress, assessment attempts and completions, certificates, and surveys, the Academy implemented a required login to access courses using a Google account. This associates Academy learning data with users’ GA account and provides aggregate data on assessment attempt and pass rates, course registrations, course completions, and completion rates. Users who registered for Academy courses were also prompted to complete an optional user profile that sets course messaging preferences and identifies their business type and job role. These may be used as Custom Dimensions to segment behavior data and potentially personalize course content in the future.

At the beginning of each course, the Academy prompts users to complete an optional pre-course survey where students rate their understanding of the course topics to be covered. Students are also asked to communicate whether they plan to complete the course for a certificate of completion. This course intent can also be used as a Custom Dimension to analyze user behavior for those who planned to complete...
the course. An optional, post-course survey follows each course; this gives students an opportunity to communicate their course satisfaction, state whether they made progress understanding the topics they rated in the pre-course survey, provide opinions on the learning formats used, and offer free-form feedback.

Standard GA metrics provide demographic data such as age, gender, country, and language; acquisition channels (what search, sites, or marketing brought users to the Academy); and device type such as desktop, mobile, or tablet. The Academy also uses standard GA metrics to measure course engagement, including new and returning users, average sessions per user, average pages per session, and average session duration. In addition, the Academy collects custom behavior data from its learning platform. Event tracking measures how far users proceed incrementally through lesson videos and demonstration videos, tracks progress in guided tours, records engagement with in-line practice activities, and tracks clicks on additional resources links at the end of each lesson.

**Understanding Academy Users**

Like users of many other open learning environments, Academy course participants are widely differentiated across business types, job roles, demographics, experience levels, and learning motivations. Examining Academy courses two weeks after course launch over a four-month period (April 1–August 1, 2017), showed a diverse array of business types.

Because some businesses have more than one business objective, users were allowed to select multiple affiliations in their profile. It should be noted that these business types often implement GA data collection in different ways, according to their distinct business goals. For example, business measurements for e-commerce...
could include the purchase of products or newsletter subscriptions. For *branding*, this could be user loyalty and recency. For *lead generation*, it could be filling out a contact form or following a site on social media. *Publishers* may want to measure the engagement of particular types of content or clicks on a specific article. An *online community* might want to know the average number of user posts or top discussion groups. An *online support site* may measure the completion of a guided help flow or how users rated a support article.

Within each of these businesses, employees often have different job roles (or straddle multiple job roles) that require relevant knowledge of GA. While skewed heavily towards marketers, the job roles represented in the Academy user profile are also varied.

As with business types, different job roles potentially interact with GA in different ways. *Marketers* tend to focus on how to best spend marketing and advertising budgets, how to understand the return on investment for past marketing campaigns, and how to determine which marketing efforts are most effective. *Data analysts* focus on making sure users have the best online experience possible, which includes identifying problems on the website or app, monitoring the strength of a business’s web presence, determining who are the most valuable customers, and deciding which metrics to improve. *Web developers* focus on supporting an analytics implementation and making code changes to support data collection. *Agencies* might look at tracking keywords for search campaigns or comparing their data with others in their industry. For small- or medium-sized businesses, these roles could easily overlap, requiring multiple domains of analytics knowledge.

According to GA demographics data, while the majority of Analytics Academy students are male, aged 25–34, speak English (according to their browser settings),
and reside in the United States, the Academy has a healthy diversity of course participants. Males participate slightly more than females (see Figure 6.3) and the majority of ages fall between 18 and 44 (see Figure 6.4).

Academy students come from many different countries (see Figure 6.5) and while the majority of Academy students’ web browsers are set to English, many other languages are represented in the top 10 (see Figure 6.6).

Product and industry experience also differ across Academy students. A learning survey conducted on the Academy (N = 1,000) in March of 2017 indicated

---

**FIGURE 6.3** Analytics Academy registrants segmented by gender

**FIGURE 6.4** Analytics Academy registrants segmented by age

that 37% of users have no previous experience with GA (unchanged from users surveyed two years prior). As expected, students in GAB overwhelmingly reported their experience as beginner in the course topics shown in Figure 6.7.

Advanced Google Analytics students (who should have either taken GAB or had basic proficiency with GA), also reported a significant degree of “beginner” knowledge in the requested course topics (see Figure 6.8).

Academy students could be writers looking to promote their blogs and online articles, employees at small-to-medium-sized businesses looking to grow, or employees at larger businesses or agencies who spend millions of dollars on advertising. While Academy users generally skew toward beginner-level
knowledge, the Academy learning survey indicated that, in general, 70.6% of respondents are self-taught (down from 78.5% two years earlier). Open-ended feedback communicated that users commonly learn GA from colleagues, friends, and relatives (including their children!). Without formal training, we can assume that even more experienced users have feature- or skill-gaps in their understanding of GA. Thus, Academy participants are generally at different skill levels, depending on their business and job role, their previous experience with the product, or the extent to which they have gone through formal training.

**Measuring User Engagement**

Academy participants are generally motivated to learn, since GA can potentially help grow and monetize their business or help get them a job. Of those surveyed, 67.56% indicated they are taking courses to improve their job performance, 48.60% want to help their career or job search, 42.5% want to improve their business, and 41.12% are attempting to fulfill a current job role requirement. Support tickets and social posts indicate that many participants find value in the course completion certificate to help find jobs or serve as proof of on-the-job-training. Course data four months after launch shows relatively high engagement for students in both courses (see Table 6.1).

Completion rates for the Academy are relatively high when compared with other MOOCs (Jordan, 2014). The pages-per-session and session duration suggest that users invest a significant amount of time and attention in the course, although the low sessions-per-user implies that, across students, this is confined to a single visit or two on average. As has been observed in other MOOCs, there is typically a proportion of users who are motivated to complete the course from the outset (Bergner, Chuang, Mitros, Pritchard, & Seaton, 2014). When students register for a course, the pre-course survey asks whether they intend to complete the course for a certificate of completion. This reflects the intrinsic motivation of users, as 86.4% of GAB students and 88.5% of AGA students report intending to complete each course. The Academy uses this survey response as a Custom Dimension to analyze data by those who intended to complete and those who did not. As expected, those who indicated they planned to complete the course had significantly higher engagement than those who did not (see Tables 6.2 and 6.3).

<table>
<thead>
<tr>
<th>TABLE 6.1 General engagement metrics for beginner and advanced course registrants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Course Registrations</strong></td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>GAB 170,859</td>
</tr>
<tr>
<td>AGA 50,936</td>
</tr>
</tbody>
</table>
FIGURE 6.7 Google Analytics for Beginners course registrants who reported their subject-matter experience level as “beginner”
FIGURE 6.8 Advanced Google Analytics course registrants who reported their subject-matter experience level as “beginner”
It is worth noting that browsers (who said they did not plan to complete) still viewed almost as many sessions-per-user and pages-per-session as completers across both courses; browsers’ session duration, however, was much lower (indicating less time spent on average). Interestingly, we see that a small percentage of browsers who did not intend to complete the course actually did. For students who indicated they planned to complete the course, only about 20% of GAB and 30% of AGA users had done so at the time of this study. Since we are unable to easily segment this data by registration cohorts, it is difficult to tell if this number is relatively low because users are still working through the courses or have simply given up. When segmented by female gender and by non-English-speaking users who intended to complete the course, engagement and completion rates were either at parity or higher for females and non-English users compared to general, intended-completers across both courses. The exception was AGA in which 26.47% of non-English speakers who intended to complete actually completed the course, compared to 30.64% for all intended completers. Also, engagement rates were slightly lower. This could indicate that the advanced concepts were harder to master in subtitles and translated text than the beginner course.

Since self-professed completers have declared their intent to deeply engage with the course, they are an appropriate population to segment and look for signs of disengagement to help understand what might be discouraging students. The Academy can use the signals business type and job role as Custom Dimensions to compare against the general segment of users who indicated they would complete (see Table 6.4).

Most of the examples in the Academy are e-commerce focused (and the GA demo account is based on an e-commerce business), so it makes sense that the e-commerce business type has the highest completion rate and some of the highest

<table>
<thead>
<tr>
<th>TABLE 6.2 Google Analytics for Beginners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>will complete</td>
</tr>
<tr>
<td>won’t complete</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 6.3 Advanced Google Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>will complete</td>
</tr>
<tr>
<td>won’t complete</td>
</tr>
</tbody>
</table>
engagement metrics (such as pages-per-session). In contrast, the completion and engagement rates for online support and online community segments are lower than e-commerce, lead generation, and branding. This could indicate the fact that course content and examples were generally written with ecommerce in mind, rather than other businesses, which constitute a smaller percentage of Academy enrollments. Segmenting by job role can identify similar drops in engagement (see Table 6.5).

Agencies enjoy the highest completion rates and some of the strongest engagement metrics, which could be due to mandatory on-the-job analytics

**TABLE 6.4** Google Analytics for Beginners (by business type sorted by completion rate)

<table>
<thead>
<tr>
<th>Business Type</th>
<th>Unique Users</th>
<th>Completion Rate</th>
<th>Sessions/User</th>
<th>Pages/Session</th>
<th>Session Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>137,314</td>
<td>19.63%</td>
<td>2.32</td>
<td>6.88</td>
<td>21:49</td>
</tr>
<tr>
<td>Ecommerce</td>
<td>55,411</td>
<td>19.37%</td>
<td>2.38</td>
<td>6.67</td>
<td>21:29</td>
</tr>
<tr>
<td>Lead Generation</td>
<td>42,883</td>
<td>19.08%</td>
<td>2.38</td>
<td>6.61</td>
<td>21:48</td>
</tr>
<tr>
<td>Branding</td>
<td>51,287</td>
<td>19.00%</td>
<td>2.40</td>
<td>6.63</td>
<td>21:30</td>
</tr>
<tr>
<td>Content Publisher</td>
<td>44,354</td>
<td>18.82%</td>
<td>2.42</td>
<td>6.54</td>
<td>21:07</td>
</tr>
<tr>
<td>Online Support</td>
<td>23,387</td>
<td>17.76%</td>
<td>2.40</td>
<td>6.53</td>
<td>20:39</td>
</tr>
<tr>
<td>Online Community</td>
<td>24,718</td>
<td>17.39%</td>
<td>2.40</td>
<td>6.43</td>
<td>20:18</td>
</tr>
<tr>
<td>Ecommerce</td>
<td>55,411</td>
<td>19.37%</td>
<td>2.38</td>
<td>6.67</td>
<td>21:29</td>
</tr>
<tr>
<td>Lead Generation</td>
<td>42,883</td>
<td>19.08%</td>
<td>2.38</td>
<td>6.61</td>
<td>21:48</td>
</tr>
<tr>
<td>Branding</td>
<td>51,287</td>
<td>19.00%</td>
<td>2.40</td>
<td>6.63</td>
<td>21:30</td>
</tr>
<tr>
<td>Content Publisher</td>
<td>44,354</td>
<td>18.82%</td>
<td>2.42</td>
<td>6.54</td>
<td>21:07</td>
</tr>
<tr>
<td>Online Support</td>
<td>23,387</td>
<td>17.76%</td>
<td>2.40</td>
<td>6.53</td>
<td>20:39</td>
</tr>
<tr>
<td>Online Community</td>
<td>24,718</td>
<td>17.39%</td>
<td>2.40</td>
<td>6.43</td>
<td>20:18</td>
</tr>
</tbody>
</table>

**TABLE 6.5** Google Analytics for Beginners (by job role sorted by completion rate)

<table>
<thead>
<tr>
<th>Job Role</th>
<th>Unique Users</th>
<th>Completion Rate</th>
<th>Sessions/User</th>
<th>Pages/Session</th>
<th>Session Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agency</td>
<td>18,917</td>
<td>21.47%</td>
<td>2.48</td>
<td>6.56</td>
<td>22:13</td>
</tr>
<tr>
<td>Marketer</td>
<td>74,185</td>
<td>19.80%</td>
<td>2.39</td>
<td>6.77</td>
<td>22:03</td>
</tr>
<tr>
<td>Analyst</td>
<td>46,287</td>
<td>19.66%</td>
<td>2.42</td>
<td>6.72</td>
<td>21:53</td>
</tr>
<tr>
<td>All</td>
<td>137,314</td>
<td>19.63%</td>
<td>2.32</td>
<td>6.88</td>
<td>21:49</td>
</tr>
<tr>
<td>Other</td>
<td>30,117</td>
<td>18.14%</td>
<td>2.24</td>
<td>6.86</td>
<td>20:30</td>
</tr>
<tr>
<td>Developer</td>
<td>17,965</td>
<td>16.34%</td>
<td>2.32</td>
<td>6.31</td>
<td>19:24</td>
</tr>
</tbody>
</table>
training typically required of roles in that type of business. However, similar to online support and online communities the developer role reveals a distinct drop in engagement. This potentially makes sense, since developers would likely only require a subset of the course curriculum in order to perform their job tasks.

To understand whether this is simply indicative of the GAB curriculum or a broader trend, we can apply the business type and job role Custom Dimensions to AGA as well (Table 6.6).

Content publishers have the highest completion rates in AGA with relatively lower pages-per-session and session duration, whereas lead generation and online support have slightly lower completion and engagement rates than other segments. This could indicate that the advanced course content is not serving those particular businesses quite as well as other segments. AGA engagement by job role reveals similar disengagements to GAB (Table 6.7).

Unexpectedly, those who included the other job role category had much higher engagement than other segments. This could be users looking for a change in profession or students who have not yet embarked on their careers, both of whom would have sufficient motivations to complete. Again, as in GAB, developers have the lowest completion and engagement rates across pages-per-session and session duration. It is significant that those who straddled multiple job roles demonstrated lower levels of engagement if they included developer among the roles selected, particularly for GAB (Tables 6.8 and 6.9).

To identify reasons for disengagement across the two courses, custom funnels can show how segments of users progressed through the course and where they exited.

| TABLE 6.6 Advanced Google Analytics (by business type sorted by completion rate) |
|----------------|----------------|----------------|----------------|----------------|----------------|
|                | Unique Users  | Completion Rate | Sessions/User | Pages/Session  | Session Duration |
| will complete (Content Publisher) | 8,621 | 31.13% | 3.02 | 6.47 | 24:53 |
| will complete (Branding) | 10,509 | 30.79% | 2.96 | 6.68 | 25:25 |
| Will complete (all) | 37,075 | 30.64% | 2.86 | 7.01 | 26:31 |
| will complete (Online Community) | 4,466 | 30.61% | 2.96 | 6.69 | 25:26 |
| Will complete (ecommerce) | 12,688 | 30.02% | 3.02 | 6.57 | 25:11 |
| will complete (Lead Generation) | 9,919 | 29.52% | 2.99 | 6.50 | 24:53 |
| will complete (Online Support) | 4,489 | 29.47% | 2.87 | 6.65 | 24:17 |
Comparing e-commerce paths to online support and online community paths, there is virtually no difference for how users navigate Unit 1 (see Figures 6.9, 6.10 and 6.11).

Comparing developers to agencies also shows no obvious variation between how users navigate (Figures 6.12 and 6.13).

In fact, what is striking is how over 90% of user traffic drops after two lessons. While there are clearly committed users going through sequentially (that account for the higher completion rates in later lessons), this two-lesson drop off is evident throughout all units of GAB and AGA. Starting with any lesson in the sequence, this pattern is remarkably consistent throughout both courses no matter where the funnel begins. This suggests that most Academy users do not view the course content sequentially. The relatively high pages-per-session and session duration

<table>
<thead>
<tr>
<th>TABLE 6.7</th>
<th>Advanced Google Analytics (by job role sorted by completion rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Unique Users</strong></td>
</tr>
<tr>
<td>will complete</td>
<td>4,379</td>
</tr>
<tr>
<td>will complete</td>
<td>10,982</td>
</tr>
<tr>
<td>will complete</td>
<td>16,101</td>
</tr>
<tr>
<td>will complete</td>
<td>4,751</td>
</tr>
<tr>
<td>will complete</td>
<td>37,075</td>
</tr>
<tr>
<td>will complete</td>
<td>3,554</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 6.8</th>
<th>Google Analytics for Beginners (by developer job role sorted by completion rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Unique Users</strong></td>
</tr>
<tr>
<td>will complete</td>
<td>37,075</td>
</tr>
<tr>
<td>will complete</td>
<td>3,554</td>
</tr>
<tr>
<td>will complete</td>
<td>2,349</td>
</tr>
<tr>
<td>will complete</td>
<td>7,318</td>
</tr>
<tr>
<td>will complete</td>
<td>5,465</td>
</tr>
</tbody>
</table>
for general users show that users are not simply disengaging, but they are jumping around the course to find what they need, possibly prospecting for relevant information, a behavior observed in other MOOCs (Guo & Reinicke, 2014; DeBoer, Stump, & Breslow, 2014; Seaton et al., 2014).

If indeed users are prospecting the course, this is corroborated by post-course survey feedback that requests more relevant and relatable content:

- “Some of this content only applies to retail and e-commerce professionals, which is tough to relate to for us content publishers.”
- “I would have liked to see more information for content producer sites or lead generation sites.”

Other users asked for content more tailored to their specific business’s needs and roles:

- “There are many others (like Business Manager, Product Managers, etc.) who would like to know the overview and power of GA so that they can leverage it, but they might not want to know each part of the GA… . A course tailored with those objectives … will also save time and fulfill the curiosity/improve the understanding of such groups.”

These users are (perhaps unwittingly) asking for a course engineered for a personalized learning experience. The lack of personal relevance across business types and job roles could be a primary reason why some intended completers are not as motivated as expected, particularly in the online support and online communities segments for GAB and the developers segment for GAB and AGA.

### TABLE 6.9 Advanced Google Analytics (by developer job role sorted by completion rate)

<table>
<thead>
<tr>
<th>will complete (all)</th>
<th>Unique Users</th>
<th>Completion Rate</th>
<th>Sessions/User</th>
<th>Pages/Session</th>
<th>Session Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37,075</td>
<td>30.64%</td>
<td>2.86</td>
<td>7.01</td>
<td>26:31</td>
</tr>
<tr>
<td>will complete</td>
<td>3,554</td>
<td>29.12%</td>
<td>2.89</td>
<td>6.40</td>
<td>23:04</td>
</tr>
<tr>
<td>(Developer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>will complete</td>
<td>1,244</td>
<td>28.86%</td>
<td>3.06</td>
<td>6.15</td>
<td>22:34</td>
</tr>
<tr>
<td>(Developer, Analyst, Marketer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>will complete</td>
<td>1,620</td>
<td>28.33%</td>
<td>3.04</td>
<td>6.14</td>
<td>21:56</td>
</tr>
<tr>
<td>(Developer, Marketer)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>will complete</td>
<td>551</td>
<td>28.31%</td>
<td>3.15</td>
<td>6.03</td>
<td>21:02</td>
</tr>
<tr>
<td>(Developer, Analyst, Marketer, Agency)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE 6.9 GAB Unit 1 path (intended completers who selected ecommerce business type)
FIGURE 6.10 GAB Unit 1 path (intended completers who selected online support business type)
FIGURE 6.11 GAB Unit 1 path (intended completers who selected online community business type)
FIGURE 6.12 GAB Unit 1 path (intended completers who selected agency job role)
FIGURE 6.13 GAB Unit 1 path (intended completers who selected developer job role)
The Promise of Personalization

With the arrival of scalable learning found in open learning environments and MOOCs, studies have attempted to better understand how user behavior differs in mass enrollment courses and to account for lower-than-expected completion rates. Some have suggested that, due to the large volume of participants, such courses are not able to effectively evaluate learning goals by traditional standards of educational expectation and measurement. Kizilcec and Schneider (2015) suggest that students taking online courses may not browse linearly through a course, situating this practice as a phenomenon of web culture, where users may casually view different learning content that suits them, similar to browsing websites or social media. Clearly, the large volume of differentiated users in course programs like the Academy make it difficult to serve all its students with personally relevant course curriculum. The open availability of the course to virtually anyone, coupled with the diversity of user goals or experience levels, creates multiple opportunities for disengagement. It is no wonder that users click one or two lessons in sequence before bouncing to a different lesson. The generic curriculum is intentionally written to appeal to the majority of users, which puts the burden on students outside of that majority to search through the content to find what they need.

The diversity of the Academy learning audience is one of the most obvious challenges for creating relevant course content. Internal user research studies at Google have demonstrated that, for beginners, getting started with GA can be daunting. Non-technical users struggle to implement JavaScript tracking code in order to collect basic analytics data. They are also overwhelmed with the amount of GA features and reports available and struggle to understand which metrics are most relevant to their business. Features and reports also evolve constantly, mandating some degree of continuing education to maximize use of the product. Because many users are self-taught, they may not be aware of particular features or reports, and need guidance on what they should use or best practices. Users already familiar with GA may not understand conceptual fundamentals like data collection and processing, which can affect decisions about implementation and therefore data quality.

Users come to open-course environments with different motivations and learning needs (Kizilcec & Schneider, 2015). For some, the objective is to earn a certificate to prove to themselves (or an employer) that they have mastered a particular domain of knowledge. Others may be motivated by the curiosity to fill knowledge gaps, to supplement what they know with additional information, or to get a better understanding of a particular subject matter. MOOCs serve a useful purpose for these types of browsers without requiring course completion. But users who have indicated they wish to complete the course should not be obstructed by learning content they do not find applicable. To make course content relevant to different types of users in an open-learning environment
requires an understanding and adaptation to each user’s specific interests and knowledge goals.

Personalization offers a way to increase engagement across all segments of users, despite different motivations or experience levels. As Chris Dede notes in this volume, there is substantial evidence that personalized course content can increase motivation and provide significant educational outcomes for differentiated users. For users at different experience levels, personalized course content can be used to counter the “expertise reversal effect” where more advanced users may actually perform worse if taught in the same way as novice users (Kalyuga, Ares, Chandler, & Sweller, 2003). By personalizing course content based on a student’s experience level, the Academy can optimize their instruction for different levels of product familiarity. Personalization can also address different business types and job roles that may use GA in different ways. Developers who deal with implementation and configuration settings will likely not need lessons on how to use GA to build marketing audiences or perform campaign analysis. Publishers who use platforms like AdSense to serve advertising on their website generally will not need to learn marketing platforms like Google Ads. And, while small-to-medium size businesses might relate to the example measurement plan for the Google Merchandise Store, agencies require a completely different example with agency-specific business goals, tactics, and key performance indicators. By personalizing course content, large, open-learning environments can appeal to different segments of their learning audience with content that is more relevant and directly applicable.

To better serve the diversity of user motivations, Kizilcec and Schneider (2015) have suggested rethinking MOOCs as an archive with content existing within a larger content ecosystem. They suggest breaking apart MOOCs into separate modules that could be tagged to allow reference-style usage. This could allow MOOCs to adapt to user needs, potentially even offering course content in different sequences for different users. Wilkowski, Deutsch, and Russell (2014) have suggested letting students create their own custom courses out of the content that interests them, while Daniel, Vázquez Cano, and Gisbert Cervera (2015) suggest using pre-defined indicators or pre-requisites to adjust course content based on user interests or educational backgrounds. These all provide valuable guidance in creating a model of personalization for large, open learning environments.

Proposed Method

There are different ways of personalizing content: implicit, explicit, and adaptive. Implicit personalization updates course content based on a user’s behavior, such as what is clicked on or a navigation path. Explicit personalization is determined by signals that the user actively communicates as part of a user profile, survey, or similar diagnostic input. Adaptive personalization predicts specific immediate responses before and during user interactions (this is often used in exams or
quizzes that require content to adapt to the skill levels of learners). Each type of personalization has benefits and limitations. Implicit personalization requires little to no effort from the user, simply that they engage with a learning application so it may intuit user signals based on behavior. However, the quality of data signals for personalization is limited by what can be accurately intuited by user behavior and can be affected by false negatives or false positives (Lu, 2004). Explicit personalization requires proactiveness on the part of the user, which may reduce the number of users willing to volunteer their preferences (and is only as good as users’ self-reported data). Adaptive personalization can potentially be more complex and requires sophisticated algorithms that constantly evaluate user inputs and provide real-time responses. Every open learning environment will need to assess methodologies of personalization against their stated goals and limits of their platform.

Implicit personalization presents a number of challenges for GA. Google Analytics accounts consist of parent properties with different child views of data underneath each property. Some configuration settings for features are applied at the property level (which impacts all the views in that property), while some features are set up at the view level. Settings configured at the property level are not associated with individual users, while settings at the view level are. For instance, if a user linked Google Ads to GA, that would be associated at the property level and would impact all users that belonged to the property. It would be impossible to use that as a data signal to personalize course content for individual users. However, if a user was to create a segment or a custom report, that only applies to that user and therefore could be used as a signal for individual personalization.

Similarly, users can have administrative rights at the account, property, or view level. Just because users may not have administrative rights on the account used to view Academy lessons does not mean they do not have administrative permissions on a different account and need to learn about administrative settings. Adaptive personalization presents similar problems as implicit personalization for using data signals to adapt content, and the engineering burden required to build truly adaptive content is high. Additionally, privacy protections within Google limit the availability of behavior data needed to perform implicit or adaptive tracking.

Although explicit personalization places an added burden on the user to communicate preferences to the learning environment, it can be a very direct way of connecting relevant course content with users’ stated preferences. In an explicit model, Academy users could complete a series of short questions to provide information about their specific learning goals. Users could indicate whether they were interested in online marketing and the course could include lessons on Google Ads based on whether the user checked yes or no. Users could select whether they had administrative rights (or wanted to learn about administrative settings) and the course would show the relevant lessons based on that input.
There are multiple opportunities for personalization in MOOCs including learning paths through a program and targeted course content within lessons.

**Personalized Learning Paths**

To personalize which lessons a user sees, the Academy envisions an optional user diagnostic that exists as part of the Academy user profile. This diagnostic should be optional so as not to disengage learners who find this added survey burdensome. The diagnostic would ask a series of questions that communicates content preferences and creates a model of the learner based on specific user attributes (Pea & Jacks, 2014). Some examples of data signals to target Academy learning content include:

- **GA experience level:** [just starting out] [somewhat familiar] [very familiar]
- **GA360 customer or interest (paid/premium product):** [yes] [no]
- **Account administrator rights:** [yes] [no]
- **Job Role:** [Analyst] [Marketer] [Developer] [Agency] [Student] [Other]
- **Website Type (Business type):** [Publisher] [Ecommerce] [Lead Generation] [Branding] [Online Support] [Online Community] [Other]
- **Goal to advertise online:** [yes] [no]
- **(Other common business goals)**

Different Academy lessons would map back to the response criteria. In cases like **GA experience level**, the learning path would suggest course lessons based on skill level for users to begin their training. This could be considered core curriculum shared by every user at that experience level. All assessments would map back to this core curriculum in order to provide a shared means of testing personalized lesson sequences in parity. Adding **GA360 interest** would show additional course content related to GA’s enterprise-level paid product. Administrator rights show additive content of administrative settings and other features that require administrator access. **Website type** and **job role** responses show or remove lessons mapped back to the relevance of those particular signals. **Goal to advertise online** would show additive lessons on marketing tools and reports. These signals could be supplemented with further-defined business goals that can be mapped back to lessons teaching supporting features.

Note that some user needs may not fit neatly into the prescribed diagnostic response options. Categorical user input selections can be notoriously fickle in communicating the nuance of user needs. As in the current application, job role and website type will be multiple-answer selections to indicate overlaps across broader businesses and tasks. If more than one selection is checked (for example, **marketer** and **analyst**) that creates a conflict between the mappings of user inputs and suggested lesson outputs, the system will default to overwrite any muted suggestions. Thus, if both **analyst** and **marketer** are selected, the analyst role will...
include lessons that the marketer role would have otherwise muted, thereby overriding lesson content to show. Similarly, if users find their job role and website type do not fit into any of the feedback selections, they can check other, which will exempt that particular signal from influencing the personalized mapping, thereby mitigating false positives or negatives. In order to manage assumptions about what constitutes answer options (such as different job roles or website types), the Academy can include information icons next to each answer with descriptions of those selections. This will allow users to more accurately align their choices with the Academy’s definitions of each answer type.

A particularly compelling learning path model is one that responds with real-time feedback to diagnostic response criteria. When users choose each diagnostic selection, the learning path offers a list of suggested lessons across available courses in real time. As users make more diagnostic selections, the system immediately indicates which lessons have been added or muted. This will be framed as a suggested learning path through the course. Users will have the option of manually customizing this path at any time by adding back in muted lessons or muting lessons that are not of interest (Figure 6.14).

The importance of showing instantaneous lessons mapped to diagnostic criteria should not be understated. This provides a powerful way of making explicit the relationship of user needs to a library of content knowledge. The ability of users
to customize their path provides an additional opportunity for self-reflection and awareness of users’ own learning needs and knowledge gaps.

All user inputs, as well as which lessons users add or mute to customize their learning paths, should be saved and associated with that user. For returning users (who may have already completed coursework and received certificates), the learning path could provide callouts to show which lessons have been updated since the user last logged in or new lessons that have been added. The customized learning path will also include a way for users to indicate when they have completed a lesson and to track progress. This offers another engagement signal for learning analytics data. Additional Custom Dimensions can be created from the responses to the diagnostic inputs to help analyze course behavior. These diagnostic mappings could be leveraged for personalizing course content within lessons as well.

**Personalized Lesson Content**

To personalize lesson content, a new data structure and model will need to be built that maps user input selections from the diagnostic to tagged blocks of HTML in lesson content using a content management system. Tagged blocks of course content will only be served to users with the corresponding mapping attributes associated in their user profile. If users were to change those input selections in the diagnostic, then the learning path and course content would automatically update to show any course content tagged with those mappings. For example, when showing the Google Store’s measurement plan, the Academy could show different versions of this plan based on whether the user indicated they were a publisher, an e-commerce business, or an agency. Similarly, when providing real-world examples of how a feature might be used, the Academy could offer specific examples to individuals based on the business type and job role indicated. This type of personalization is not limited to lesson content. Activity questions for different job roles would be much more applicable in terms of practicing particular tasks in GA, and even assessment questions could be customized for individual users (provided they were assessing the same underlying course concepts across all types of users).

**Success Metrics**

If personalization creates custom learning paths for users and makes lesson content more relevant, then this should be evident in the learning analytics data collected. Some studies have suggested that MOOCs and similar large-scale learning environments should not be bound by the same models of assessment and success as traditional courses due to the sheer diversity of participants, particularly in knowledge backgrounds and intention (DeBoer, Stump, & Breslow, 2014). However, this is based on the idea of a generic curriculum trying to serve
students with vastly different backgrounds and intentions. If the course content better addresses the diversity of participants in large-format online courses through personalization, students may navigate MOOC courses in a more traditional fashion, since the content has been streamlined for relevance.

In the Academy model, analytics should be collected on the diagnostic itself, tracking engagement and the customizations that users make to their learning paths. If personalization truly makes courses and lessons more relevant, there is an expectation that engagement and completion rates would rise across all segments, particularly those underserved segments that show evidence of disengagement. However, it is worth noting that some of the traditional metrics by which engagement has been measured may change. For instance, users who engage with the diagnostic may actually have decreased session duration and pages/session metrics if they find what they need faster and spend less time searching for content.

**Risks/Challenges**

While the prospect of personalizing learning paths and course content for large, open learning environments offers obvious benefits for users, it is not without risks. It is unknown whether the potential benefit of customized learning paths and course content will justify the engineering effort to build out the user diagnostic and the data structure for personalized learning. Users also may not be sufficiently motivated to complete the diagnostic, particularly if they feel the process is too time consuming or burdensome. The diagnostic categories may not be specific enough to capture the subtle differences that characterize different variants of users. Are there user differences that may not be captured in the semantic design of the diagnostic? User studies should be run early in the process to ensure that diagnostic categories are applicable to the diversity of the learning audience.

There are additional considerations involved in personalizing course content. Creating high-quality interactive courses designed to engage learners is time-consuming and involves significant effort. The personalization of course content requires instructional designers to develop additional content for each of the user audiences it wishes to target. This has the potential to extend the course development process and make it more difficult to update and localize content. Course developers will need to be judicious in deciding which content is worth personalizing. This may depend on additional factors like course format and the ease with which that content may be edited and updated. Also, designers will have to write content in a manner that fosters seamless transitions between the core curriculum shared by everyone and extra lessons included for select individuals. Regardless of how a personalized learning path is ordered or what lesson content is shown, personalized content needs to make contextual sense with the content that comes before and after. Do the potential engagement gains justify the
development effort? How should return on investment be measured compared with the extra instructional- and platform-development time needed for such an endeavor?

Another consideration for personalized course content is assessments and certificates. While a common assessment can test on a core curriculum, should assessments also include questions to assess the mastery of personalized content if served to select individuals? And if users complete different assessments based on their diagnostic inputs, does the course certificate need to reflect the specialized subject matter completed? These are important questions that should be considered and which could potentially impact the student experience and amount of engineering work required.

**Future Investigations**

Personalized learning paths mapped from user diagnostic inputs are only as good or predictive as the lesson mappings created by their designers. These mappings have no assurance that they are indeed the best lesson suggestions for each combination of user inputs; this is why it is so important to let users manually customize learning paths. Machine learning (ML) can ensure these mappings are as accurately predictive as possible. Machine learning comes out of the field of artificial intelligence, but rather than striving to make machines intelligent, ML instead seeks to make algorithms learn from experience. Instead of instructing a machine what to do, ML models extract patterns out of vast amounts of data and applies those patterns to future situations. This has the potential to provide predictive user experiences based on the many users who demonstrated patterns of previous behavior. Google has used machine learning to advance fields such as speech recognition, language translation, and image classification.

Machine learning works well for defined tasks in which relevant variables can be communicated to the learning algorithm, such as predicting a user’s ideal learning path based on previous user inputs. When users communicate particular diagnostic criteria and the Academy suggests a learning path, machine learning can track whether users add or mute lessons in their path based on their combinations of other inputs like *business type* and *job role*. Over time, ML will begin to associate the various combinations of inputs with those particular business types and job roles. As more users engage with the diagnostic, users who input the same combinations of diagnostic inputs will either validate or adapt that path based on whether they add or mute lessons.

Longitudinally, ML algorithms will come to understand which lessons are added or muted and extrapolate the most common learning path for each combination of diagnostic inputs. If a learning path is not customized by a user, the algorithm will assume the suggested learning path was what the user wanted and use that as a data point to validate that particular path with the user inputs given. As an example, let us say hundreds of thousands of users indicated they were
marketers with administrative access who were measuring a lead generation website. While the Academy would suggest muting lessons related to the *enhanced e-commerce* feature in GA, perhaps marketers for lead-generation businesses actually *are* interested in using this feature. If enough of them unmute this feature, adding it back to their learning path, over time the Academy’s ML algorithm would associate *enhanced e-commerce* lessons to that diagnostic criteria.

Machine learning could effectively overwrite the original Academy lesson mappings with better suggestions. This allows users to communicate what they need and to provide predictive guidance for similar users in the future. If successful, course developers would spend less time manually customizing learning paths as the algorithm increasingly predicts correct paths for different types of users. This could even inform instructional designers of what different segments of users wish to learn.

**Conclusion**

The convergence of technologies such as cloud computing (big data), learning analytics, and ML is affording online learning programs new opportunities to personalize education at scale (Siemens, Dawson, & Lynch, 2013). A diagnostic mapping that suggests learning paths with tailored lessons and a means of refining those suggestions over time using ML has the potential to better engage different segments of learners in large, open learning environments who intend to complete a course. To be successful, the next generation of MOOC-type courses will need to create personalized learning experiences to better serve their differentiated learning audiences. Learners growing up in an Internet-connected, digital world will increasingly expect online courses to intuit their individual needs and serve them content that makes their learning experience more personally relevant and meaningful.

**Notes**

1 Segmentation in Google Analytics is a way to view a subset of data in a report and may be used to compare different segments together.
2 The Google Merchandise Store sells Google-branded promotional items in a store on its Mountain View campus as well as online.
3 A session begins when a user navigates to a page that includes the GA tracking code and generates a pageview hit. It will end after 30 minutes if no other hits are recorded. If a user returns to a page after a session ends, a new session will begin.
4 A Custom Dimension is a way to organize data based on additional user information collected that is not available by default in GA.
5 The registration form and each lesson and assessment represents a “page” in the Academy data.
6 Events are generally user interactions in an interface that can be collected in GA.
7 Google Analytics for Beginners includes two clickable demonstration videos built in Adobe Captivate that walk users through creating a GA account and applying filters to views.
Based on a Goal completion in GA of passing all assessments for a given course.

Users generally set up separate properties in GA to collect data for each website, mobile application, or device (Google Support).

Users set up views with filters to derive different looks at their data within a property such as a view that only includes Google Ads traffic under a property that collects website data.

This concept has been inspired by the Intelligent Wizard developed by Google’s Engineering Education group for Google Computer Science Education (Google Computer Science Education: Explore Resources). This wizard lets users interested in computer science enter multiple criteria and view matching education results in real-time.

Acknowledgements

The author wishes to thank Justin Cutroni, Adam Feldman, Helen Huang, Andrei Lazarchuk, Michael Lenaghan, Christina Macholan, Paul Muret, Lizzie Pace, Katie Richardson, and Kevin Trotter.

References


Motivations: Learning Assistance in Online Education

Massively Open Online Courses (MOOCs) are rapidly proliferating. According to Class Central (Shah, 2016), in 2016 more than 58,000,000 students across the world registered for more than 6,800 MOOCs offered by more than 700 institutions. Further, these numbers continue to grow rapidly. Today, MOOCs cover almost all disciplines and education levels, and their students cut across most demographics groups such as gender, age, class, race, religion, nationality, and so forth.

However, the effectiveness of learning in many MOOCs is questionable as the student retention ratio typically is less than 50% and often less than 10% (Yaun & Powell, 2013). Although there are several reasons for the low student retention, a primary reason is the lack of interactivity in MOOCs (Daniel, 2012). Thus, one of the principle recommendations for improving the effectiveness of learning in MOOCs, and thereby also improving student retention, is to enhance the interaction between the teacher and the students (Hollands & Tirthali, 2014).

As an example, consider Georgia Tech’s recently launched online section of CS 1301: Introduction to Computing based on the Python programming language. This online section is in addition to traditional, residential sections of the Introduction to Computing class. The online class itself has two sections. In Spring 2017, the accredited section was available only to 45 selected Georgia Tech students who had access to three teaching assistants (TAs) in addition to course materials provided by the instructor. The three TAs provided several kinds of support to the online students, such as answering questions, tutoring on the course materials, evaluating student progress, and so forth. The open and non-credited section of the online Introduction to Computing class—the MOOC—currently has more than 40,000 registered students. The students in the MOOC have access to all the same course
materials as the students in the other online section. However, the 40,000 MOOC students do not have access to any TA (or the instructor, except indirectly through the standard course materials). Given that computer programming is a technical skill that many students find difficult to master on their own, it is unclear what percentage of students in the MOOC section will successfully complete the course. It seems safe to say the percentage of students who successfully complete the MOOC section without any teaching assistance will be significantly lower than the students in the online section with teaching assistants.

Of course, most humans are capable of learning some knowledge and some skills by themselves. However, reliable estimates of autodidacts with the capacity to learn advanced knowledge and complex skills are not readily available. For the purposes of the present discussion, let us posit that the vast majority of learners can benefit from learning assistance: perhaps more than 90% of the 58 million students taking a MOOC worldwide may need or want some learning assistance, and perhaps as many as 99% may significantly benefit from learning assistance. If we assume just one teaching assistant (TA) for 50 students for a typical MOOC, then we need at least 1 million TAs for supporting the 58 million students registered for a MOOC! It is highly doubtful that anyone can organize or afford such a large army of human TAs. The Georgia Tech CS 1301 MOOC itself would need about 800 TAs to support the 40,000 students, more than the number of TAs in all other Georgia Tech classes in computing combined. This raises a profound problem: how can we provide meaningful learning assistance to the tens of millions of learners taking MOOCs?

In response to this question, MOOC teachers, researchers, and service providers are engineering online learning by building on several technologies for learning assistance, such as E-Learning (e.g., Clark & Mayer, 2003), interactive videos (e.g., Kay, 2012; Koumi, 2006), intelligent books (e.g., Chaudhri et al., 2013), intelligent tutoring systems (e.g., Azevedo & Aleven, 2013; Polson & Richardson, 2013; Van Lehn, 2011), peer-to-peer review (e.g., Falchikov & Goldfinch, 2000; Kulkarni, Berstein, & Klemmer, 2015), and autograding. Of course, many of these technologies were developed prior to the start of the modern MOOC movement with Stanford University’s MOOC on artificial intelligence in 2011 (Leckart, 2012; Raith, 2011). Nevertheless, MOOCs are extensively developing and deploying these technologies to assist online education.

Another strategy for engineering online learning is to design and develop virtual teaching assistants to augment and amplify interaction with human teachers. These virtual teaching assistants may help with many of the tasks human teaching assistants do, for example, cognitive tutoring, question answering, question asking, autograding, formative assessment, and metacognitive tutoring.

In this chapter, we describe a virtual teaching assistant called Jill Watson for the Georgia Tech OMSCS 7637 class on knowledge-based artificial intelligence (KBAI). Jill Watson (JW) has been operating on the online discussion forums of different offerings of the KBAI class since Spring 2016. At the time of writing this chapter in
June 2017, some 750 students and some 25 (human) TAs had interacted with different versions of JW. In the Spring 2017 offering of the KBAI class, JW autonomously responded to student introductions, posted weekly announcements, and answered routine, frequently asked questions. Thus, JW is a partially automated, partially interactive technology for providing online assistance for learning at scale. In this discussion of JW, we describe the motivation, background, and evolution of the virtual-question-answering teaching assistant, focusing on what JW does rather than how she does it.

**Background: An Online Course on Artificial Intelligence**

In January 2014, Georgia Tech launched its online Masters of Science in Computer Science program (OMSCS). OMSCS is a fully accredited, highly selective Georgia Tech graduate degree offered to select students from across the world. The online courses are developed by Georgia Tech faculty in cooperation with Udacity staff, offered through the Udacity platform, and supported by a grant from AT&T. The goal of the OMSCS program is to offer the same courses and programs online that are offered through the on-campus Masters program while maintaining equivalent depth and rigor (Joyner, Goel, & Isbell, 2016). In Spring 2017, the OMSCS program enrolled an order of magnitude more students (approximately 4,500) than the equivalent residential program (approximately 350) that cost far less (approximately $7,000) than the residential program (approximately $30,000) (Carey, 2016; Goodman, Melkers, & Pallais, 2016). By now a few hundred students have successfully completed the OMSCS program, and the diploma awarded to them does not mention the word online.

As part of the OMSCS program, in 2014, we developed a new online course called CS7637: Knowledge-Based Artificial Intelligence: Cognitive Systems (KBAI). The first author of this article (Goel) had been teaching an earlier face-to-face KBAI course on the Georgia Tech campus for more than a decade. While the online KBAI course builds on the contents of the earlier on-campus KBAI course, we rethought the course for the new medium and developed many of the course materials from scratch (Goel & Joyner, 2016). The second author (Polepeddi) took the online KBAI course in Summer 2015 and was a TA for the course in Spring 2016.

The online, semester-long KBAI course consists of 26 video lessons developed from scratch that help teach the course material (Ou, Goel, Joyner, & Haynes, 2016), a digital forum (Piazza) where students ask questions and participate in discussions as illustrated in Figure 7.1, a learning management system through which students submit assignments and receive grades (Sakai), a proprietary peer-feedback tool developed at Georgia Tech where students read and submit feedback on each other’s assignments, and a proprietary autograder tool developed by Udacity that helps grade the source code of programming projects. The course is administered by the instructor (typically Goel), who is assisted by a small team of TAs. The TAs generally answer questions and facilitate discussions on the digital forum, and they grade assignments, projects, and examinations.
FIGURE 7.1 While the video lessons in the OMSCS KBAI course are like a textbook, the class forum is like a virtual classroom where students ask questions, discuss ideas, and give feedback. Here, a student asks a question about whether there is a word limit on an assignment.

Length of first assignment
Okay so when first looking at this assignment I said to myself how on earth am I going to come anywhere close to 1000 words. Now that I am deep into it I am going WAY over 1000 words. I feel that by completion I might be close to 2000 words. Is this going to be an issue and yes I was expecting to have the opposite problem.

Thanks,
Since Fall 2014, we have offered the OMSCS KBAI course each fall, spring, and summer term. Enrollment in the class has ranged from about 200 to 400 students each term, so that at the time of writing, about 2,000 online students have enrolled in the course. For the most part, student surveys of the online KBAI course have been very positive (Goel & Joyner, 2016; Ou, Goel, Joyner, & Haynes, 2016). In addition, in the fall terms of 2014, 2015, and 2016, we have offered the same KBAI course to residential students at both graduate and undergraduate levels. The performance of the online students on the same set of assessments using blind grading has been comparable to that of the residential students (Goel & Joyner, 2016; 2017). The retention ratio in the online section has been 75–80%, only slightly lower than the 80–85% in the residential sections.

The OMSCS KBAI course has provided us with a research laboratory for conducting experiments in pedagogy for online education. For example, we have experimented with programming projects based on real artificial intelligence research to promote authentic scientific practices (Goel, Kunda, Joyner, & Vattam, 2013) as well as the use of peers as reviewers and TAs as meta-reviewers (Joyner et al., 2016). We also developed and deployed about a hundred nanotutors for teaching domain concepts and methods (Goel & Joyner, 2017). A nanotutor is a small, focused AI agent that models students’ reasoning on a particular problem engaging a domain concept or method to be learned. Given a student’s answer to the problem, a nanotutor first classifies the answer as correct or incorrect and then provides an explanation on why.

A Challenge in Scaling Online Education: Responding to Student Questions

Teaching the OMSCS KBAI class in the Fall 2014 and Spring 2015 terms revealed a new challenge for the teaching staff: the discussion forum for the online class was very active and thus took a large amount of staff time to monitor and respond. Table 7.1 provides the data from the discussion forums for the online and residential sections from Fall 2016. As Table 7.1 indicates, the discussion forum for the online section had more than 12,000 contributions compared to fewer than 2,000 for the residential class. One obvious reason for this six-fold increase is that the online class had three times as many students as the residential class. Another, perhaps less obvious reason is that the discussion forum acts as the virtual classroom for the online class (Joyner, Goel, & Isbell, 2016). It is on the discussion forum that the online students ask questions, get and give answers, discuss the course materials, learn from one another, and construct new knowledge.

While the abundant participation on the discussion forum of the online class likely is an indication of student motivation, engagement, and learning (and thus is very welcome), the higher levels of participation create a challenge for the teaching
staff in providing timely, individualized, and high quality feedback. On one hand, the quality and timeliness of TAs’ responses to students’ questions and discussions are an important element of providing learning assistance and thus play a part in the success of student learning and performance. On the other hand, given the high rate of student participation on the discussion forum, the TAs may not have time to respond to each message with a high quality answer in a timely manner.

A Potential Answer: Virtual Teaching Assistants

In reading through the students’ questions on the online discussion forums of the OMSCS KBAI class in Fall 2014 and Spring 2015, we recognized (as many teachers have done in past), that students often ask the same questions from one term to another and sometimes even from one week to another within a term. For example, questions about length and formatting of the assignments, allowed software libraries for the class projects, and class policies on sharing and collaborating have been asked in different ways every semester since January 2014. Perhaps more important than that, from the online discussion forums of the Fall 2014 and Spring 2015 OMSCS KBAI classes, we had access to a dataset of questions students had generated and the answers TAs had given. Thus, in Summer 2015, we wondered if we could construct a virtual teaching assistant that could use the available dataset to automatically answer routine, frequently asked questions on the online discussion forum. We posited that if we could create a virtual TA that could answer even a small subset of students’ questions, then it would free the human TAs to give more timely, more individualized, and higher quality feedback to other questions. Also, the human TAs may have more time to engage in deeper discussions with the students.

Our thinking about the virtual teaching assistant was also inspired by IBM’s Watson system (Ferruci, 2012; Ferruci et al., 2010). Independent of the OMSCS KBAI class in Fall 2014, IBM had given us access to its Watson Engagement Manager\textsuperscript{4} for potential use in support of teaching and learning. We successfully used the Watson Engagement Manager for teaching and learning about computational creativity in a residential class in Spring 2015 (Goel et al., 2016). Building

\begin{table}
\centering
\caption{The level of participation of online students in the OMSCS KBAI class on the digital forum is much higher than that of residential students. Table 7.1 compares four participation metrics between online students and on campus students during the Fall 2016 offering of KBAI class.}
\begin{tabular}{llll}
\hline
                      & Residential (Fall 2016) & Online (Fall 2016) & \% Change \\
\hline
Number of students  & 117                    & 356               & +3x \\
Total threads       & 455                    & 1201              & +2x \\
Total contributions & 1,838                  & 12,190            & +6x \\
\hline
\end{tabular}
\end{table}
on this educational experience with the Watson Engagement Manager, in Fall 2015, IBM gave us access to its newer Bluemix\textsuperscript{5} toolkit in the cloud. Thus, we were familiar with both the paradigm of question answering and some of the Watson tools.

**Jill Watson and Family**

At the time of writing (June 2017), we have developed three generations of virtual teaching assistants. We have deployed these virtual teaching assistants in the discussion forums of the online KBAI classes in Spring 2016, Fall 2016, and Spring 2017, as well as in the residential class in Fall 2016. All actual experiments with the virtual teaching assistants have been in compliance with an institutional review board (IRB) protocol to safeguard students’ rights and to follow professional and ethical norms and standards.

We call our family of virtual teaching assistants Jill Watson because we developed the first virtual teaching assistant using IBM’s Watson application programming interfaces (APIs). However, the names and tasks of specific virtual teaching assistants have evolved from generation to generation as described below. More important, starting with the second generation, we have used our own proprietary software and open-source libraries available in the public domain instead of IBM’s Watson APIs (or any other external tool). We made this shift to cover a larger set of questions as well as a larger set of tasks.

**Jill Watson 1.0**

**Design**

In January 2016, we deployed the first version of Jill Watson, Jill Watson 1.0 (JW1), to the Spring 2016 offering of the OMSCS KBAI class. Although we included JW1 in the listing of the teaching staff, initially we did not inform the online students that JW1 was an AI agent. As noted above, we built JW1 using IBM’s Watson APIs. JW1 is essentially a memory of question-answer pairs from previous semesters organized into categories of questions. Given a new question, JW1 classifies the question into a category, retrieves an associated answer, and returns the answer if the classification confidence value is greater than 97%.

Initially, we deployed JW1 on the discussion forum with a human in the loop; if JW1 was able to answer a newly-asked question, then we would manually check that her answer was correct before letting her post that answer to the class forum in reply to the question. In March 2016, we removed the human in the loop and let JW1 post answers autonomously.

Every 15 minutes between 9 a.m. and 11 p.m., JW1 checked the discussion forum for newly-asked student questions. We chose this time interval to mimic the working hours for most human TAs as well as to monitor to JW1’s
performance throughout the day. If there was a question that JW1 could answer and that another human TA had not already answered, she would post an answer.

**Performance**

Figures 7.2 and 7.3 illustrate some of JW1’s interactions with the online students on the discussion forum of the OMSCS KBAI class in Spring 2015. (Note that we have blackened some portions of the exchanges to maintain student confidentiality.)

We found that while JW1 answered only a small percentage of questions, the answers she gave were almost always correct or almost correct. We wanted to both increase the range of questions covered by JW1 as well as the task she addresses. The latter goal led us to develop the next generation of Jill Watson.

**Jill Watson 2.0**

**Design**

In the first week of the KBAI class, we ask students to introduce themselves on the discussion forum by posting a message with their name, their location, why they are taking KBAI this semester, other OMS classes they have taken, activities outside of school, and one interesting fact about them. Human TAs then reply to each student, welcoming him/her to the class. However, it is time consuming to respond individually to 200–400 students within one week. Thus, we built the second generation of Jill Watson, Jill Watson 2.0 (JW2), to autonomously respond to student introductions.

Unlike JW1 that was built using IBM's Watson APIs, we developed the software for JW2 in our laboratory from scratch, using only open-source, external libraries available in the public domain. Further, unlike JW1 that used only an episodic memory of question-answer pairs from previous semesters, JW2 used semantic processing based on conceptual representations. Given a student's introduction, JW2 first mapped the introduction into relevant concepts and then used the concepts as an index to retrieve an appropriate precompiled response.

In August 2016, we deployed two separate, virtual TAs to the discussion forums of the Fall 2016 offerings of the KBAI class that included both an online section and a residential section. We redeployed JW1 to answer routine, frequently asked questions as a TA named Ian Braun and we deployed JW2 to respond to student introductions as a TA named Stacy Sisko.

Just like Ian Braun, every 15 minutes between 9 a.m. and 11 p.m., Stacy checked for newly posted student introductions. Just as with routine questions, if there was a student introduction that Stacy could reply to and that another TA had not already replied to, she would autonomously post a welcome message.

Once again, while we listed both Ian Braun and Stacy Sisko among the teaching staff, we did not inform the students that they were AI agents. To
FIGURE 7.2 In this question about a class project with a coding component, the student asks whether there is a limit to their program’s run time. Jill Watson 1.0 correctly answers that there is a soft 15-minute run time limit.
In this question about a class assignment involving a writing component, the student asks whether there is a maximum word limit. Jill Watson 1.0 correctly answers that there is no strict word limit. Another student then has a follow up question asking for elaboration, which a human TA handles. After this exchange, one student in the class speculates whether Jill Watson is human.
prevent students from identifying the human TAs among the teaching staff through Internet searches, all human TAs operated on the discussion forum under pseudonyms.

**Performance**

Stacy Sisko autonomously replied to more than 40% of student introductions. Figure 7.4 illustrates Stacy’s responses to student introductions.

Figure 7.5 illustrates Ian Braun’s interactions with students on the online discussion forum. We found that although Ian Braun was a redeployment of JW1, he performed better in the Fall 2016 KBAI class than JW1 did in the Spring 2016 class both in the coverage of routine, frequently asked questions and in the proportion of correct answers. This improvement likely was because by Fall 2016 we had a larger dataset of question-answer pairs since the class had been offered a few more times by then.

**Jill Watson 3.0**

**Design**

Given the success of Stacy Sisko in using semantic processing to reply to student introductions, we created a third generation of Jill Watson, Jill Watson 3.0 (JW3), that uses semantic processing for answering questions. Unlike JW1, JW3 does not use IBM’s Watson APIs. Instead JW3 relies solely on an episodic memory. Given a student’s question, JW3 first maps the question into relevant concepts and then uses the concepts as an index to retrieve an associated answer from the episodic memory of questions organized into categories.

In January 2017, we deployed two separate, virtual TAs to the Spring 2017 offering of the OMSCS KBAI class. We redeployed version JW2 (or Stacy Sisko) to respond to student introductions as a new virtual TA named Liz Duncan, and we deployed version JW3 to answer routine questions as a virtual TA named Cassidy Kimball. Once again, while we listed both Liz Duncan and Cassidy Kimball among the teaching staff, we did not inform the students that they were AI agents. To prevent students from identifying the human TAs among the teaching staff through Internet searches, all human TAs operated on the discussion forum under pseudonyms. We also increased the time interval during which Cassidy checked for newly-asked questions to 6 a.m. and 11:59 p.m., based on our observations of the activity on the discussion forum.

**Performance**

Liz Duncan replied to 60% of all student introductions, a performance superior to that of Stacy Sisko in the earlier generation. Figures 7.6 and 7.7 illustrate Liz’s interactions with the online students.
FIGURE 7.4 In this introduction, the student expresses interest in learning more about artificial intelligence. Stacy responds that she also shares a similar interest in AI.
Yes, we need to follow any specific naming convention for naming our assignment pdf file? For instance naming it gatsbyname.pdf. Assignment.pdf will be ok?

Hello,

Hi, do we need to follow any specific naming convention for naming our assignment pdf file? For instance, naming it [name].pdf? Assignment.pdf will be ok?

Thanks.

Ian

Follow-up discussions

Replied 8 months ago

On 27 May 2021, 13:06

In your document, feel free to put your name and leave out a GTD. T-Square will associate your submission with your GTD Account when you submit.
FIGURE 7.6 In this introduction, the student shares that they live in Atlanta. Liz responds by inviting the student to visit Georgia Tech in person if they are in the area.
FIGURE 7.7 In this introduction, the student shares that they are currently taking another OMS course called Computer Architecture in addition to KBAI. Liz incorrectly processes that the student took Computer Architecture in a previous semester, and responds and asks what they thought of the class, prompting the student to reiterate that they are currently taking the class.
We found that Cassidy Kimball performed much better than JW1 and Ian Braun. For example, of the questions that students asked about KBAI’s three class assignments, Cassidy autonomously answered 34%, and of all the answers Cassidy gave, 91% were correct. Figures 7.8 through 7.11 illustrate Cassidy’s interactions on the online discussion forum.

**Student Reaction**

In the KBAI classes in Spring 2016, Fall 2016, and Spring 2017, we shared the true identities of the virtual AI agents towards the end of the term. Student reactions to our use of virtual teaching assistants in online discussion forums have been uniformly and overwhelmingly positive. Figure 7.12 illustrates a small sample of student reactions from the KBAI class in Spring 2016 after the students learned about the true identity of Jill Watson towards the end of April.

**Discussion**

There are several questions about the virtual teaching assistants that we have not fully answered in this chapter. The first question is how does Jill Watson work? As we briefly indicated above, Jill Watson 1.0 uses an episodic memory of questions and their answers from previous episodes. We developed JW1 using the IBM Bluemix toolsuite. In the second generation of Jill Watson, Ian Braun was a redeployment of JW1 for answering questions. However, Stacy Sisko used semantic-information-processing-technology developed in our laboratory to reply to student introductions. In the third generation of Jill Watson, Cassidy Kimball too uses semantic-information-processing-technology developed in our laboratory for answering questions as does Liz Duncan for replying to student answers.

Second, is the Jill Watson technology transferrable to other classes with different student demographics and using different educational infrastructures? To answer this question, we are presently building a new version of Jill Watson for a new Georgia Tech class, CS 1301 Introduction to Computing MOOC, that at present has 40,000 students but no TA support whatsoever.

Third, is the Jill Watson technology effective in lowering the demands on the teaching staff? While it is too early to determine the answer to this question for the task of question answering, anecdotally there is some evidence to suggest that Jill Watson did reduce the load on the teaching staff for responding to student introductions and for posting messages to the class.

Fourth, is the Jill Watson technology effective in enhancing student performance and improving student retention? We are presently conducting studies and collecting data to answer this question about student engagement, learning, and performance; it is too early to have insights into the issue of student retention.

Fifth, what ethical issues arise in using Jill Watson as educational technology in an online classroom? As we mentioned above, we obtained IRB approval in
FIGURE 7.8 In this question about a class assignment involving a written component, the student asks whether there is a preferred format for citations. Cassidy correctly responds to part of the student’s question that the APA format is recommended. A human TA responds to the other part of the student’s question.
FIGURE 7.9 In this question about a class project involving a coding component, the student asks for more feedback after submitting their assignment to the automated grading system. Cassidy incorrectly answers this question as if the student was asking about which problem sets are graded. The student asks someone else to help, to which a human TA responds.
FIGURE 7.10 In this question about a class assignment involving a writing component, the student asks about whether there is a preferred format to name files. The student also inserts a sentence asking human TAs not to respond, possibly in an attempt to discover the identity of the virtual TA. Cassidy correctly responds to this question.
FIGURE 7.11 In this question about the class midterm involving a written component, the student asks about the level of detail they should include in their responses. Cassidy correctly replies to the question, but the student second-guesses her answer and asks another human TA for confirmation.
FIGURE 7.12 Students react to our class post at the end of KBAI Spring 2016 class announcing the true identity of Jill Watson.
advance of the Jill Watson experiments. Nevertheless, these experiments have raised several additional, ethical issues. For example, when is it appropriate to use AI agents without telling human subjects about them? Does the use of a feminine name for an AI agent implicitly promote gender stereotypes? Might the use of AI agents as virtual teaching assistants eventually result in reduced employment opportunities for human teachers? These are serious questions that require investigation.

Conclusion

We may view the Jill Watson experiments from several perspectives on learning engineering. First, we may view Jill Watson as an educational technology for supporting learning at scale. In fact, this was our primary, initial motivation for developing Jill Watson, and this is also how we motivated the discussion in this chapter. As indicated above, Jill Watson uses AI technology for supporting learning at scale by automatically answering a variety of routine, frequently asked questions and automatically replying to student introductions.

Second, we may view Jill Watson as an experiment in developing AI agents so that for highly focused technical domains, highly selected subject demographics, and a highly-targeted context of human-computer interaction, it is difficult for humans to distinguish between the responses of AI and human experts. We found that in order to improve coverage, the design of Jill Watson gradually moved from using an episodic memory of previous question-answer pairs to using semantic processing based on conceptual representations.

Third, we may view Jill Watson as an experiment in human-AI collaboration. The KBAI class has become a microsociety in which humans and AI agents collaborate extensively and intensively, living and working together for long durations of time.

Notes

1 For a full description of the class, see their website at http://www.cc.gatech.edu/academics/degree-programs/bachelors/online-cs1301.
2 For the selection of classes, see https://www.udacity.com/courses/georgia-tech-masters-in-cs.
3 For the course description, see https://www.omscs.gatech.edu/cs-7637-knowledge-based-artificial-intelligence-cognitive-systems.

Acknowledgements

We thank IBM for its generous support of this work both by providing us access to IBM Bluemix and through multiple IBM faculty awards. We are grateful to
Parul Awasthy for her contributions to the development of Jill Watson 1.0 in Spring 2016. We thank members of the Design & Intelligence Laboratory for many discussions on the Jill Watson project. We also thank the students and (human) teaching assistants of the KBAI course in Spring 2016, Fall 2016, and Spring 2017. This research has been supported by multiple Georgia Tech seed grants for research and education. We especially thank Georgia Tech’s Office of the Provost, College of Computing, C21U, and the OMCS program.

References


CREATING PERSONALIZED LEARNING USING AGGREGATED DATA FROM STUDENTS’ ONLINE CONVERSATIONAL INTERACTIONS IN CUSTOMIZED GROUPS

Scott M. Martin and Matthew L. Trang

Introduction

The conventional method to increase enrollment within traditional, residential, post-secondary academia requires expenses that undercut the additional revenue. For example, to add one additional section of a 30-seat course, institutions need to augment their teaching capacity. This requires overloading payment to current faculty, adding additional full-time faculty, or hiring a competent adjunct to teach the extra course. Also, the additional 30 students need additional student support services, such as advising and counseling, which require supplementary financial resources (Cini & Prineas, 2016).

To counter increasing expenses, higher-education institutions have utilized mass lecture halls for required, foundational courses, scaling these to hundreds of students at a time. This pedagogical innovation did reduce instructional costs but again increased expenses for student-support services, negatively affected student-learning achievement levels from limited student-teacher interactions, and reduced personalized learning options (Chen, 2000). To further cut expenses, institutions of higher education began increasing the number of mass lectures in first and second year undergraduate major courses, replacing full-time or senior part-time faculty with graduate students as instructors who were often unaware of best practices of classroom management for large groups of students (Archibald & Feldman, 2011).

In large lecture halls, an immense number of peers may intimidate most students, except perhaps the most gregarious, so they are less likely to speak up for fear of saying something unintelligent (Chen, 2000). In addition, a large number of students within a class minimizes participation, as it is difficult to find time to call on every student to speak (Geske, 1992). Many students in a classroom also
means a lack of faculty connection with each student. For example, when there are 100 or more students in a classroom, it is difficult for faculty to call students by name to encourage participation and students see it as impersonal if the instructor says “hey you” (Chen, 2000).

To educational institutions, the promise of transposing their academic offerings to an online platform seemed an economic godsend, having struggled the past 25 years with shrinking endowments, reduced state support for public colleges, and increasing loan and bond interest payments brought on by a competitive student-experience building frenzy (Mitchell, Leachman, & Masterson, 2016). However, legacy digital learning systems were designed to provide better physical classroom management of course content for the instructor, such as facilitating assignments, tests, and grades, and were never intended to be used to teach online-only courses. Chris Etesse, one of the founding members of Blackboard, related that “Blackboard was designed to replace the metal basket on the teacher’s desk, to help the teacher better manage course assignments, quizzes, and tests. It was never designed to be used to teach” (C. Etesse, personal communication, May 16, 2016).

As discussed in Chapter 3, “Tinkering Toward a Learning Utopia,” with the rise and general acceptance of online instruction as an alternative educational pathway to traditional on-site classroom instruction, the field of learning science has begun to study datasets based on the copious amount of information collectable in digital environments. Research has thus far produced unclear results in understanding the multifaceted causes behind student attrition, course completion metrics, or overall student learning in asynchronous online-learning environments (Tyler-Smith, 2006). However, one consistent insight in almost all study and survey results—as well as in responses from instructors, learners, and researchers—is that lack of real-time interactivity has negative effects on student learning and motivation (Angelino, Williams, & Natvig, 2007). One-way transmission of academic content is roughly comparable to the historical, broadcast-television education model; one can scale the course content to a higher degree, but it is still an isolated learning experience for the student (Abdous & Yen, 2010; Croft, Dalton, & Grant, 2010).

Research demonstrates that scaling course content either by increasing class sizes within a physical classroom or by using an online, live-streaming, one-way platform produces very similar results: lower mean academic performance output, lower satisfaction survey results, and higher attrition rates (Angelino, Williams, & Natvig, 2007). However, online, education-focused platforms that provide live-streaming of instructors and provide some form of communication channels for real-time learner interactivity may potentially enhance the learning experience for a greater number of the traditional student demographic, as well as possibly better serving historically underserved populations and the non-traditional student (Angelino, Williams, & Natvig, 2007). As discussed in Chapter 2, “The Role of the Learning Engineer,” data-mining and surveys are starting to shed light on understanding the online educational experience: the expanding demographic of
student users for online-education offerings, why certain segments of the population have higher attrition rates, and what types of remediation may work for whom and when.

**Modern Online Learning Systems**

Social media companies and streaming, education-focused platforms (Udemy, Khan Academy, Lynda) have offered pre-recorded and live streaming content (video on-demand) for over a decade but only recently began integrating chat channels, private messaging, and other peer-to-peer functions so users can communicate with each other and sometimes with a content creator (e.g., YouTube in 2016, Twitter in 2016, Facebook Live in 2015, and Snapchat in 2012). In 2011, Twitch, the game-viewing and streaming-media platform, launched chat features that allow viewers to comment publically within a game player’s channel (Ewalt, 2013). In education, 2012 was the year streaming academic-related content was initiated by organizations and companies such as Coursera, Udacity, Udemy, and edX, which gave rise to the Massive Open Online Course, or MOOC (Smith, 2012). Legacy online-education-platform companies like Blackboard and Canvas Network followed suit, also offering a live-streaming toolkit for educators.

Thus, on the heels of social media companies embedding live-streaming content into their platforms, combined with interactive chat and AV features, updates of education-focused, legacy-software platforms added similar interactive channels and features. Further, in recent years a new generation of online-learning platforms have been launched, potentially changing how millions of learners worldwide gain access to and engage in education and training opportunities. Two such platforms are described next to illustrate their features.

Udemy.com is a primary, streaming, online-learning platform—based in San Francisco and founded in 2009—that is largely focused on education for professional adults. Udemy offers on-demand, not live-streaming, courses in a variety of disciplines, such as marketing, design, health and fitness, and IT and software. Udemy houses a library of over 55,000 mostly non-credit, non-certificate courses taught by primarily part-time and freelance instructors from around the globe, with each course ranging from 30 minutes of content to days’ worth of lectures. Udemy is not an accredited institution; courses are generally viewed in order to learn or improve upon a specific skill rather than as a module for credit that can be used toward a degree (Kay, Reimann, Diebold, & Kummerfeld, 2013). Each Udemy course is created by an instructor, and class material can be viewed by a learner at any time after registration and enrollment. Instructors record each Udemy lecture individually, and then each lecture is archived in a course library for the learner to complete at their own pace. If learners have questions regarding the subject matter, they can post a query within the question-and-answer channel embedded in each lecture video.
Although Udemy courses do not have a live-streaming course coupled with an interactive communication chat channel, learner interaction within the question channel is still commonplace. Learners assist each other by answering each other’s questions before, during, or after the streaming course, when the instructor may be offline. As with many highly viewed, archived courses, extensive interaction occurs sometimes between hundreds of students in real-time, chatting about subject-matter course content and also interacting socially about non-subject-matter topics. Figure 8.1 exhibits a screen shot of the Udemy learner User Interface (UI), featuring a streaming, archived course and the question channel. The question channel embedded in each lecture video also gives the learners options to insert images or graphics, to help clarify a topic or question, or to hyperlink a website to provide further extended information to their peers.

LiveEdu.tv is one of the newest, education-and-training-focused, live-streaming platforms. Founded in San Francisco in 2016, LiveEdu offers a semi-curated collection of non-credit courses and modules in technical disciplines such as programming, game design, data analytics, augmented reality, and artificial intelligence that are taught by part-time and freelance instructors from around the globe. Each instructor creates a personalized channel that hosts their personal and professional information and their live-streamed content that learners can join during each live broadcast. Each broadcast is archived within each instructor’s project playlist so learners can view or revisit archived videos of each previously live-streamed course. Figure 8.2 exhibits a screen shot of the LiveEdu learner’s user interface (UI) with one grouped class of 12 learners.

Similar to Udemy, LiveEdu has no limit on the number of learners in each instructor channel, and learners communicate with each other and the instructor through a chat channel in the student UI. LiveEdu also borrows from social media platform conventions and allows learners to add additional content besides text—such as images, hyperlinks, and emojis—into the chat channel.

In addition, LiveEdu incorporates database-driven, natural-language bots, better known as chatbots, within the learner and instructor chat channels to stimulate conversation and provide various levels of assistance and support to the instructor. The LiveEdu chatbot structure uses a relational database model and SQL language to store and manage the keyword pattern matching of text entered. The chatbot is designed to respond to a learner’s or instructor’s query without taking into account the larger conversational context. LiveEdu’s chatbots also semi-randomly submit chat-channel prompts to help facilitate learner discussion and to assist an instructor with deeper student engagement.

**Issues with Chat Channels in Large Online Courses**

When using a single, mass-chat channel for large numbers of participants in social media platforms, chat confusion can occur and disrupt the learning process (Fuks, Pimentel, & de Lucena, 2006). Some learners may ask pertinent questions, but
It's been used in radio astronomy for over half a century and is now emerging as an optical technique.
FIGURE 8.2 Screen shot of LiveEdu student interface with instructor live-stream
others add semi-relevant commentary, and still others contribute drivel and irrelevant emojis. Add a misplaced chatbot prompt, and the chat-channel content can quickly deteriorate into interactional incoherence (Cornelius & Boos, 2003). Chat confusion within a synchronous chat channel generally occurs in courses of more than 30 students (assuming 50% of learners are observers and not contributing to the discussion at any one instance) (Dennen, 2007). Capping the number of students in each course to increase chat coherence would improve subject-matter discussions, but would also greatly hinder the ability to scale the course.

One strategy to resolve chat-channel incoherence is the development of mediated, text-chat systems. Several methods have been used to various degrees of success. For example, there are manual conversation techniques, whereby the instructor outlines how and when learners contribute to a discussion. There is also the message queue, whereby the chat-channel server withholds the public publishing of a message for a certain timeframe before publishing the next message (Pimental, Fuks, & Lucena, 2005). Redesigns of the chat-channel layout to improve visibility and declutter the user interface have also been implemented. This can include replacing student names with short nicknames or icons, lengthening the vertical dimension, limiting the amount of text, and including greater space between chat messages (Vronay, Smith, & Drucker, 1999). All of these innovations are incremental advancements to improve the chat incoherence and confusion that occurs with large numbers of learners on an online-learning platform.

Whether derived during a live-streamed course or from a streamed-archived course, synchronous-learner, chat-channel data logs offer promising sources of learning data previously unavailable from physical classroom learning models or asynchronous chat forums and bulletin boards. Learner-chat data may provide insights about the academic performance of each student to a learning engineer or an instructor. Also, an overall increase in the rate of questions at a particular point in a streamed lecture may signal to a learning engineer a need for content revision.

Within a live chat channel, if learners discover non-subject-matter social commonalities, this may inspire the sharing of greater and more personal information, exhibiting elements of the Social Penetration Theory first proposed by Altman and Taylor (1973). When people discover some initial common interests and activities, a desire to develop closer relationships emerges, which then requires an additional disclosure of private and sensitive personal information (Altman & Taylor, 1973). We see this exhibited on social media platforms daily. If linked to academic achievement, these new data sets can exploit a unified analysis that combines academic subject data with non-academic socially shared data points for understanding a student’s pattern of educational performance.

**Aggregate Student Learning**

*Aggregate student learning* is an expansion of the concept of *whole student learning*, which is generally understood in post-secondary education to be an extension of
the classroom and lab academic experience to include integrated activities and support from the offices of student affairs, student counseling, and student life in the overall learning plan of a student (Sandeen, 2004). Aggregate student learning (ASL) can be defined as a unified consideration, analysis, and assessment of academic subject data and non-subject, socially-shared data points in measuring achievement within a student’s overall academic performance. ASL, for the purpose of this chapter, implies the cohort consideration, analysis, and assessment of data generated virtually by learners through communication channels within a digital learning platform. In addition, ASL may include additional data points derived from virtual student support services, such as academic advising, professional mentoring, and even student counseling—whether provided by a live-streamed professional or instead via machine-learning, artificial-intelligent chat bot.

Data, But What Data?

The term big data (sets) was coined in the 1990s, primarily due to the immense amount of lab and computer data produced by the ambitious, international human genome sequencing, mapping, and analyzing project led by the NIH’s National Human Genome Research Institute (NHGRI). Since a single human genome has between 20,000–25,000 genes made up of over 3 million base pairs, sequencing and mapping just a few human genomes produced thousands of petabytes of data, requiring farms of servers to store the genetic codes (Hood & Rowen, 2013). Analysis of the genome code exponentially increased standard, data-storage requirements and necessitated the creation of new pattern recognition software and database algorithmic models to provide biosciences researchers with useful results.

As in the biosciences, where big data has allowed us to map our individual genetic predisposition to certain diseases so we can adjust our decisions for a healthier life, education-focused, software-platform innovations can collect big data-sets of student learning, potentially revealing the predisposition of a student’s dynamic learning state relative to their current personal condition and academic setting. These new granular data-sets may also enable personalized instructional interventions and individualized academic support mechanisms, thus helping the teacher and the learning engineer together to sequence and “map” the road to academic success for each student.

Learning engineers and education researchers struggle to determine exact factors that will help increase online education persistence and help improve a learner’s academic performance, as they are missing data from the many external, non-educational factors that also influence a learner’s success. The research literature is rich in studies that demonstrate that cognitive processes (Thompson & O’Brien, 1991; Matthews, 1996), organizational ability (Lawrence, Volet, Dodds, & McGraw, 1985), and time management skills can have as much an impact on persistence rates and academic performance as do
motivation (Vallerand et al., 1992) and personal characteristics (Schiefele & Csikszentmihalyi, 1995). Moreover, student academic achievement can also be negatively affected if the instructor’s teaching methodologies and pedagogical approach differs greatly from the student’s learning style (Cahrkins, O’Toole, & Wetzel, 1985).

Thus, the constrained sources of learning data-sets, conditions, and metrics used to conduct studies related to student success may be the problem in limited effectiveness for evidence-based, instructional interventions (Grant-Vallone, Reid, Umali, & Pohlert, 2003). Scholars are typically not looking at the entire aggregate student learning experience: how a student’s personal, professional, and other external experiences, status, and situations may negatively or positively affect a learner’s persistence and academic performance. It may be that a particular student’s financial status, physical health, personal or romantic relationships are affecting their academic achievement performance. Or a student’s social media, crowd-sourcing activities, personality traits, mood (sentiment), or political beliefs may not be conducive to learning a certain subject within certain academic environments, at certain periods of an academic experience, or under particular learning circumstances.

Big data sets capturing these characteristics, historically unavailable from the typical post-secondary residential student, can now be captured and studied alongside traditional assessment measures. This enables integration of all this data within newly created education-focused algorithms and machine-learning, artificial-intelligent models to generate new categorical data, correlation coefficients, and potential causality relationships that might substantially improve the outcomes of learning engineering and the findings of learning science. “Systems like Moodle, Sakai, and Desire2Learn provide the virtual spaces or class containers. But what happens when the course management systems and the social networking systems merge? Who will oversee it? Where might this lead?” (Bonk, 2009, pp. 345–346). As Dede and Ho (2016) discuss, “We believe that what is happening with data-intensive research and analysis today is comparable to the inventions of the microscope and telescope. Both of these devices revealed new types of data that had always been present but never before accessible” (p. 32). These aggregated, student-learning data-sets have also always existed but rarely have been accessible or quantified, nor have they been collated within a learning experience. The advent of online education platforms and the copious data that they can generate opens up this opportunity. Incorporating and analyzing these new data sets as complements to historical learning research data-sets may be equivalent to the increasingly precise data provided by the inventions of the electron microscope or the geo-stationary space telescope.

Scriyb.com, a (2016) live-streaming, online-learning platform with a coupled, interactive chat channel, attempts to solve some of the issues of scaling learning and teaching outlined above. In particular, Scriyb has developed
several learning-engineering solutions to address the issues of chat confusion and incoherence from large numbers of learners using chat channels to better foster social engagement and communication between learners.

The initial innovation incorporates algorithms that segment large numbers of students into smaller clusters, or groups, of learners formed by using private and public variable data-sets that encourage peer-to-peer and social-learning cohorts (Babad, 2009; Martin, 2015). Each small-grouped class of learners, no more than 30 learners per class in each larger course, is walled-off from all the other groups. In effect, learners in one grouped class are not aware they are actually being taught in real-time in a class of hundreds or thousands of other learners, as they can only see the name or synonym of other learners within their own grouped class. Figure 8.3 exhibits a screenshot of the Scriyb instructor’s UI with four grouped classes of 20 students each. Each grouped class has its own communication chat channel, and the instructor can view all the groups’ chats at the same time or separate groups at any one time.

The datasets that determine each grouped class are divided into three variable groups: basic (r), education (d), and personal (z). Basic private variables can consist of a learner’s name, home address, birthdate, sex, or any litany of requested registration data at login. Education (d) data could include school, major, current GPA, highest GPA, graduation year, or similar data sets. Lastly, the third variable, personal (z), includes interests and activities, hobbies, or any number of additional personal data points. The z algorithms also scrape the public web (Internet) to both confirm and enhance the overall data-set. These private and public variables feed the algorithms to determine each student’s current academic achievement levels, which generates a silhouette to represent each student during and after an academic experience.

“Idealized” Learner Group Composition

After a learner’s silhouette is created, prior to the launch of an academic experience, the next algorithm forms each idealized class group. Since the intent is to create groups of learners that encourage social engagement and social learning, silhouettes are identified and tagged as high, mid, or low academic achievers. Social interests and commonalities between learners are also compared and matrixed, and they are blended together to create an idealized academic/social environment. Class composition between high, mid, and low achievers is critical to success: a blended formula of the peer-effect models found in the boutique (students perform higher when surrounded by students similar to themselves), shining light (a single high achiever provides a great example for all other students), and the rainbow (where an assortment of student abilities benefits the entire group) peer-effect model (Sacerdote, 2011).

The algorithmic models that create the idealized learner groups do not just place all the most gregarious learners into the same grouped class—or the learners
FIGURE 8.3 Scriyb Instructor User Interface, with Group 2 chat viewable on the right-side column
from the wealthiest or poorest neighborhoods—but they compile learners based on the following academic percentages:

- **High-Achievers:** 30% (+/- 3%)
- **Mid-Achievers:** 48% (+/- 3%)
- **Low-Achievers:** 22% (+/- 3%)

Student academic achievement percentage schemata are inspired by the modern multi-player entertainment game platform, whereby through chat (and sometimes voice and video), the more experienced players within a gaming team guide the less experienced players to win a challenge. In an online-learning environment, the analogue is for a team of students to understand and successfully complete an academic assignment. Figure 8.4 shows an example of how a grouped class is compiled, with high, mid, and low achievers each with visually identified r, d, and z variable-connected commonalities.

Caroline Hoxby and Gretchen Weingarth’s seminal paper, “Taking Race Out of the Equation: School Reassignment and The Structure of Peer Effect” (2006), discovered through their 20-year study at Wake County School District, North Carolina, that low-achievement students appeared to greatly benefit academically from interacting with high-achievers with similar interests (personality characteristics and commonalities) in the same classroom. Additionally, they also discovered that high achievers measurably benefited from studying alongside low achievers with the same interests and commonalities (the same type), demonstrating that when combining students with disparate academic achievement levels in order to raise the academic threshold of an overall class, the students within a class must “maintain continuity of types” (Hoxby & Weingarth, 2006, 33). Further, at the higher education level, Stockdale and Williams (2004) examined the differential academic effects on high, middle, and low achieving students working together within a five-member group project metric and discovered that low-

![Example of Grouping Variables in a Course of 30 Students](image)

**FIGURE 8.4** Example of grouping students (clustering) using r, d, z variables
and mid-tier students improved significantly over a two-semester introductory psychology course from their previous mean academic scores.

The social engagement and peer-to-peer learning strategies that initially grouped learners have fostered personal and group friendships, personal relationships, and virtual social bonding. When realized over time, virtual student bonding produces both subject matter and non-subject matter semi-public sharing of private and sensitive personal information, illustrating the Social Penetration Theory mentioned previously (Altman & Taylor, 1973).

**Dynamic Online Class Regrouping and Rebalancing**

If the algorithmic assumptions that initially grouped learners within a class were incorrect, or the interpolation mathematics was off-target in creating the initial groups for a class, the dynamic-regrouping, machine-learning algorithms can correct these placement errors (Martin, 2016). The dynamic regrouping innovation allows for active-learner performance tracking and analysis, in order to regroup or subgroup learners within a live-streaming course. All instructor-to-student communications and all peer-to-peer student academic subject matter and non-subject matter social communications are parsed, tagged, and indexed. This data is then compared and measured within a matrix against instructor assignment results posted on the Scriyb platform.

For example, if a target learner had previously been defined as a high achiever prior to the launch of an academic experience (initial grouping), and during their academic experience their silhouette stays in the upper 30% in academic-achievement output compared to the other students in the same online grouped class, then the student remains in the same virtual synchronous class. If the same target student does not perform academically at or above the 30% threshold and begins to demonstrate a 48% mid-tier achievement rating compared to the other students in the online grouped class, then the learner is moved to a subclass (high, mid, low) mid-achievement tier but stays in the same online grouped class. However, if the online grouped class has no openings to accommodate a mid-achieving student tier, as all the other originally grouped mid-achieving students have been successful at maintaining their original academic rating, then the target student is dynamically moved seamlessly to a different idealized online grouped class prior to the next synchronous classroom meeting. Figure 8.5 provides an example of the sub-regrouping of an academically declining student with a strong social, non-subject link between another student and the dynamic regrouping of a declining student into another online class within the same course of study.

The goal of the dynamic regrouping innovation is to maintain the initial grouping based on similar-type students of 30% high-achievers, 48% mid-achievers, and 22% low-achievers within the same grouped classroom during an academic term that facilitates and maintains social learning and engagement. The dynamic regrouping innovation allows for the analysis of additional measurable
attributes found in Social Learning Theory (Bandura, 1963), Peer-to-Peer Cohort Learning (Davidson & Goldberg, 2009), as well as Team-Based Learning (Michaelsen & Sweet, 2008) combined with traditional grading mechanisms to influence a learner’s academic outcome. Integrating all these dimensions provides a deeper analysis in the diagnostic report of a student’s understanding of the subject matter taught.

**Deep Academic Learning Intelligence**

The Scriyb, interactive live-streaming platform previously discussed also employs Deep Academic Learning Intelligence (DALI) to assist an online learner with artificial-intelligence-driven student-support services and guidance as it pertains to their academic journey (Martin, 2017). DALI consists of multiple, deep, neural language networks (DNLN)—artificial intelligence algorithmic models that provide students academic advising, personal counseling, and individual mentoring derived from the massive dynamic group learning that includes academic performance history, subject-based and non-subject based communication content understanding, and social and interpersonal behavioral analysis. Data-modeling capabilities allow DALI to build learner communication and social-interaction maps, to derive cohort links between these interactions and academic performance, and to inform a learner, for example, that class absences, missed assignments, or misdirected social interactivity may be a detriment to their class’s academic goals.

Advising and student counseling can be as important to academic success as actual classroom instruction, especially for first-time learners and historically underserved populations (Towles, Ellis, & Spencer, 1993; Nichols, 2010). Figure 8.6 shows an example of the DALI Deep Machine-Learning Model with \( n \) number of inputs, hidden layers, and outputs.

![FIGURE 8.5 Sub-regrouping and dynamic regrouping of students](image-url)
Over time, DALI learns, or more specifically is trained about each learner’s evolving external environment, condition, state, and situation (non-subject matter) as they impact, or may impact, their academic performance within the Scriyb online-learning platform. In order to understand the differences between academic and pure social non-academic chat communication within an academic learning experience, DALI models need pre-trained classifiers about both, as well as about the academic scoring matrix used and the intervention methods and solutions available to offer a learner.

Upon detecting a potential issue, shared through a communication channel with an instructor or another learner in their same grouped class, DALI will make corrective recommendations to the learner to help remediate and modify potential negative academic outcomes. A learner personally trains their DALI models by responding to an intervention recommendation with a simple click of either “Yes I will,” “No Thanks,” “Maybe,” “Ignore,” and if it was “helpful.” A learner’s initial response options from the recommendation are limited to “Yes I will,” “No Thanks,” “Maybe,” and “Ignore,” but a follow up on the “helpful” solicitation allows DALI to receive an even greater entropy vector to more fully offer accurate and impactful recommendations in the future (similar to an Amazon.com verified customer product or service feedback or review). Figure 8.7 shows an example of an initial DALI recommendation, and then the follow-up “helpful” solicitation with resultant learner response. Every learner’s initial grouped class and dynamic regrouping data, every DALI recommendation, and every learner response are stored in a learner’s Omega Learning Map (ΩLM).

**Omega Learning Map and DALI**

The Omega Learning Map (Martin, Naidich, Martin, Trang, & Mohamed, 2017) is a joint, artificial-cognitive-declarative-memory map designed to dynamically
store, retrieve, and recall machine deep-learning data-sets derived from an Aggregate Student Learning (ASL) rubric. This new memory storage model is required for useful mapping and retrieval of the immense student data-sets resulting from utilizing multiple, interleaved, deep-learning algorithms to parse, tag, and index academic, communication, and social learner data cohorts derived from academic achievement, social (personality), and behavioral analysis. In order for deep-learning models to dynamically store and retrieve parsed and classified data-sets in an artificial cognitive memory map, such as the ΩLM, these data-sets must arrive into the model pre-assigned, processed, and classified, mimicking a biological cognitive memory model.

Knowledge Acquisition System

The ΩLM’s Knowledge Acquisition System (Martin, Naidich, Martin, Trang, & Mohamed, 2017) is an expanded and refined version of the original, after-image, primary-and-secondary memory model first proposed by William James in 1890. The design of the ΩLM transposes James’s after-image memory model into a sensory memory module that contains both current and historical learner’s data, defined as semantic inputs, and the communication channels (text, spoken, visual) of the learner’s experiences and defines them as episodic inputs. The Knowledge Acquisition System (KAS) also divides James’s primary memory model into working and short-term memory modules. Further, James’s secondary model is translated as a unique long-term memory module containing a learner’s declarative memory experiences—including sensory and episodic memory inputs—and includes a recommendation best-practices module called the universal memory bank. Figure 8.8 outlines the ΩLM Knowledge Acquisition System (KAS).

Knowledge within this system is defined as information required to make recommendations to a learner based on a set of pre-defined academic states and social conditions. To accurately record and store this information for every

FIGURE 8.7 Demonstration of an initial DALI recommendation and follow-up “helpful” solicitation with resultant student response
FIGURE 8.8 Complete Scriyb Knowledge Acquisition System Model
learner, a detailed and organized data recognition and storage process must be implemented. The goal of the KAS is to perform this recognition and storage function that mimics the various memory systems of the human pre-frontal lobe and hippocampus. The separation of the memory process into several independent and parallel memory modules is required as these separate memory systems serve separate and incompatible purposes (Squire, 2004). The Knowledge Acquisition System is divided into four, prime-variable groups, each representative of a human hippocampus model including sensory memory (v), working memory (w), short-term memory (m), and long-term memory (l). Another unique feature of the KAS invention is the Universal Memory Bank (UMB). The UMB tags and indexes all learners’ parsed data from an integrated cohort experience, which is the sum of the DALI suggestions, recommendations, responses, and the follow-up helpful responses that may represent potential universal conditions that another learner may experience in the future, outside any one learner’s $\Omega_{LM}$, decoupled from any one learner’s silhouette, and within a generic long-term memory (LTM) UMB. If a tagged cohort experience is recognized as similar by DALI as another learner’s conditional experience, DALI may only provide previously successful recommendations to help ameliorate the issue or conflict, thereby using other learners’ $\Omega_{LM}$ resultant data to solve a different learner’s co-similar issue.

**Summary**

Preliminary qualitative evidence derived from two Scriyb online courses held in the summers of 2016 and 2017\(^1\) demonstrated that an average of 84% of properly-grouped learners in their cohorts (high $z$-variable with similar commonalities and interests) volunteered answers to questions that their peers posted in the platform communication chat channels prior to the instructor answering the questions, and that an average of 79% of the time the answers provided by learners for learners were correct (or near-correct) according to the post-course, chat-log-data analysis conducted by DALI and the course instructors. Furthermore, based on post-class assessments, it appeared that both subject matter and non-subject matter social engagement equally had an effect on academic achievement levels and that the more socially engaged a student was with other students within and outside the learning experience, as expressed within the platform communication chat channels, the greater their academic achievement results.\(^2\)

Although much research still needs to be conducted to prove the efficacy of the hypotheses that using interactive, live-streaming, online-education-related platforms improves learner engagement, and that real-time social learning may have a direct impact on overall academic success, there is a multitude of scholarly literature from traditional classroom and residential student studies that substantiates assumptions and preliminary evidence that these recent innovations may transform the learning
experiences for many learners. Moreover, interactive, live-streaming, online-learning platforms coupled with integrated, machine-learning solutions may provide not just new granular-level learning-science data-sets to harvest and comb for new insights into the learning process, but they may also may provide learners with new vehicles of academic, personal, and professional direction, guidance, and support to help them improve academic performance. Deep machine-learning models applied to teaching and learning may hold the potential to help finally scale educational opportunities for a greater number of learners globally and also may provide true personal-learning pathways to help facilitate academic success.

The authors believe the innovations outlined in this chapter hold promise to overcome some of the most pressing and common historical, online-learning impediments and obstacles that have negatively affected online-learner retention, satisfaction, and academic performance, and to help facilitate the online learner’s joy of learning and help recapture that learning magic moment that so often occurs within the traditional physical classroom.

Acknowledgements

The authors wish to thank Vera Lichtenberg, PSTH, Grace Davis, James Casey, Stephanie Kane, Rizza Villacorte, Thanh and Phuong Trang, and the Aspiring Scientists Summer Internship Program.

Notes

1 Summer 2016: n = 12 high school and college freshman, median age = 17, Art & Animation I course, 34 hours of instruction over five days. Summer 2017, n = 15 high school and college freshman, median age = 18, Cybersecurity I course, 30 hours of instruction over five days. Both courses were taught by experienced, but non-tenured, university-level faculty. Questions and answers and their frequency were analyzed for the aggregate as well as both classes individually using descriptive statistics. Where relevant, mean and standard deviation were calculated.

2 The greater the number of messages posted by a student, subject and non-subject aggregated, the mean assessment score increased 9.9% in Summer 2016. The mean assessment score increased 12.3% in Summer 2017.

References


EXECUTING THE CHANGE TO LEARNING ENGINEERING AT SCALE

Bror Saxberg

Introduction

As we can see throughout this volume, there are many learning-science results with promising applicability to education at scale. The challenge we face—in addition to expanding these results and understanding under what conditions they work the best (e.g., for which learners and for which topics?)—is how to actually implement these results at scale. Our industry has no experience with what Herbert Simon (as described in the preface to this book) called a “learning engineering” approach to improving short- or long-term learning outcomes: systematic ways to creatively apply what is empirically known about learning (i.e., learning science) at scale, in real-world situations, under real-world constraints.

This chapter describes a number of issues that arise as we try to take learning-science opportunities to application at scale. It discusses how we at Kaplan engaged in local and global organizational change to make evidence-based approaches to learning improvements “the way we do things around here,” meaning the application of learning science results and good learning-evidence practices to make measurable progress in real-world situations under real-world constraints—learning engineering. We hope more organizations engage deeply in these approaches and ideas and believe our whole sector will benefit by more of us collaborating on this work.

Getting started with change at scale is challenging. We will discuss four key phases we put our own organization through to make the change to learning engineering at scale:

1. Expose the organization to possibilities.
2. Educate the key action takers.
3. Accelerate systematic effort.
4. Embed cyclic evaluation and improvement.

**Expose the Organization to Possibilities**

Looking across the landscape of learning delivery at scale, there is an unfortunate lack of awareness of the evidence that exists about how learning works (and does not work) and very little experience in translating any of this evidence into practical approaches at scale to learning environment improvement. To generate the energy needed to start the change, we can first expose key decision-makers (managers, developers, teachers, and more) to the potential of these approaches, as well as to examples. With this, we can excite a few early adopters into working on their own examples within their own learning environments. When successful, such efforts can then motivate a wider internal audience into thinking “this can work here.”

**Lack of Use at Scale of the Evidence about Learning**

Empirical research on how learning works has made remarkable progress over many decades, and a detailed, comprehensive review is well beyond the purview of a single chapter (Clark & Mayer, 2016; Hattie & Yates, 2013). However, a few well-supported hypotheses are worth noting because of their implications for learning at scale.

From a cognitive standpoint, one of the most profound insights is the division (roughly) of how the mind works between working memory and long-term memory.

**FIGURE 9.1** Working memory—long-term memory model
memory. Working memory is the part of our minds that houses our verbal selves—the “talky-talky” part of our mind that runs our internal monologues as we live in the world. It is where sensory input from auditory and visual channels (and more) is first engaged. It is also the most flexible part of our mind, handling the difficult things we have to decide and do, including much learning. Unfortunately, it has very limited capacity—just a few things can be handled at once—and, to boot, is error-prone, and often becomes overloaded. Finally, it only stores information for a short period of time.

If we only had working memory to draw from, we would be in trouble. Fortunately, the other part of our mind steps in to help—long-term memory. Long-term memory (as its name suggests) keeps things around longer (months and years are possible), but also can run many processes in parallel, and, if trained properly, makes very few errors. However, it is also very rigid in what it can process (but it can draw immediate attention from working memory if the situation steps out of bounds), and, very inconveniently for instructional design, is not easily verbally accessible. Conveniently, it has a near-perfect connection to working memory, so that drawing on processing from long-term memory feels effortless, obvious, and intuitive. (Which contributes to the difficulty of using experts to teach or design teaching—much that is obvious to an expert is arcane for a novice.)

Some kinds of competence can reside completely in long-term memory. For example, many people have the experience of planning to drive to one place, only to begin thinking about something else in their lives, and realize too late they have driven to another. In those circumstances, we rarely pause to wonder, “Who drove us to the wrong place?” What this shows is that very complex processes can run automatically in long-term memory and independent of working memory.

Often, however, expertise is a collaboration between working memory and long-term memory, even after years of practice and feedback. For example, although you can plan a summer vacation while driving to work, you will most likely never be able to write a persuasive essay while planning that summer vacation at the same time. Many important expert decisions and tasks require practiced collaboration between things mastered in long-term memory with complex considerations that take up working memory.

There are a variety of techniques (Tofel-Grehl & Feldon, 2013) for unpacking expertise derived from cognitive science protocols like think-alouds developed from the 1960s forward. In addition, a wide array of laboratory experiments suggests ways to design instruction to limit the burdens on working memory while taking advantage of the parallel wiring of audio and visual information channels to lift learning performance (Clark & Mayer, 2016). Indeed, there are now well-supported frameworks for how to link evidence-based instructional practices to different types of learning outcomes (Koedinger, Corbett, & Perfetti, 2010), which we deployed at scale within Kaplan. (See later sections of this chapter.)

Beyond the cognitive dimension of mastery and learning, we also have to take into account motivation: what gets in the way of students starting, persisting, and
putting mental effort into (presumably) well-designed learning activities? In a survey of a variety of motivation research undertaken by Richard Clark of USC (Clark & Saxberg, 2018), four different factors seem most problematic for people to be motivated (defined as being willing to start, persist, and put in mental effort): valuing, self-efficacy, attribution, and negative emotional states. (Note that liking is not a part of this definition of motivation—for learning purposes, if you start, persist, and put in mental effort on well-designed learning activities, your brain will in fact change and learn regardless of whether you like it or not. This parallels (Dweck, 2006) how muscles work—you may not enjoy weight training, but if you start, persist, and put in the effort, your muscles will, indeed, change.):

- Valuing. If you value what you are doing and/or how you are doing it, you are likely to start, persist, and put in mental effort. This could be because you simply enjoy the activities themselves (the liking bit), but it could also, or instead, be because you see the value of it for your future. You may also value what you are doing because it puts your own existing expertise to work—people simply enjoy exercising their competence.

- Self-efficacy. If you do not think you can do something, regardless of how valuable you find it, you may not start, persist, or put in mental effort. This is different than valuing it, and the solution to self-efficacy issues is also different: instead of focusing on how important an activity or an outcome is to a learner’s future, the right intervention is to refuse to agree with the sense that “you can’t do this” and instead show how you have already done things related to these outcomes, and how others like you have also managed to master these outcomes.

- Attribution. Related to a failure of self-efficacy, another motivational failure ties to factors learners believes are in their way—outside themselves and their own control. “My teacher hates me,” “No one could understand this text/media,” “I don’t have a enough time,” “I don’t have a good place to study”—the reasons are legion. Here, the correct intervention is to problem-solve around the claimed issues, helping to resolve the specifics, while at the same time modelling for learners how one handles this kind of issue. In other words, tenacity among all parties (educators as well as learners) is an important part of overcoming attribution issues.

- Negative emotional states. The most difficult and complex of situations is when a learner is preoccupied with a negative emotional state: they are angry, scared, depressed, and so forth. Such emotions will naturally prevent learners from starting, persisting, and putting in mental effort to challenging but important learning tasks and will likely require a full armamentarium of social services and other approaches to resolve.

What this work on motivation reveals is the possibility of a diagnostic and treatment framework for motivation issues to apply alongside cognitive diagnostic
measures—to understand what gaps in a person’s motivation or mastery might impede progress. It is tricky—a learner who is not making progress could be suffering from a cognitive issue or from a motivation issue or both. If you assume the issue is cognitive when it is really motivation, you run the risk of rushing in with new learning work that the learner simply does not have the steam to tackle. On the other hand, rushing in with the wrong intervention for a motivation issue may not help—and may make matters worse. For example, treating a learner with a self-efficacy issue—“I can’t do math”—with an intervention on value—how important mastery of mathematics is for people in careers like theirs—can make things worse.

So there is much already known about how expertise is put together and how the cognitive and motivational dimensions of learning work. That makes it especially unfortunate (not to say alarming) that very little of this decades’ old work has made its way into practice at scale: not to university faculty, nor publishers, nor ed-tech developers and financers, nor to schools of education, nor to regulators and accreditors.

A detailed analysis of why this is the case is also beyond the purview of this chapter, but it may be related to the history of professional education as a whole (Ravitch, 2001). In the early part of the 20th century, when both schools of medicine and schools of education were being set up to professionalize these careers, medicine had the advantage of events like the impact of sulfa drugs on death rates in World War I to push strongly the incorporation of science into new approaches to teaching medicine. At the same time, cognitive science was in a bit of a dead end because of the complexity of the biology of the brain and was retreating to a simplistic behaviorist model of mind that left little beyond stimulus and response.

That meant schools of medicine were set up from the beginning to pay attention to science as a foundation for medical training and to the methods of science (including randomized controlled trials, careful understanding of the biological foundations of function and disease, and requirements to demonstrate efficacy before at-scale deployment). By contrast, schools of education, seeing little of relevance in cognitive science at the time, did not focus early teacher education on an equivalent science of learning. Unfortunately, even after cognitive scientists discovered a new path forward by looking to model higher-level information processing in minds rather than low-level neural functioning (the cognitive revolution in the 1940s and 1950s drew on parallels with descriptions of information processing in then-new computers), schools of education remained apart from the rapidly growing body of evidence that resulted.

This lack of focus on evidence-based approaches to learning carried through to the whole enterprise of higher education, even though the new cognitive science research work emerged from university departments. Arguably, the strong faculty career focus on research in many institutions, privileged over the work required to measure student success over time and optimize it, slowed down adoption of evidence-based methods for delivery of student learning in the higher-education sphere.
Finally, corporate training, perhaps drawing on experiences from higher education, in spite of the material (including financial) benefits of improving performance of employees in high volume, high value, high variance job categories from the median to top levels, has also not yet taken advantage, at scale, of the full array of evidence-based possibilities. This is harder to understand, given the usual capability of markets to take advantage of more efficient, higher value execution and delivery, but the fact that most experts in corporate training come not from cognitive science backgrounds, but rather from a wide variety of other often traditional education backgrounds, may have held them back.

Over time, though, the dead hand of Keynes is likely to force the issue (Edmondson & Saxberg, 2017): As the nature of work begins to change more rapidly under the influence of information-rich appliances and robotics, organizations will gain more and more value from the ability to systematically and reliably change the skills of their workforce. They will not be able to simply “fire and hire” to get the competence they need because the life-cycle of skills will become too short (down to quarters vs. years) to make firing and rehiring employees rapidly a workable solution when compared to continuous retraining.

The incorporation of technology into learning has not helped. As discussed in Chapter 3, the use of technology to enhance learning has gone through repeated cycles of enthusiasm, adoption, and then disgruntlement since the 1930s (Hess & Saxberg, 2014). From the radio broadcast of lectures to the use of film or television for something similar and on into the computer-centered innovations of the last few decades, the fanfare that greets “the new way to learn” has not led to real improvements in learning at scale.

The problem is that technology in itself does nothing for learning (Clark, 1983). What it does do is make good or bad learning more affordable, more reliable, more accessible, more data-rich, and more personalized. However, you need to start with good learning solutions to hope to have technology make a positive difference to learning itself.

We now have enough information about how expertise and learning work that we can begin to look at technology-enhanced learning in a better way. Instead of starting with cool technology, we should focus on what the actual learning problems are and what new activity by each mind would likely be measurably better than what that mind is doing now. Once we have a sense of what would be better for the mind to be doing—more practice and feedback, more personalized connection to interests and background, less cluttered visual and auditory inputs, and so forth—we can then see if technology provides any advantage for the delivery of that better learning experience.

**Jump-Start the Change with Early Adopters**

Given the lack of experience at scale in applying evidence-based methods in most organizations, we have to start by building some exposure to the potential
benefits and then begin working with individual early adopters to bring the new
techniques to their work, to show the difference it makes.

At Kaplan, we began with an overall framework to describe the approach that
evidence suggests works. (See Figure 9.2.) This framework brings together four
key aspects of evidence-based instructional design that we can use to explain why
an organization would follow this approach:

- Making the ultimate learning outcomes clear is key not just to designing the
  learning well but also to making clear the value of the intended learning. Unfortunately, as we have seen, what expert minds actually decide and do is not easy to unpack from their minds—they themselves do not have verbal access to all that has been automated into long-term memory. Investment is needed here to ensure we are targeting a full set of what experts decide and do, including in cognitive-science-based techniques like cognitive task analysis (Tofel-Grehl & Feldon, 2013).
- Evidence-based instructional design methods then provide the route for learners to more efficiently and effectively master these learning outcomes. Even though the evidence often comes from laboratories, not real-world situations, there is real promise in trying to apply what this evidence suggests about how learning works.
- For a learner, well-designed, valid, and reliable learning measurements over
time are key both to evaluating how it is going and to seeing if changes to improve difficult parts of the learning environment are in fact making learning work better.
- Finally, systematic use of relatively rapid, well-designed pilots, especially with technology-enhanced learning tools, closes the loop on iterative improvements.

This allows for a “learning engineering” process

**FIGURE 9.2** Wheel of evidence-based learning
Well-designed pilots (e.g., randomized controlled trials when we can do them) also provide strong arguments for how new learning approaches add value to an organization by making it clear that only the changes made could account for the changes in improvement observed. Such pilots are powerful learning-engineering tools.

Once we have exposed people in the organization to what evidence exists about different approaches to learning and assessment, and we have begun to establish a framework by which the organization can start to systematically apply this evidence, we can begin to work with early adopters across the organization who see the potential for the benefits in application to their areas.

Within Kaplan, we had early engagement from our Kaplan Test Preparation (KTP) group on potential benefits from applying research on cognitive science results. KTP is a national leader in preparing learners for the Law School Admission Test (LSAT). One segment of this high-stakes assessment includes challenging logical reasoning problems with complex question stems and distractors, which KTP has known for decades are very difficult for learners to master. KTP has become very good at training learners for these assessments and, as they contemplated how to use multimedia to help, felt the obvious approach was a video on the Kaplan approach to logical reasoning problems on the LSAT, together with a workbook.

However, one of the KTP learning engineers realized that the kind of very complex reasoning challenge posed by the logical reasoning problems would tend to overload the working memory of most learners. He knew there had been decades of research by John Sweller and colleagues (Sweller, Ayres, & Kalyuga, 2011) on the use of worked examples as a straightforward way to help learners deal with these kinds of problems: learners look through a number of worked-out problem examples with brief, expert commentary. He proposed providing learners with worked examples of reasoning by expert LSAT logical reasoning problems instead of the long-form video plus a workbook.

Initially met by skepticism, the decision was made to run a randomized controlled trial to compare the use of worked examples to the long-form Kaplan video plus workbook. (See Figure 9.3.) As reported at the American Educational Research Association (AERA) in 2013 (Rudman, Niemi, & Sweller, 2013), Kaplan found that the 90-minute video plus workbook did not significantly improve learners’ performance on logical reasoning problems over no preparation (a somewhat distressing outcome for a test preparation organization!). However, the two worked example conditions, one using 15 worked examples (13 minutes of effort by learners) and one using eight worked examples (around eight minutes of effort by learners) worked statistically significantly better at helping learners master these complex problems.

This led to serious introspective conversations within the KTP product development organization. It was so “obvious” that video would be better for
learners—yet the evidence demonstrated clearly that use of worked examples took less time for learners (eight minutes for eight worked examples), led to higher logical reasoning problem success, and was very much less expensive to develop as a learning intervention (eight Power Point slides vs. an hour of professionally produced video plus a workbook).

The KTP product development team took away the correct conclusion from this experience—there was a clear need to incorporate learning science results early in the product development process. Where there was relevant evidence to guide a learning intervention, it needed to be brought to bear early in the prototyping phase of product development—intuitions about what worked for learning did not always lead to the best results.

**Educate the Key Action-Takers**

With energy built up from awareness of external work showing the promise of evidence-based approaches to learning and learning measurement, and with the success of key initial internal projects consistent with this external promise, the stage is set for wider use of these methods.

To efficiently train and coach the learning developers around the organization, we began to build out systematic processes and job aids for designing instruction
Learning science pushes us to backwards design

![Backward design diagram](image)

**FIGURE 9.4** Backward design

tied to the well-known concept of backward design for both instruction and assessment. (See Figure 9.4.) The basic idea is to start from the detailed description of what we intend learners to be able to decide and do at the end of a segment of instruction and work backwards to design what evidence/assessment data we want to collect to demonstrate mastery, along with the required practice, feedback, demonstrations, information, and overviews.

We also created streamlined guidance grounded in learning science on how to design towards mastery. (See Figure 9.5.)

- Learning starts at a declarative level, where everything lives in working memory with no support from long-term memory. This is a very frustrating stage for learners—they feel as if they are processing very slowly (they are), they are forgetting key things very quickly (they are), and that they are making many mistakes (they are). There is little help for it right at the beginning: they need support by having visual aids to keep reminding/refilling working memory with stages/steps.

**Design starts from how expertise gets acquired**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Implications for Instructional Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>Clear information displays, e.g., job aids, examples, reference material</td>
</tr>
<tr>
<td>Procedural</td>
<td>Build varied Practice tasks, and rich feedback/coaching</td>
</tr>
<tr>
<td>Automated</td>
<td>Repeated frequent practice to build speed and accuracy</td>
</tr>
</tbody>
</table>

**FIGURE 9.5** Stages of mastery
After a few passes in the declarative stage, learners begin to gain familiarity with how to handle similar situations—there are the beginnings of mastery within long-term memory to draw on. It is at this stage that a learner starts to benefit from feedback—they can liberate enough space in working memory to pay attention simultaneously to the task at hand and to feedback from prior attempts. This is a stage that also benefits from a wide array of settings to cement and deepen mastery.

For some aspects of learning, another stage can be reached where a mind’s processing becomes essentially automatic, completely resident in long-term memory. Some complex tasks, but not all, can indeed reside entirely in long-term memory: for example, as mentioned earlier, an experienced driver navigating to work leaves plenty of space in working memory to plan a summer vacation. For those components of mastery to become automatic requires repeated practice and feedback over an extended period of time.

We also built a framework to help our instructional designers think through the connection between types of learning outcomes and related, evidence-based approaches. (See Figure 9.6.) We based this on a simplified version of Carnegie Mellon’s KLI framework (Koedinger, Corbett, & Perfetti, 2010). The basic idea is to make an advance on Bloom’s Taxonomy, which was developed decades ago before there was much awareness of how learning actually works. (Bloom’s Taxonomy is a bit like Aristotelian physics—a fine armchair approach to thinking about different types of learning outcomes, useful in its own way, but missing connections to evidence about how the world, in this case learning, actually works.)

At the top level, we are seeking to give learners mastery over complex, cognitive, procedures. These really are complex (e.g., “Make a one page summary of...
10 or more pages of information.”). A job or profession is made up of dozens of these kinds of complex cognitive problem-solving capabilities.

Each of these complex procedures is supported by essential knowledge components: facts, concepts, processes, and principles. In turn, each of these kinds of outcomes has research supporting best practices for mastery (e.g., spaced repetition for facts). We wanted to give our learning engineers a straightforward way to stay aware of evidence-based approaches to the different types of learning outcomes.

We have also begun to instill a similar structured set of approaches from good psychometric practice to ensure that evidence gathering ends up with sufficient validity and reliability to trust the measures. (See Figure 9.7—AERA, 2014.)

With these frameworks as foci, we can build training for the entire organization to begin to effectively use evidence-based approaches. Ultimately, we have to be aware that the people working on instruction have minds with the same

![Figure 9.7: Approach to validity and reliability, grounded in NCME](image-url)
needs as their learners have for motivation, practice, and feedback, and good coaching and collaboration to achieve new outcomes.

We designed a hybrid training program to accomplish this. (See Figure 9.8.) After an asynchronous, online, instructional-training program of around 20–30 hours, we ran coached projects to increase the mastery of evidence-based instructional design. This work was a part-time effort for a significant fraction of instructional designers (40% of the total), giving them opportunities to collaborate with other Kaplan instructional designers on instructional-design projects of interest to them, designed in an evidence-based way that matched their online training and coached by a more experienced evidence-based instructional designer.

Changing human skills is difficult, so at scale requires a significant investment of time and effort. In our case, this effort was a multi-month, part-time commitment by the instructional designers (and their learning organizations), but it was consistent with the evidence on how to reliably change human expertise.

As we will discuss later, in addition to this one-time, up-front training, we also had our instructional designers evaluate each other’s work across the organization using an evidence-based checklist. While providing good feedback to development teams on how to change/improve their own learning services, a secondary impact was to provide instructional designers with a chance to problem-solve in new learning areas with the same framework on learning they had been trained on and intended to work with in their service areas. Consistent with research on concept mastery, applying new concepts to a new area is an excellent way to deepen mastery. Reviewing and suggesting improvements to colleagues about the application of learning science to their areas did this.

**Accelerate Systematic Efforts**

With blueprints and training in place for evidence-based approaches to improving learning environments, the hardest part of the work begins: actually using these approaches to make a difference to learning in the operations at scale. Larger organizations like Kaplan often have multiple, distinct groups working on learning environments. With different domains and outcomes, together with context-specific feedback from teachers, learners, and other stakeholders, we should not be surprised to end up with different learning solutions across the organization. What we want, however, is for these different learning groups to still share common underlying principles that they apply to their own unique learning and other needs. We also want them to actively discuss and learn from each other (and from outside resources) using these principles to get new ideas and to generate evidence about how to improve.

To encourage this collaboration over time, in addition to the common training described above, at Kaplan we pulled all the learning engineers around the organization into a community of practice. This community of learning engineers
It’s real work to alter how a large number of IDs build...
(the Kaplan Learning Architect Community) is beginning to work across boundaries (geographic, subject, learning context) to deepen their understanding of evidence-based methods and issues. The community has access to experts (internal and external), access to common training, opportunities to learn from each other’s pilots and experience, and the chance to work together on problems/opportunities of common interest (e.g., motivation issues, assessment design issues, learning analytics, mobile learning, and more).

However, embedding these practices as part of the culture in a large organization requires serious commitment and time from many layers of the organization not just from the learning engineers. We also need to involve the general managers, those who allocate capital, people, and other resources. They have to make trade-offs among a range of specialized functional areas (marketing, IT, delivery, real estate, etc.) including resources for learning-environment development and improvement. We need these general managers to become as comfortable engaging in decision-making about evidence-based learning approaches as they are in other areas. Unfortunately, very few general managers in the education sector have any background in evidence-based approaches to learning.

To make such a change requires strong intentional support of the most senior managers and even the board of directors. Within Kaplan, our global CEO and CFO (Andrew Rosen and Matthew Seelye, respectively) supported this vision of general managers being able to engage more fluently in evidence-based trade-off conversations about learning.

To push this forward, we designed a series of twice yearly conversations between the global CEO, the global CFO, the global CLO, and each general manager supported by their learning team. (See Figure 9.9.) This was a large commitment of time from the global CEO and CFO—approximately twenty hours, twice yearly—but that very commitment sent a signal about the importance of this new work. It also made it easier for the global CEO and CFO to compare the different learning organizations and their stage of maturity in thinking clearly about trade-offs—having side-by-side conversations created a rich context for discussions about relative progress.

**Embed Cyclic Evaluation and Improvement**

Having exposed people to what’s possible with evidence-based approaches, educated key personnel on how to use these in practical circumstances, and expended significant effort in getting widespread adoption, we can begin to systematically improve learning over time. We have to make sure these approaches stick—become a part of the culture and lead to on-going cycles of improvement grounded in evidence, both from within the organization and outside. For Kaplan, this has meant investing in regular reviews of instructional design, assessment quality, and the generation and evaluation of portfolios of pilots.
Established a General Manager review process to focus on learning tradeoffs and essential ingredients for quality.

**Learning Outcomes**
- Derived from experts or expert associations and/or regulatory bodies
- Tied in with domain specific research (if any)
- Clearly described and tied to what learners are intended to be able to decide and do

**Assessment**
- Research-based development processes
- Expert reviews
- Item, test and scoring reliability studies
- Evidence that assessment does what it's supposed to do (e.g., accurately predict future test scores; measure competence)

**Instructional Design**
- Lesson components are tightly aligned with learning outcomes
- Course proactively addresses student motivation

**Data**

**Focus for Sept 2016**

**Focus for April 2017**

**Focus for Sept 2017**

**FIGURE 9.9** Systematic approach to twice-yearly conversations about learning priorities, challenges, trade-offs
One of the approaches Kaplan is using has been mentioned before—they have all their learning engineers use a learning checklist derived from good research about learning to evaluate each other’s learning environments over time. (See Figure 9.10; United States Department of Education, 2015.) The checklist includes a variety of instructional design components (e.g., learning outcomes, assessment, practice and feedback, demonstrations, overviews) but also ties into motivation and integration.

This checklist works several ways. First, it makes visible how a learning-service team can efficiently improve. One of the earliest results brought to the senior leadership team at Kaplan was a review of nine high-volume learning services, showing that all was not well. (See Figure 9.11.) In a few cases, we had not stated learning objectives very clearly (experts simply taught, without clarifying what the students were supposed to be able to decide and do) and, in a wider number of cases, we had not clearly linked assessment to the learning outcomes (implicit or explicit) for instruction. From a learning-engineering perspective, not everything in the checklist needed to be done right away—for example, personalized learning can be a quite expensive proposition—but having objective information about learning-design gaps helped to focus attention first on less expensive but high-impact issues.

The second way using this checklist helps is that the reviewers are repeatedly applying the checklist in new areas. As mentioned earlier, all the research about building mastery applies to our learning engineers as well. By applying the principles in unfamiliar learning areas, the Kaplan learning engineers further build and generalize their conceptual mastery of evidence-based approaches.
In addition to using a systematic review process to embed an iterative approach to evidence-based learning improvement, with technology-enhanced learning environments running at scale, we have the opportunity to set up systematic pilots of how evidence-based improvements can make a difference. (See Figure 9.12.) We mentioned an early example of a single randomized controlled trial (RCT) run within Kaplan Test Preparation. It is an entirely different matter to manage dozens of pilots at once.

Running many RCTs at once has really taken off within Kaplan University (KU), a virtual higher education institution in the United States with around 35,000 students, focused on serving an older (25-years-old and up) working age population.

**FIGURE 9.11** Key Kaplan products evaluated with checklist

![Checklist Diagram](image)

In addition to using a systematic review process to embed an iterative approach to evidence-based learning improvement, with technology-enhanced learning environments running at scale, we have the opportunity to set up systematic pilots of how evidence-based improvements can make a difference. (See Figure 9.12.) We mentioned an early example of a single randomized controlled trial (RCT) run within Kaplan Test Preparation. It is an entirely different matter to manage dozens of pilots at once.

Running many RCTs at once has really taken off within Kaplan University (KU), a virtual higher education institution in the United States with around 35,000 students, focused on serving an older (25-years-old and up) working age population.

**FIGURE 9.12** High volume technology-enhanced courses invite systematic controlled trials

<table>
<thead>
<tr>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For illustration purposes: Based on CM 107 (College Composition)
Each band represents 8 sections, each with 25 students (a total of 200 students)
population seeking to advance or change their careers (Brown & Kurzweil, 2016). A number of early college courses at KU have significant scale—500, 1000, or even more learners beginning every month in topics like mathematics or composition. With 1,000 learners beginning, say, a composition course in January, in theory we should be able to run five pilots of 200 learners each that month, and then repeat every month throughout the year. A single course at this scale should potentially deliver dozens of randomized controlled trials (RCTs) exploring improvements to learning performance. Kaplan is not yet operating at that scale, but KU has run more than a hundred RCTs in just a couple of years across a number of courses.

With this kind of scale, we can consider a multitude of different kinds of interventions to check from a number of different directions: assessment validity and reliability pilots, work on instructional methods, work on motivation methods, and more. (See Figure 9.13.)

However, if we want to run many (dozens) of pilots in parallel, we, like Kaplan, need to put in place management processes to keep track of multiple stages of multiple pilots—and to ensure decisions are well-made about whether to continue to pilot, to stop a pilot series, and/or to widely disseminate a piloted intervention. (See Figure 9.14.)

With all of this in place, we can begin to see the development of a typical innovation portfolio distribution of project success. (See Figure 9.15.) Most of Kaplan’s current efforts are inconclusive—ideas have promise, but have not yet landed either in the success or failure column. Many ideas remain in process, with a few either settled as not working or as appropriate to disseminate. Note that for learning-engineering purposes, it is just as valuable to know what not to disseminate (saving costs and teacher/student time), as it is to know what is worth disseminating.

**Conclusion**

Moving from case-by-case investigations of learning improvement to a systematic learning-engineering approach and culture at scale requires a range of different interventions to achieve. We need to provide systematic training at scale for the many individuals engaged in designing or supporting learning environments, as well as those preparing our learning environments for systematic evaluation and change over time. We have to move beyond the people directly involved in learning activities, the learning engineers, the teachers, the assessment professionals, to include the general managers (at least) who have to allocate capital, resources, and people to a wide variety of different activities in a learning organization. Ultimately, we need processes reinforced by the whole organization to evaluate evidence about what works and to make evidence-based changes that are piloted until they work.

If we actually do achieve this level of change in our learning organizations, it will make for dramatic changes in the impact of various roles within the organization on learner success. (See Table 9.1.)
With scale, we have the option to continue to improve further

<table>
<thead>
<tr>
<th>Adaptive</th>
<th>Assessment &amp; CLA</th>
<th>Learning Strategies</th>
<th>Media Principles</th>
<th>Motivation &amp; Self Efficacy</th>
<th>Social Norming</th>
<th>Worked Examples</th>
<th>Others</th>
<th>Open Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>SmartBook</td>
<td>CLA Re-Scoring</td>
<td>Self-Explanation</td>
<td>Advanced Organizers</td>
<td>Judgments of Learning</td>
<td>Badging</td>
<td>Early</td>
<td>Section Size</td>
<td>Challenge Exams</td>
</tr>
<tr>
<td>MyLab</td>
<td>Assessment</td>
<td>Self Summary</td>
<td>Coherence</td>
<td>Motivational Priming</td>
<td>Iconographs</td>
<td>Later</td>
<td>Orientation</td>
<td>Faculty Dashboard</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self Test</td>
<td>Continuity</td>
<td>Self Efficacy</td>
<td>Self Efficacy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Redundancy</td>
<td>Modality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Multimedia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 9.13** Variety of areas to explore through RCTs at KU
Running dozens of trials requires careful tracking

1 Study In Prep
Adaptive Learning (Mcgraw Hill Connect AB219)

3 Studies In Development
Attribution I (SOFI) Standing Out Fitting In (AB/MT140)
Indigo Live Seminar Delivery (AB/MT203)

19 Studies In Term
Adaptive Learning (McGraw-Hill Connect CM107)
Adaptive Learning (SOOMO PS124)
Attribution II (KU160)
CLA Anchor Paper (PA106)
Competency (CS113)
Campus LearnLab (CM220 Blended)
Mind Set Challenge (CM220)
PERTS Mind Set Challenge (MM150)
Science Center Evaluation Study (SC156)
AS Section Size (PS124)
Self Regulated Learning (CE100)
Indigo Live Seminar Delivery (GB512)
Stereotype Threat Surveys (CJ100, CJ101, MM150, MM207, MM212, MM255)

8 Studies In Assessment (5 Analysis in Progress; 2 Data Requested; 1 Awaiting Data Availability)
Attribution I (HS100 Pre-Intervention)
Harvard Study Supporter (CM107 and MM150)
CLA Blind Scoring (CJ101 and GB512)
Faculty Dashboard (SS310)
Social Norming (HU300, MM150, IT301)
Worked Examples (CJ130)

34 Studies Completed
Advanced Organizers (HN144)
CLA Reliability
CLA Assignment Order (CJ100)
CLA Training (CS204)
Orientation Facilitation
Bus Section Size
Seminar Delivery (LS)
Stereotype Threat Historical Review
Library Use Historical Review
Judgments of Learning (HS101)
V4 Learning Platform (IT133)
Math Center Feedback (MM207)
Motivational Priming (HS100 and PS124)
Self Test (HS200)
Adaptive Learning (SmartBook HS120)
VoiceThread (PA205 and PA201)
Worked Examples (14 instances)

FIGURE 9.14 Example of process used to track progress of pilot
At scale, we can look at the overall outcome of many pilots

Kaplan University Research Pipeline Focus and Progress (11/2015)

- Four key focus areas; Dozens of randomized control trials over past two years.

- Several early studies proved inconclusive – led to more structured pilot design process.

- Yielding several “go / no go” decisions based on evidence of improved outcomes.

FIGURE 9.15 Innovation portfolio at Kaplan University
<table>
<thead>
<tr>
<th>Role</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learners</td>
<td>• Focused on racing against schedules</td>
<td>• Focused on effort to reach serious goals</td>
</tr>
<tr>
<td></td>
<td>• Unsure how outcomes, instruction, and evidence-gathering tie to long-term success</td>
<td>• Confident that outcomes and evidence-gathering are meaningful for long-term success</td>
</tr>
<tr>
<td></td>
<td>• Unable to distinguish what a learner is doing wrong vs. what the environment is doing wrong—tendency to blame selves with no solution</td>
<td>• Are learning, themselves, what they need to do to be successful learners long-term</td>
</tr>
<tr>
<td>Families</td>
<td>• Find learner progress to be mysterious—unclear how it is going and what it is for</td>
<td>• Able to see real progress towards important goals over time</td>
</tr>
<tr>
<td></td>
<td>• Difficulty knowing best ways to help from home front—for specific learning outcomes, and for longer-term tie-ins to learner career goals</td>
<td>• Have well-defined roles that tie directly to the rest of the learning environment and to learners’ own goals for near- and longer-term</td>
</tr>
<tr>
<td>Teachers</td>
<td>• Only able to provide one form of instruction—unable to personalize well for entire class</td>
<td>• Are guided to individualize instruction (type and content) with trustworthy evidence</td>
</tr>
<tr>
<td></td>
<td>• Using methods that are traditional rather than evidence-based—how they learned, how they were trained vs. how learning works</td>
<td>• Become comfortable working with evidence and resulting suggested individualized instructional methods for each learner</td>
</tr>
<tr>
<td></td>
<td>• Often expected to do it all: generate the learning activity, the media to support it, the assessment to evaluate it and then to teach it, interpret outcomes, provide feedback—overwhelming</td>
<td>• Can focus on learners’ additional needs as they work through well-designed, well-targeted content</td>
</tr>
<tr>
<td>Role</td>
<td>From</td>
<td>To</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Teacher Managers/Principals/Local Administrators | • Often using evidence that may not be valid or reliable to judge outcomes that may not be well-linked to long-term learner success  
  • Aside from usability ("Can my teachers/learners do this?"), often have no systematic way to evaluate if a new approach really is an improvement or not  
  • May be unclear what professional development (PD) would be most helpful and which designs of PD are most impactful                                                                 | • Have trustworthy evidence about near- and long-term success for different teachers with different types of learners  
  • Able to meaningfully review impacts of changes to learning environments  
  • Can use evidence with confidence to identify areas where teaching staff can use PD and can look to see which PD makes systematic differences to learner outcomes                                                                                                                                                                                                                                           |
| Instructional Designers          | • Often have little/no valid/reliable information to highlight which activities are not leading to learner success—smile sheets are not enough  
  • Executing careful pilots takes too much effort and too much time to be useful  
  • Are not sure how to interpret evidence generated elsewhere for use with own learners                                                                                                                                                                                                                                                                                                                                                       | • Have rich evidence available about which parts of learning environment are working better or worse for which types of learners  
  • Are able to set up to pilot different interventions with good designs and good evidence, able to run rapidly at scale  
  • Can rely more clearly on evidence generated elsewhere, given information about types of learners                                                                                                                                                                                                                                                                                                                                 |
| Visual/Media Designers           | • Work based on tradition and own preferences for visual design  
  • Model often “if we get them to pay attention, they learn more, right?,” missing the issue of distraction  
  • Have little experience either reading/finding evidence-based work on learning, or in setting up pilots to compare different approaches for own learners                                                                                                                                                                                                                                                                                                         | • Have clearer guidance about what approaches are likely to work better (especially for hardest, most important outcomes)  
  • Can generate evidence to compare different media approaches (for different types of learners) that can be shared and compared with confidence                                                                                                                                                                                                                                                                                         |
<table>
<thead>
<tr>
<th>Role</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
</table>
| Subject Matter Experts      | • Often assume what they think to say (or how they remember being taught) is what learners need to know, missing the fact that experts’ internal monologue is already driven by a rich array of mastered outcomes, and that our experts now are the rare people who persevered to mastery with inefficient learning methods for most people  
  • Tend to work solo, and are asked to build out instruction, assessment, specific materials, and more—all without any background in how learning works for most folks  
  • Do not regularly receive evidence nor work with a team to identify where, exactly, learners are being tripped up on their way to master hard, important outcomes | • Engage in more effortful, but more productive, exploration of their expertise to inform learning environments  
  • Work closely with evidence-based learning engineers and others to suggest and develop best approaches for these outcomes  
  • Participate in iterative change processes to improve learning and to keep up with new developments                                                                                                                                                                                                                      |
| Managers of Instructional Designers/Course Production | • Are usually not able to guide priorities for effort in development based on evidence for where the hardest, most important parts are and which parts are not going well for learners  
  • In the absence of reliable evidence, often end up discussing with general managers only the cost side of production, without regard to outcomes | • Have good evidence to identify which outcomes are the hardest, most important across a portfolio of projects  
  • Able to engage in cost-benefit discussions with general managers about what it will take to solve business/learning problems with learning solutions and pilot these to determine real likelihood of success |
<table>
<thead>
<tr>
<th>Role</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
</table>
| Ed-Tech Developers          | • Find generating evidence of effectiveness to be enormously challenging—too slow, too expensive—to fit within purchasing/financing cycles  
• Are not rewarded by purchasers now for demonstrating and improving outcomes resulting in tendency to invest more on cosmetics of learning than efficacy | • Have lower barriers to piloting and generating high quality usability/cost/outcome evidence  
• See more demand tied directly to learning-performance outcomes, rewarding up-front focus on learning engineering and piloting against well-measured outcomes |
| Higher Level Administrators/  | • Cannot tell which courses/programs are actually delivering long-term learning value for learners  
• As a result, have to use other reasons to decide where to invest—personal interest or squeaky wheel  
• Mostly unable to reward faculty/designers with the most effective/efficient learning environments or push people towards these because they do not have the information to do so | • Good evidence indicates which learning environments are delivering near- and longer-term success for learners  
• Allows for more focused application of resources to improve/upgrade learning environments to match needs  
• Able to investigate benefits for different costs of learning investment for different types of learners and learning outcomes |
| Academic Managers/CLOs       |                                                                                                                                                                                                     |                                                                                                                                                                                                 |
| General Managers            | • Treat learning development as a black box—managed by headcount or cost—not based on outcomes over time  
• As a result, have difficulty trading off investments (e.g., IT, marketing, operations vs. learning investments)  
• Have very little understanding or exposure to ways to improve learning outcomes or even what those learning outcomes themselves really need to be | • Have a better (high level) understanding of learning engineering opportunities and tradeoffs  
• Allows them to make reasoned tradeoffs in solving various business problems with learning investments vs. other investments  
• Continuously monitor learning environment performance (and pilots) at a high level, to stay ahead of changes in definition of expertise and to ensure progress on problem areas |

Table 9.1 (Continued)
As we can see, an evidence-based culture of learning improvement at scale means many different roles need to think differently about how they use evidence and interact with learners (and teachers in some cases) to lift learning outcomes.

Especially with the advent of technology-enhanced learning capabilities (either on their own or to augment teacher-delivered instruction), we have a new and remarkable opportunity to much more systematically understand and deploy what works for learning at scale. Much of what will be most effective will require close attention to “for whom, on what?” kinds of questions, but the careful blending of what is known, now, about learning from learning science together with better learning-assessment practices can allow organizations at scale to become much more systematic (and acculturated) in their approach to improving learning outcomes for a wide variety of learners.

We should not underestimate the effort and processes required to make these changes at scale, but learners, teachers, developers, and more will benefit as we do so.

Acknowledgement

Special thanks to Kaplan, Inc. for allowing the use of materials and examples.

References


We are living in a very interesting time. Technological advances such as artificial intelligence (AI), the Internet of Things (IoT), advanced manufacturing, and robotics are impacting our lives at an unprecedented speed. Companies like Uber, Lyft, and Tesla are heralded as disrupters in the transportation industry. But they are also disrupting the very way in which people work and exchange as they revive the trust economy (Diekhöner, 2017), a practice as old as humanity itself.

We are witnessing a change in not only what work looks like, with rapidly changing types of jobs and skills required, but also a change in who (if anyone) is needed to do that work. For instance, when Tesla produces its first self-driving commercial truck, we will see a significant impact on the economy, with the number and type of driver jobs forever changed. Similarly, an anxious buzz fills the air when we talk about robotics. Admittedly, the evolution and the implications of robotics on the workforce are significant. Our focus is often on the impact of such automation on the manufacturing industry, but if we look at retail a similar picture emerges. Amazon is going to put major retailers out of business in the next five years. For all the talk about coal miners—and with all due respect for these hardworking people—there are only around 50,000 coal-mining jobs in the United States, whereas department stores have shed over 100,000 jobs in the last year (Thompson, 2017). (And even as I edit this, ToysRUs has just shut down.) What we are seeing is tectonic disruptions in industries across the United States. This will mean massive changes in the way people work and, as importantly, the way people are educated to prepare for that future, evolving economy.

Yet in academia, we seem stuck in an old world that, on the one hand, has instigated much of this technological and societal change, while, on the other hand, clings to a highly monolithic structure that is under duress. Academia is here to serve society, but society is changing—rapidly. The education industry
itself is in the throes of change, yet many traditional universities may not recognize it. In the frenetic loop of admission cycles, rising tuition, shrinking federal research dollars, and balancing budgets, universities can lose sight of the horizon line of change ahead.

Society is issuing a warning that academia must heed if we are to remain relevant to the education of the next generation. We must radically re-examine our tried and true assumptions about learning and teaching, return to and revive forgotten models of education, and create a new orchestra of instruction. Let us blend the best classical and modern instruments of learning and teaching for students and educators to pick up and compose their own scores, ones that are deeply inspiring and personal to them, and that speak to one of our most fundamentally shared human impulses: to learn.

In the evolving landscape of work, education bears a responsibility to prepare creative, nimble, and ethical problem-solvers and entrepreneurs who can adapt to and embrace continuous change. With these shifts, the role of formal education is very much changing. We must move away from passive learning, where we demand students simply regurgitate facts, and toward active learning, where students work with peers and experts on real-world problems, earning credentials based on their acquisition of skills and knowledge through hands-on practice, research, and collaboration. At MIT, we are investigating how and why we learn, using learning science to engineer the online and on-campus courses that we teach and the emerging digital learning technologies, like augmented and virtual reality, that we build.

The future of digital learning, whether for university or lifelong learning, will be increasingly modular, personalized, and embedded in real-world applicability, offering students and workers the opportunity to learn and apply knowledge as needs arise. Our brick and mortar K-12 schools, colleges, universities, and workplaces will become flexible sites for in-person experimentation and collaboration cutting across disciplines and generations. Schools will serve as local hubs for a global student body, faculty, surrounding communities, and industries working together on regional and global challenges. In my vision, online, residential, and blended learning form a new orchestra of education that inspires and incubates creativity and curiosity. But how will we get to this vision for the future of education? What do educators today need to consider as we design and engineer this future of learning?

The Pendulum Swing: Four Modes of Education

As we look to the future of education with these many questions, it is vital to know our past and to understand that our modern education system is only around 200 years old. Modern Western education, as we know it, started with universal education, credited to Napoleon’s mandate for state-directed public education. In the 19th century’s push of imperial progress, various European
educational approaches prepared people from the countryside to operate in new industrial factories, become soldiers in increasingly technical warfare, and help operate the colonial empires. Education provided enough knowledge and skills to be a good foot soldier or laborer but not so much capability in critical thinking that the people could question why they were following state orders. It is no coincidence that our traditional education system—what we call direct instruction—draws from this same factory ethos.

In direct instruction, teachers give lectures, hold recitations, and assess students’ assimilation of their presentations. Let us look at that word recitation—meaning to repeat. There is a history of imperial subservience embedded in that term—you recite, sticking to rote repetition rather than thinking. The professor (or general or factory supervisor) tells you something, and you recite—that is the foundation of the direct instruction method we use today. 19th-century American educational reformer Horace Mann is the one credited with bringing the direct instruction model to the United States—a system that remains in place in many schools and universities across the country.

Yet, in these direct instruction lecture classrooms, how much learning is happening? How are students grappling with complex ideas and debates? There may be some disputation in the classroom when a brave student challenges the teacher, but not too much, because how can you have real disputation and discussion in a one-to-many environment? This is particularly true today in crowded classrooms around the world, where rapidly growing youth populations and historic teacher shortages leave the teacher-to-student ratio woefully unbalanced.

You may be surprised to know that at the same time direct instruction was taking hold in Europe and then in America, the idea of simulated work, learning by doing, or project based learning as we call it today, began to emerge. The apprentice model of working alongside an expert existed from the beginning of formal education, if not the beginnings of humanity itself, but the idea of simulation, not in the real world, but through projects and labs, is relatively new. This only came to be in the second half of the 19th century both in Europe and America. Simulated or project based learning was pioneered at places like Rensselaer Polytechnic Institute (RPI) and the Massachusetts Institute of Technology (MIT) in the United States.

In fact, the founding principle of MIT was *mens et manus*, mind and hand. The idea was to bring the factory into the campus. If you go to the “new” MIT campus, which itself is about 100 years old, all the ceilings are very high, designed so that you could put industrial workshops inside the educational enterprise. When MIT was first established, it was seen as a bit of a vocational school. But it was at MIT, strangely, that the first high tech school emerged, when “high tech” was not yet in the common parlance as it is today. It was here that the idea of having a chemistry laboratory was invented—and so industrial chemistry was born as this lab experience generated a new generation of chemists with practical experience. Before that, aspiring chemists, like their other classmates, learned in
Latin. You recited in Latin, met the teacher’s approval, and then you declared victory over your education. The idea of actually doing and making as part of your education was then, and in some schools remains, a foreign idea.

While learning by doing through simulations and projects is a relatively recent innovation in modern education, long before the modern system we had on-the-job training and apprenticeships. Do you think Michelangelo ever attended a lecture? Was he assigned homework? Tests? Recitations? What about the Scottish inventor, engineer, and chemist James Watt? He apprenticed with his father in carpentry and, in London, in instrument making, repairing and making telescopes, barometers, and other scientific devices. On-the-job training is, in some ways, the most natural form of learning as apprenticeship. That is how humans learned prior to the modern education system. People trained in guilds; medieval craftsmen and merchants formed associations to train and pass on knowledge and skills. Apprenticeship even exists among other animal species. That is how a kitten learns from its mother how to hunt. First, the cat brings home a dead mouse to her young. And then the next time, the cat brings a not-quite-so-dead mouse. And then the third time, a pretty live mouse. The kitten gets what we today call scaffolding to learn how to deal with these mice. These cats knew about apprenticeship and scaffolding before we humans ever discovered the concept explicitly.

In fact, apprenticeships continue to survive. The German system today is highly apprentice-oriented. Today, as American industry struggles to adapt in this age of robotics and automation, we are looking to the German apprenticeship system as a model. Apprenticeship is very powerful. In professions where your life as a customer is at stake, you prefer people who have apprenticed. You would rather your surgeon went through some apprenticeship, in the form of a residency, rather than just learning everything theoretically. Pilots, like surgeons, also still use the apprenticeship system, training and shadowing master pilots until they are ready to take the controls themselves. The modern aviation industry would not be what it is today were it not for flight simulators (quite literally simulated learning). The reason for this is that it is simply too expensive to exclusively train pilots in the sky. Pilot training is scaffolded on cyberphysical flight simulators. After these initial simulated trainings, the apprentice sits with the expert to learn the advanced tricks. That is how the famous pilot Sully Sullenberger learned that if your plane gets hit by a bird and loses an engine, there are strategies to use in response.

Apprenticeship, simulation, and direct instruction: these three modes of education are fairly familiar to many of us. But there is one other system that we do not give enough credence to: mutual instruction. Lev Vygotsky, the great Russian philosopher and psychologist, wrote much about the Zone of Proximal Development. The idea is that you gain the most if you learn from someone who is close to you in skill but just a little bit better. Mutual instruction, or peer-to-peer learning as it is also called, has a long history. When the British went to India, the Englishman Andrew Bell noticed that, in Indian schools, young kids
would learn from slightly older kids through mutual instruction. Bell observed that this method turned out to be very successful; he called this the Madras System and brought it back to England, where Joseph Lancaster adapted it as the Lancaster System. It was a valuable model because it was efficient during the late industrial revolution, as the number of young people enrolled in schools began to skyrocket and the number of trained teachers was low. Mutual instruction as an educational approach survived for many years, and then it gradually faded away as direct instruction grew in vogue, much like apprenticeships eroded.

A modern revival of mutual instruction is in the Ecole 42 in Paris, France. Ecole 42 was set up by the French entrepreneur Xavier Niel in 2013. It was named “42” after *The Hitchhiker’s Guide to the Galaxy* (1979), so you can see the school is quirky to begin with. The idea of the school, which is dedicated to computer programming education, is built on the core principles of gamifying learning and mutual instruction. There is no standard curriculum. There is no teacher-led instruction and no professors in the traditional sense of the “sage on stage.” The whole university is run by around 37 people, including the janitorial staff. The “professors” write computer programs that create an educational experience like a treasure hunt for students to navigate and explore. The students at Ecole 42 learn by navigating resources like MIT OpenCourseWare and other online material such as Open Classrooms. The students also learn by asking each other: mutual instruction. All they have to do is accomplish certain technical goals. The process is gamified: with each step and skill students acquire, the program unlocks a path. When students finish the whole program, they earn a credential. The first year of the program is on campus, and the rest students can do anytime, from anywhere. There is no degree, and there is no tuition. That’s mutual instruction, that’s learning by doing, that’s disrupting education. Though the question remains: how deep is the learning?

If we look at trends in education today, we are bouncing around among these four extremes: direct instruction, project-based simulated learning, apprenticeship, and mutual peer-to-peer instruction. The pendulums keep swinging back and forth, with teachers and students alike often exhausted by the churn of education fads and fashions. After a few years of heading up digital learning at MIT, I began to realize that many fads, trends, and directions are in response to a central problem, which is that direct instruction is so addictively convenient and so baked into our education system that it has become very hard for us to change it. Just confronting it, we veer back and forth between resolve and resignation.

**A Richer Orchestra**

As we look at our education system today, we see two dominant tracks, like a two-instrument orchestra. One instrument is residential. You get into college to live and study residentially. In medieval days, this used to be the monastery, and now it is a university. This is the residential model. The other instrument in our
educational orchestra is the real world (see Figure 10.1). While many who are reading this will have gone to college, there are many more people who have not. They learn through working in the real world. They go to work for a farmer or welder, and they learn how to become farmers and welders—the apprenticeship model.

Our traditional education system is very rigid across these two worlds. If you are not blessed with the resources to go to college, you might finish high school—though in many places in the world, you do not—and then you go to work. If you go to college, the track is pretty much set. You go to school for two to four years and earn an Associate’s or Bachelor’s degree. You work for some time, then maybe go back and get an MBA or other graduate degree, and then you go back to work. Or maybe you stay in academia, which, of course, is a form of work. But, that is the limit—these two instruments are all we have had in terms of education today. However, today there is a third instrument in the orchestra: online. Now, suddenly, this one new thing starts changing the degrees of freedom available to us. And we can start composing new music, ending up with a more complex orchestra.

Online provides an opportunity to blend and mix those two modes of education. Let us look at residential learning. For instance, the flipped classroom is trying to inject more simulated and apprentice-based learning and is using online to take the direct instruction out of previous in-person time. Moreover, when online enters the picture, you not only can flip the classroom but also can bring the classroom to students away from campus. When students go away from campus to do internships or to study abroad, they could also get credit for courses they can take online simultaneously. The internship becomes not a semester delayed—it is a semester enriched.

Once students graduate and earn that degree, they go to work. But does going to work now mean that you must give up on your education? This whole one-and-done attitude to education is passé. In the last three years, money, the most regulated industry on the planet, has been seriously disrupted by Bitcoin. But how many of us really understand Bitcoin? We need continual learning to keep pace with these rapidly changing industries and technologies. We are on a treadmill, and the treadmill is gaining speed. Not learning is no longer an option.

**Existing Pathways**

![FIGURE 10.1 Existing educational pathways. Is this the limit to our design language?](image-url)
When you are working, you need to be learning in parallel. If you are on that treadmill, you cannot just stay stationary, otherwise you can forget about advancing your career. And, at some point, maybe you fall back and you realize, “Oh my God, I need to go back to college.” And even then, as you are getting another degree, you may realize that you need to add knowledge in tangential fields such as blockchain, or machine learning, or genetics, or self-driving cars, or autonomy, or AI, or IoT, or digital music, or new payment technologies. Your learning must go on and on.

While many of these innovations like blockchain and AI seem unreal, they will cause real changes in real industries. For example, for a computer scientist quantum computing is going to change the game. If you are a computer scientist today and you do not know what is happening in quantum computing, you have to be very careful about the future of your career. This will always continue: the reality is that you are going to have to learn new areas outside of your main field of expertise, both while you are in formal education and while in work. We can offer more opportunities to do that learning online, or to dip back into college again, maybe for a week to get a refresher, and so on.

Returning to my music metaphor, does this start to look like music? Think about the state of music several hundred years ago. Today, we have a rich diversity of musical traditions and genres, in large part because of human creativity in inventing new instruments to produce new sounds and combinations. Yet we are not showing the same level of guts in education today. We are still stuck in the types of rigid educational systems invented hundreds of years ago. If music has reinvented itself so many times, it is about time the academic industry began to reinvent itself. At our office, the Office of Open Learning at MIT, we are trying to spark that reinvention for MIT and to make our learnings available to the world.

**MIT Open Learning**

The roots of the Office of Open Learning go back about 18 years. I was involved in the creation of OpenCourseWare in the early 2000s, when MIT decided to give its entire curriculum away to the world for free. Think about this: in its history, OpenCourseWare has had in the range of 200 million unique people come and download courses. But how many users does Facebook have? Over 1.8 billion. Here is the difference: on Facebook, you say, “Here’s a picture of my cat.” On OpenCourseWare, you download a course on linguistics by Noam Chomsky or condensed matter physics by Wolfgang Ketterle, the Nobel Laureate.

In 2012, we launched MITx. We launched the first online course from MIT at around the same time that Stanford launched courses on Udacity and Coursera. We launched our first course with Anant Agarwal, a brilliant and forward-thinking professor at MIT. He developed a massive open online course (MOOC), Introduction to Electrical Engineering, essentially on circuits and systems. MIT announced in December 2011 that they were launching a new pilot
online learning program based on this one course. Agarwal had nothing at the
time—no code, no digital platform, none of the online content. They just made
the announcement and then, over the course of the next three months, Agarwal
and his team built the software and content, launching MIT’s first MOOC in
April of 2012. I remember running into Agarwal and asking, “How’s it going,
man? How many students do you expect?” He said, “Listen, I hope no fewer
than 5,000 because that will be embarrassing. But I hope no more than 10,000
because I will not—I will just die.” He ended up with 155,000 students. That is
how MITx became big.

We then went to our dear friends down Massachusetts Avenue at Harvard,
who were thinking in the same direction, and that is how Harvard and MIT
launched edX. Agarwal is on leave from MIT as he now runs edX. It is a won-
derful story of academic creativity and gumption. edX has now reached about 14
million unique registered users with their courses. MIT content has reached about
4 million people; we have issued about 140,000 certificates and developed about
150 courses.

These digital learning technologies present an incredible platform for bringing
together diverse participatory communities of learners to tackle our world’s
greatest challenges. In our MITx and edX MOOCs, we find that success in
online learning is in part facilitated by the learners’ communities. New spaces
need to be built within course platforms for communities of learners to work
together to develop collective knowledges, build bridges between global divides,
and apply their course work to real-world application.

Online also enriches students’ time on campus. Everything we do with
MOOCs and OpenCourseWare, we bring back to blended learning to flip or
activate the classrooms on campus. We are making our courses available to stu-
dents while they are doing internships, so they can now earn credit while away
from campus. We have also introduced new credentials like the MicroMasters,
which is half of a Master’s degree earned online. We took a full two-semester
professional Master’s degree (which costs a pretty penny) and made the first half
fully online at an accessible cost. We invite people around the world to take that
first semester—this is an opportunity open to anyone. There is no admission—
just take the courses and see how you do.

If you finish, and you pass the proctored exams, MIT will award you this new
credential we invented called an MITx MicroMasters. This credential is afford-
able, costing far less than the standard ticket price of a residential Master’s degree.
Such new forms of credentials are increasingly recognized by industries and
companies who value lifelong learning. Importantly, if learners finish the Micro-
Master’s credential, they can then apply to MIT and join our residential Master’s
program. It is great for us from the admissions side, because now we know how
good these students really are. It is not just a GPA from a school we have never
heard of or a recommendation from a professor we will never meet. We have
your record of success through your performance online.
If you get into MIT through this program, we give you credit for the semester completed. And then you can finish the full Master’s degree in another semester. This MicroMasters is a way for us to tap into the workforce and make education more affordable. And, at the same time, for us, it is a win-win. We get to pick really great candidates from this process. The first MicroMasters cohort of 39 students arrived on campus earlier this year from the Supply Chain Management (SCM) MicroMasters. We found that these students are both more international and more seasoned, with more work experience than our typical SCM Master’s students.

Once MIT launched the first MicroMasters, two things happened. One, other schools came to us and asked to launch MicroMasters at their institutions. We agreed, and through edX, other universities started launching MicroMasters. Two, other schools came to us and asked if they could admit the MITx MicroMasters learners not admitted to MIT. We agreed to this as well. With the MicroMasters, edX has now reached 400,000 people, with MIT reaching 200,000 of those enrolled users. Now, there are 25 universities on four continents, and there are 40 MicroMasters and counting. The concept has really taken off. For an industry that is so glacial, this is pretty fast: two years. But, in some ways, the ultimate success story is that, through the MicroMasters, we found at least 39 students around the world who have now come to MIT, people who have been admitted whom we may never have seen without this new system. This is a new way of discovering talented people. We call this inverted admission, and it is one of the great breakthroughs online technology is enabling for education at MIT.

Another new program we introduced was MIT bootcamps. We asked, “What if people came to MIT or to an MIT-run event somewhere in the world for just a week?” We piloted a bootcamp program focused on innovation and entrepreneurship. It turned out to be really successful. We have people take MITx MOOCs on entrepreneurship and, once they finish the MOOC, they apply to this intensive, one-week program. A small percentage gets selected to come to a bootcamp. We have found real impact in these short programs, from participant testimonials to the startups that have come out of the program. In fact, bootcamp alums have raised about $70 million of investment funding for ventures that came out of the program. We have run about a dozen bootcamps in locations around the world in Australia, Turkey, Taiwan, and Brazil; most have focused on entrepreneurship, though the catalogue is growing to include IoT, blockchain, the future of food, and sustainability. These blended bootcamps are yet another example of the new affordances allowed by online as we convene learners and leaders around the globe. With programs like these hybrid MIT bootcamps, this new orchestra is more flexible and dynamic (see Figure 10.2).

Everything we do online, we also apply on campus, using it to flip our classrooms and to make the learning experience more engaging. MIT has a long history of getting away from lectures by designing physical studio spaces like, for example, the Technology Enabled Active Learning (TEAL) pedagogy and classroom space developed out of the MIT Physics Department. Still, in the
early days of digital courseware, I had to convince all my colleagues to participate. But today, we have about 100 residential classes that use these digital tools. Professors love it because of the new space and time opened up in their classrooms for discussion and hands-on labs and activities. Students love it because through online tools like rapid assessments and adaptive hinting, they get very quick feedback on their work. Students do not have to wait three weeks for the assignment to come back in order for them to understand that they did something wrong. Everything is live.

Our office, the Office of Open Learning, is composed of the Office of Digital Learning (ODL), where we do all the MITx MOOCs; OpenCourseWare development; residential experiments; and xPro, our professional education unit. We also have the MIT Integrated Learning Initiative (MITili) that looks at the science and economics of learning. My colleagues and I at MIT authored a report called the Online Education Policy Initiative Report (Wilcox, Sarma, & Lippel, 2016), which looks at this landscape of learning science, educational tradition, and digital learning.

In summary, many practices we take for granted in education are not necessarily compatible with the science of learning. Online flipped classrooms, project based learning—they all give us opportunities to better align our practices with the science of the learning brain. The MIT Integrated Learning Initiative is looking at everything from how dyslexia occurs to ADHD to the policies and inequalities that impact school choice. What is the science behind learning? How does spacing between letters and spacing between sentences impact learning? Does mindfulness impact learning? How do we make learning more effective online, in our classrooms, and in our schools?

Along these same lines, we are increasingly turning to learning engineering, the iterative design process of implementation and testing discussed in this book, as we incorporate learning science into the design of our online and in-person learning experiences. At MIT, we are nurturing a new profession of learning engineers out of our Digital Learning Lab, a cohort of postdoctoral students who serve as key liaisons between faculty experts, instructional designers, and digital course teams. Their emerging role in the intersections of research, application, and pedagogy is increasingly vital to the future of education.
Designing this New Instrument

If online is a new musical instrument for education, what should it look like? How should it be designed? At MIT, we are turning to learning science and learning engineering to fundamentally re-examine our assumptions about how we design the curriculum, pedagogy, and tools that make up education. The concept of the lecture has not fundamentally changed in 700–800 years, but what has changed is that we have turned it into a blunt weapon for everything. We sit people down in lecture halls and go at it for 90 minutes and declare victory; that is really what the education complex has become. We need to start rethinking this because there is a fundamental assumption here that is wrong, particularly as we look at learning through the lens of cognitive psychology. In this approach to education, the assumption is the professor holds the pen, and the student is a mere sheet of paper. And all the professor has to do is write on that sheet of paper.

But a better way to think about it is the student is creating a model of the world—sort of like a plant growing. You have to feed that model as and when the student wants it. The professor is there to give the plant potassium when it needs it, nitrates when it wants them, water when it wants it, sunlight when it wants it. But when it wants it—that is a very different model than what we have currently. It is so inconvenient that we just ignore it. The fact is that the delivery model we have based our education system on is misguided. Anyone who is a parent knows this because we are trained as human beings to teach and to learn—and parenting is a fundamental aspect of that. Learning is the one thing we are hardwired for. Science actually explains and shows that our instincts about learning as parents are closer to the truth than the prevalent approaches to education that we take for granted in a classroom.

Do Not Tire the Prefrontal Leprechaun

What makes learning difficult? Think of a treasury as a wonderful palace of knowledge, a tree of knowledge, a library of knowledge that we want to construct. The problem is information comes to you in nuggets—a few coins here and there. Now here is the difficulty: the challenge in between is our digestive system for information that, unfortunately, is not a sheet of paper that the professor can write on. To use another analogy, learning is like having a leprechaun in your brain who is a combination of (among other things) your prefrontal cortex and your hippocampus. This leprechaun can only pick up so many coins of knowledge at a time. And what he has to do is run back and forth, smelt, reshape this knowledge, put it away, and organize it. When understanding crumbles, he has to bring that knowledge back up (sort of like a shelf collapsing), reconcile misunderstandings, and deal with the things that do not fit together. And, by the way, he must also eat, sleep, and do other things. This little leprechaun gets tired in how long? What do you think your attention span is? Be
honest: eight to 14 minutes? This is the problem that we face in delivering educational content, whether in person or online.

So, what happens to this little leprechaun in the prefrontal cortex when it gets tired? It takes a break. There is actually a technical term for it: mind wandering. You can actually see mind wandering in a functional MRI (fMRI).\textsuperscript{5} Mind wandering is the state in which the brain seems to bounce around in something called default mode. The leprechaun is actually doing all the other things you do not give it time to do—connecting the dots. Mind wandering is very good for spontaneity and creativity, which we do not particularly encourage in classrooms, by the way. We kill and decimate creativity. Because, after all, we were trying to create soldiers to go fight the war, not question the general. Uncontrolled mind wandering, however, also seems correlated to depression (Smallwood & Andrews-Hanna, 2013). Mindfulness seems to address this somewhat (Mrazek, Smallwood, & Schooler, 2012).

Cognitive load theory might explain an aspect of the stamina of this little mental leprechaun—or the cognitive load that the brain can sustain. This leads to all sorts of principles for designing instruction: for example, when to use fill-in-the-blank problems (for novices) and when to use more open-ended problems (for experts) (Kalyuga, Ayres, Chandler, & Sweller, 2003). This also informs multimedia instructional design. For example, when you have a graphic, does it help to play music? Does it help to add audible narration (Mayer, 2005)? These are the kinds of questions we are testing as our faculty and Digital Learning Lab fellows design digital course experiences and learning tools.

**Short Attention Spans are Human**

All sorts of principles come out of this work in fMRI imaging: principles like mind wandering and retrieval learning that need to be integrated into our educational design. For instance, make your lectures short is a principle many of us often violate—present company included. Why? Because it takes us 10 minutes to get everyone to the room, so I could not just give you a 10-minute lecture. But you can do that segmentation online. That is the beauty and convenience of online delivery. Take, for example, a paper by Guo, Kim, and Rubin (2014) on student engagement watching videos from edX. They looked at the length of the video and how much people watched. It turns out that on average, learners will watch up to about six to nine minutes. If you increase the length of the video, people will watch less. If you go any further than ten minutes, not only will they stop watching, but they will also give up much earlier.

Formal education might declare them inattentive. Yet, reading is a very recent skill in evolutionary terms. It is a recent imposition on the human brain. With their short attention spans, people are simply being (gasp) human. Paying attention beyond a few minutes is actually a skill humans are not very good at. Do not tempt mind wandering—make your lectures short, which can be easily achieved online.
Here is a second tip. At the end of a 5–10 minute video segment, you will learn something. What is the one thing you ought to do at that point? Let me give you some options: play some music? Ask yourself questions about what you learned? Re-watch the video? What do you think? Ask questions: this activates what is called retrieval learning, or the testing effect. The learning gains from the testing effect compound, like interest, because the more you learn and must retrieve that knowledge, the more knowledge adds up. One of the other things we must combat is the illusion of learning. If a student rereads material, they gain familiarity and they think they have learned. But actually, they do not retain that knowledge in the long term. However, if you test them, they suddenly feel like they did not learn as well in that immediate moment, but actually, their long-term retention is better. We can do this kind of retrieval learning in a classroom, but we choose not to. However, online we can quiz learners just as they are on the brink of forgetting, making them retrieve that knowledge from their short term memory and encode it into their longer-term storage.

This leads me to the great German psychologist Herman Ebbinghaus, a pioneer in the study of memory. In a self-designed experiment, he memorized nonsense words he had made up, and he tested himself every day on these words, and he measured how quickly he forgot them. His work resulted in what we now call the Ebbinghaus Curve, or the forgetting curve. We forget things over time. The point being is that you will forget. That is why you should not wag a finger at a child and say, “I can’t believe you forgot that.” Because, trust me, you are forgetting too.

Perhaps the reason you forget is that the brain is naturally cautious and frugal in its deployment of neurons. When you first form a memory, the brain forms a chemical. The memory is created by chemical connection, a neurotransmitter connection between the neurons. Over time, the chemical basically dries off and you lose memory. However, if you remind the brain of the knowledge after some time, just when it is getting rusty, then the brain actually drops a second physical connection. It creates a second synapse, then a third synapse. That is what forms long-term memory. The best way to learn is to learn once, wait until you are about to forget, then learn again, and then wait until you are about to forget again, and then learn it again. Over time, that knowledge becomes permanent memory.

I would recommend cascading out questions. Let us say you have 50 problems in your back pocket that you can give to your students as a teacher. Do not give all 50 problems that day. Give five that day and then work one of those problems in a week later. And then work five of them in a month later. That is when you get the Ebbinghaus effect. The language app Duolingo does this really well, and we are building these principles into our online courses. Do we ever do this in a physical classroom? No, because it is inconvenient.

With the arrival of the IoT, things change even more. For example, can the Alexa or the Echo play a part in learning? Consider these devices with Quizlet,
which is actually the product of an MIT dropout: Andrew Sutherland. Sutherland was at MIT when Quizlet surpassed *The Wall Street Journal* in the number of visits per day. Imagine Quizlet on your Amazon Echo. Your child comes down for breakfast and Echo engages them in a little bit of back and forth conversation, reminding them of what they learned the night before, triggering Ebbinghaus forgetting and recovery. Augmented and virtual reality technologies offer even more instruments for the future of learning, a frontier MIT is actively exploring.

**Spark Curiosity**

There is a circuit that is activated by neurotransmitters called dopamine. One of my colleagues John Gabrieli, the Director of the MIT Integrated Learning Initiative, found that if he could get the dopamine circuit activated and a certain set of brainwaves aroused, learning was significantly better. But he could not explain how to get the circuit activated. Meanwhile, Charan Ranganath and his colleagues from UC Davis discovered that the state most associated with this dopamine circuit is a familiar English word that we associate with dopamine flow. That English word is *curiosity*. So, if you make someone curious, they are more likely to learn better (Gruber, Gelman, & Ranganath, 2014).

We know this as parents as we try to spark curiosity in our children. As Plutarch said, “A mind is a fire to be kindled, not a vessel to be filled.” Plutarch got it right, yet how often as educators are we focused on sparking that creativity. Again, our instincts are sometimes better than some of the historical educational traditions we are taking for granted. With all the online content out there, students can get the content without the teacher. The one thing that the teacher can really do that no one else can—and the one thing I remember my best teachers for—is to inspire their students, get the dopamine circuit going, sparking curiosity and learning.

**The Future of Education**

Several years ago, I had the honor to be involved in the development of a new university in Singapore: the Singapore University of Technology and Design (SUTD). We decided that we would throw the rulebook out and start again. In the design of this new university, we basically reinvented quite a bit based on experience from MIT and elsewhere, reimagining everything from lectures and assignments to university architecture. At SUTD, we have lectures once every two or three weeks. Most of the classes have labs and projects, especially design projects. Most of the curriculum is highly scaffolded. But, in the intervening years since we launched SUTD, the landscape has continued to dramatically change.

If you were to design a university or school today, how would you go about it? MIT is looking at that very question with the launch of the Abdul Latif Jameel World Education Lab (J-WEL). J-WEL looks at the application of learning
principles to everything from pre-kindergarten to grade 12, to higher education, and all the way to the workplace, working with schools, universities, and companies around the world. It combines the best of learning science and engineering together to develop and share tested best practices in education from across the globe, building on a long history at MIT of education capacity building.\textsuperscript{8}

The great writer Roy Amara noted, “We tend to overestimate the impact of technology in the short term and underestimate it in the long term.” Education change is upon us. We are only at the beginning of the beginning—to misquote Churchill, who said the end of the beginning. We are in a very interesting and musical time. For me, the future of learning is blended, individuated, fluid, hands-on, and responsively designed. Change is coming, and it sounds like music to my ears.

Notes
1 A quick history of MIT’s campus: In its founding in 1861, MIT was originally located in Boston’s Back Bay. In 1916, MIT moved across the river to the more industrial landscape of Cambridge where there was more space to expand the campus and accommodate high-ceiling, open-floor labs and workspaces.
2 Even the word technology, as we know it today, was only a few years old when MIT was established.
3 Watt did spend time at the University of Glasgow, but he was a technician there, not a student.
4 In 2016, Ecole 42 opened up a second school in Fremont, California.
5 MIT now has four MicroMasters in various stages of launch. The second one was in economics and development policy. The third one is principles of manufacturing, and the fourth is data science.
6 One of the innovations that has really made the field so exciting is the evolution of new imaging technologies such as functional MRI (fMRI). fMRI looks at where the oxygen is going. It can tell which part of the brain is processing at different points in time. It is sort of like, if you could see which part of a computer-processing chip was hot when you were doing different things, you could tell whether it is the CPU or the memory. With fMRI, we are learning some amazing things about the brain.
7 I highly recommend the book Make It Stick: The Science of Successful Learning (2014) about retrieval by Peter Brown, Henry Roediger, and Mark McDaniel.
8 MIT has a long history of designing other universities. I am actually the graduate of a university that MIT helped develop in the 1960s called the Indian Institute of Technology, Kanpur.

Acknowledgements

This chapter is adapted from a talk delivered at the LearnLaunch conference in 2018. Dr. Lana Cook was instrumental in editing this piece.

References


This page intentionally left blank
PART III
Conclusions
This page intentionally left blank
In this concluding section, we first discuss crosscutting themes through the chapters and synthesize the current state of learning engineering. Next, we propose an overarching framework for contextualizing learning engineering—the Digital Teaching Platform—as an emerging model for an infrastructure that can realize learning engineering’s full benefits. Finally, we advance a research agenda to help realize this aspirational vision in online higher education.

The Current State of Learning Engineering

As the workplace rapidly evolves, we have to do more to ensure that people throughout their lives can reliably and effectively shift their skills. Because people are increasingly working with information-rich tools of various kinds (e.g., AI, database-informed decision-support, robotics), rapid technological shifts will drive alterations in human skills because of their complementarity with the digital infrastructure. In turn, the number of sophisticated jobs that are satisfying and financially rewarding will increase.

It is still early in the development of learning-engineering approaches. Very little current work in this field makes use of empirical results on how education can increase the reliability and effectiveness of online and blended learning at scale. Also, it is still rare for learning science researchers to collaborate on problems and data with practitioners in the field. Often, too little attention is paid to valid and reliable methods for gathering evidence about learning and motivation for learning that are required by iterative improvement and the personalization of learning environments.

Indeed, it has to be said that our understanding of how learning actually works is still piecemeal. Often, we do not know when results can generalize or when they are applicable only to particular contexts, and we struggle to identify the
learning outcomes and evidence gathering relevant for long-term student success. Getting these things right is especially important as we increasingly personalize learning environments to develop those cognitive and motivational assets valued in the real world.

Personalization is likely to be much more important for generating successful, learning environments at scale than it was early on in an area like medicine. Medicine benefitted materially by big changes in health driven by uniform application of treatments to specific, uniformly identifiable disease states. This helped drive investment in more empirical work on population averages to find additional bulk treatment effects. (Now, there is increasing awareness that population effects even in medicine are hiding substantial subgroup differences—personalized, “precision” medicine is coming.) For learning, however, there are unlikely to be highly successful bulk treatments for minds—how individual minds differ when they come into a learning environment is much more likely to require customized interventions.

The good news is that there is a range of learning science available to draw on to improve our educational environments, to make them more and more effective and efficient over time; the chapters in this book provide examples and evidence. What we need, and are only just starting to develop, are iterative development platforms and approaches grounded in learning science and linked with detailed data about the learner, including success measures across multiple time-scales. This means going beyond one-shot assessments at the end of a module or course to also build evidence over time that the new skills and approaches acquired are applied in later situations, in school and beyond.

To do this at scale, we need to prepare a range of professionals (such as instructional designers, teachers, purchasing agents, managers, regulators) to work in new ways. This will be its own at-scale learning project: we need double or even triple-vision as we develop interventions to recognize not only that the learner’s mind has to be handled in new ways but also that all the other professionals involved need training, with sufficient practice and feedback, to change their practices. This is another kind of learning-engineering opportunity that is still in its infancy yet will be crucial for at-scale success.

Building new ways at scale for researchers and learning engineers to collaborate is an example of the kind of professional change needed. To maximize the iterative opportunities of technology-delivered instruction, we need to develop fluency around bi-directional flows of problems and information between these two groups. This means connecting research-lab efficacy with at-scale effective product development through the crucible of usability, while satisfying fidelity of implementation measured by at-scale, well-designed learning, motivation, and behavior measures with frequent iterations.

This kind of collaboration can provide new sources of value beyond traditional research publications. It is just as valuable to find things that do not work with these approaches as it is to find things that do: there are massive savings in cost.
and effort by avoiding things that are ineffective. Similarly, it is very valuable to find more efficient ways to reach the same levels of mastery—engineering is about effectiveness and efficiency, within the constraints of real-world delivery, to free up resources to do even more to help learners.

As can be seen from this book’s chapters, there are examples already of this kind of collaboration happening—a wide variety of investigations and collaborations are beginning. What is needed now is for these efforts, and the value they provide to learners and the delivery system, to become more visible, more frequent, and more in demand. The opportunity for researchers to gain through experiments at scale, allowing them to more deeply understand the influence of learner context, is largely unrealized, as is the opportunity for delivery environments to provide measurably better learning at scale, rather than guessing.

Ultimately, as this promising approach to learning engineering gains traction, learners, their families, their employers, and society at large will be the winners. Instead of reactive floundering as technology and information accelerate change, people can develop real pride and experience around their continuing potential to reinvent themselves and send the same message of hope and progress to next generations as well.

To realize the full potential of this aspirational vision, learning engineering should be contextualized in an educational infrastructure designed to maximize the power of online and blended learning. The next section describes a framework for such an infrastructure.

The DTP Model

The Digital Teaching Platform (DTP) is a model originally designed to provide a framework to understand the dynamics of a K-12, teacher-led classroom. In this model, the digital environment is the major carrier of curriculum content and functions as the primary instructional environment in today’s technology-intensive classrooms (cf. Dede & Richards, 2012; Richards, 2017). We propose that a modified version of this model can guide the evolution of an infrastructure to realize the full benefits of learning engineering across all types of online and blended learning and to address the challenges laid out in the first section. Our discussion uses the DTP model as a framework to identify some of the important research areas for learning engineering as applied to online courses at the higher education level.

The DTP model has two feedback loops (see Figure 11.1). The main classroom loop displays the dynamic between the instructor, students, and curriculum. In its initial conception, this is a continuous feedback system as the teacher manages a K-12 classroom, facilitates learning, and monitors student progress. The data generated here (including baseline information about learners and teachers as well as evidence about cognitive and non-cognitive aspects of learner performance) can feed a variety of iterative processes discussed above.
The inner, technology-intensive, personalized practice loop is individualized dynamically for each student. This loop is where the personalization goals talked about above are addressed, and it can also feed iterative improvement processes.

As argued in Dede and Richards (2012, pp.1–2), a full-fledged DTP addresses three major requirements of contemporary classrooms:

1. A DTP models a completely realized, networked digital environment that includes interactive interfaces for both teachers and students. Teachers use the administrative tools of this digital environment to create lessons and assignments for students and to manage and evaluate the work the students return, including evaluations of non-cognitive aspects of the learner experience. A DTP includes specific tools for assessment: creating tests, assigning them to students, and reviewing the results, ideally both immediately and longitudinally (and, again, including non-cognitive aspects of learner progress). The teacher tools also provide timely reports on student progress or their remedial needs. The administrative tools for students allow them to complete assignments and assessments. More importantly, these tools allow for both individual and group work. Some students can work independently on individualized assignments, while others can work collaboratively on shared assignments.

2. A DTP provides the content of the curriculum and assessments for teaching and learning in digital form. This content includes all of the information in the curriculum—the instruction, the exercises, and the assessments—targeting specific cognitive development as well as non-cognitive progress. The content also includes interactive elements: manipulatives, interpersonal, role-playing, and other non-digital activities; special-purpose applications; multimedia materials; and so on.
3. A DTP supports real-time, teacher-directed interaction in the K-12 classroom. It includes special tools for managing classroom activity; for monitoring progress on assignments; for displaying student work, demonstrations, and challenges on interactive displays; for managing group discussions; and for coordinating all large-group and small-group activities with desired progress on cognitive and non-cognitive capabilities.

There are some learning environments that are not far from this vision already (e.g., the work of Reasoning Mind or Carnegie Learning), so these characteristics are not impossible to achieve. All of these aspects of a DTP are designed to function effectively in the give-and-take atmosphere of interactive instruction. The teacher can shift quickly from large group demonstrations to small group activities to individualized practice and assessment. Students move seamlessly from using their devices for these activities to closing their computers and participating in discussions, both peer-to-peer and teacher-led. The teacher is fully in control of student activities by making assignments, mentoring individuals, and leading discussions. In DTPs, the pedagogy of the curriculum is designed using principles of evidence-based, guided, social constructivism as a theory of learning for both cognitive and non-cognitive progress, and the system provides the support for a transformation of teaching and learning (cf. Dede, 2008).

**Online and Hybrid Courses**

There are a variety of open questions raised by the chapters in this book. We propose that there be controlled experimentation guided by an infrastructural model for online teaching and learning—a DTP 2.0. How does this differ from the description of a K-12 DTP above?

**Distance Versus Face-to-Face Learning**

Many of the assumptions of a technology-intensive, face-to-face environment are also the assumptions of online courses. The digital environment is the major carrier of curriculum content and functions as the primary instructional environment. The major difference is the nature of the feedback loop between students and teacher, identified as the third requirement above (as well as the more constrained nature of peer-to-peer interactions for the students). In a face-to-face environment, the teacher-student interaction is direct, and feedback to an observant teacher is immediate. Note the larger, outer loop (in Figure 11.2). In this figure, the part of the loop identified by “A” needs to be created. How can learning engineering help in closing this loop? How does the student provide feedback to the teacher? How does the teacher notice that the online class is confused, bored, or intrigued? There are early explorations of detecting student attention through analyses of student interaction data (e.g., Ryan Baker’s affect detector (cf. Baker, Gowda, & Corbett,
2011), but there are many other ways under investigation (and many underlying theories of affect) including machine-processed video and voice, motion, and biometric data (Calvo and D’Mello, 2010). Over time, more integrated models of non-cognitive learning will emerge to blend different measures together in a more coherent fashion (Stafford-Brizard, 2016).

Quality teaching depends on an instructor’s recognizing problems or opportunities during the delivery of a course and making a precise and timely response. Sensing and reacting to the rhythm of a course while teaching is constrained by time, workload, and attention. Consequently, improving best practice in delivering a course depends on amplifying an instructor’s capacity for maintaining situation awareness, making a diagnosis, and implementing a response.

Note that the nature of the technological infrastructure in Figure 11.3 is similar to a K-12 DTP with face-to-face classroom interaction. With direct interactions, the teacher can glance at the students and quickly assess the nature of their participation.

 Via learning engineering, in online courses the system can be monitoring hundreds of students, or many more in the case of MOOCs. In addition, there can be multiple teachers facilitating the same course.

**Big Data and the Community of Teachers**

Online courses provide the opportunity to monitor teaching and student participation. The ongoing analysis allows the system to leverage working at scale. There can be massive numbers of students and massive numbers of courses, with a
variety of ongoing changes. In theory, all of this can be monitored, and the effectiveness analyzed. But what are the acceptable measures? A challenge for learning engineering is to evolve new processes and tools for efficient, systematic improvement.

In courses with multiple sessions and multiple instructors, there is a third loop that is three-dimensional between classes to allow for the teacher to be a learner within the community of teachers.

Learning communities of teachers are aided when participants can ground high-level strategies (e.g., answer a question with a question) in specific situations (Dede, Eisenkraft, Frumin, & Hartley, 2016). This contextualization illustrates how these approaches vary depending on specific factors. A DTP is ideal for a community of practice among teachers because the infrastructure can provide a steady stream of teacher–student–system interactions for analysis, leading to the collective development of instructional insights.

**Synchronous Versus Asynchronous Models**

The DTP not only disintermediates the relationships of teacher to student—and both of them to content (the feedback loops)—but also disintermediates the relationship between the present and the past or future. This begins to provide a framework for the longitudinal data collection ideas in the first section. It is only with scale over time that we can begin to address the patterns that are occurring in the online courses. By collecting both cognitive and non-cognitive evidence about learner progress over time, we are setting the stage for examining multi-dimensional trajectories of learner progress, and we can begin to look for patterns of success in both cognitive and affective domains for different kinds of interventions applied to different starting points in the high-dimensional space. Two students with a similar level of writing performance may in fact be at very
FIGURE 11.4 Multiple courses can provide data for the Analytics Engine and can foster a community of support for the instructor.
different starting points based on identity and motivation variables—they may need very different interventions to make progress in writing.

These patterns come into play particularly when thinking about actionable predictive analytics and/or formative assessment. What type of feedback helps the dynamics of an asynchronous classroom? In our judgment, many of the research areas in the next section address these dynamics.

A Research Agenda to Advance Learning Engineering in Online Higher Education

Next Steps for the Research Community

As the book chapters illustrate, researchers, instructional designers, and practitioners have found that learning engineering has great potential to deepen students’ engagement with courses, help teachers improve the efficiency of their instruction, and enable a broader range of learners to succeed because courses are tailored to their individual needs. However, much work remains to be done before these learning-engineering processes will be optimized for achieving these objectives. That optimization will require targeted and sustained support for research and development in learning engineering and in the learning sciences that are foundational for its success. At present, the most important priorities for research that advance learning engineering are (a) design-based research on the development of digital teaching platforms for online and blended learning, and (b) the development of models for effective collaboration between the learning engineers and others involved in online and blended course development and delivery, including researchers trying to understand how learning and motivation actually work and what can influence mastery at scale.

Important Dimensions for Research Design

Applied, Collective Research

At present, an important priority is for studies that produce usable knowledge about online instruction and evidence collection in higher education, motivated primarily, not by intellectual curiosity, but instead out of a desire to address persistent problems in practice and policy (Dede, 2011). This is not to disparage basic research or theoretical work that is valuable in the learning sciences and provides a foundation for learning engineering. That said, the purpose of learning engineering is the optimization of particular instructional offerings for specific contexts; as a by-product, insights about learning theory will emerge from the outcomes of learning engineering. Essentially, we would like more work in what is called Pasteur’s Quadrant (Stokes, 1997)—the place where problems have both high practical value and high research value—while at the
same time recognizing the value of work with mostly practical value (e.g., improving efficiency of instruction).

The process of creating and sharing usable knowledge is best accomplished by a community of researchers, practitioners, and policymakers. Individually developed outlier ideas can be valuable and should be some part of the research agenda in learning engineering. However, because the goal is to optimize the effectiveness and efficiency of instruction, what is required as discussed above is a multidimensional perspective, encompassing online learning as a vehicle for academic, interpersonal, and intrapersonal capacities that have been shown to matter to young people’s long-term success. Such a grand challenge necessitates a group effort, combining various research methods and integrating knowledge from several fields, including cognitive science, professional development, curriculum studies, instructional design, research methodologies, and online learning. This may necessitate directed development efforts along the lines used by DARPA (Defense Advanced Research Projects Agency) to rapidly increase the rate of progress on important practical projects (Shilling, 2015).

Fully understanding sophisticated online learning experiences that are effective across a wide range of contexts may require multiple studies that use learning engineering for each of their various dimensions, each scholarly endeavor led by a group that specializes in the methods best suited to answering research questions along that dimension. Using such a distributed research strategy among collaborating investigators, funders could create portfolios in which various studies cover different portions of this sophisticated scholarly territory, with complementary research outcomes enabling full coverage and collective theory-building, grounded in practical at-scale practice.

Research on What Works for Whom, in What Context

Numerous studies document that no single pedagogy is optimal for all subject matter and every student (Dede, 2008). Education developers too often assume that innovations cannot be brought to scale unless they can be replicated precisely and implemented with fidelity. However, experience has shown that the successful implementation of online learning depends on adaptations to each student. Therefore, the best way to invest in learning engineering is to acknowledge that context matters. Instructional design processes should be designed for customization to serve a range of educational settings, their teachers and students, their curriculum, their motivation and cognitive states, and their culture; professional development should include building educators’ capacity to make these adjustments. Therefore, statistical power in research studies about efficacy is important because, rather than assuming that an online learning experience is effective for all students in some bulk, universal manner, research and development should focus on what works for whom, in what contexts (Means, 2006). This is the multidimensional perspective referred to above.
Research Balanced Between Design and Evaluation

A research agenda should respond both to stakeholders’ needs for evaluative studies and the learning-engineering field’s need for research that informs design. Too many studies now privilege the former over the latter. A blended, empirical research model—designed not only to answer questions about whether a particular design works well but also to provide evidence to explain why it works—seems a reasonable and effective alternative to the evaluation-centric approach now prevalent. Such a research strategy also mitigates researcher-practitioner tensions. On the one hand, the need for evaluation-based marketing to potential adopters of learning engineering is a necessary evil for many scholars; on the other hand, the more theoretical work of explaining why and to what extent design interventions work is, at best, the icing on the cake to developers and vendors of educational online learning who know that ultimate sustainability depends on creating an evidence-based, scalable product. At the heart of all successful engineering processes is the iteration from evidence gathered in one generation to changes (based on theory, insight, judgment, or even guesswork) that will be measured again. If the measures of motivation as well as cognitive progresses are good ones (and that is far too rare now), then this flow of evidence can help both researchers and practitioners make progress.

Thus, a research agenda for learning engineering for online learning in higher education (and most other learning contexts) should focus on applied, collective research; adaptation of learning environments based on what works for whom and in what context; and a balance between design-based studies and evaluations of effectiveness.

Illustrative Research Questions

In no particular order, below are illustrative research questions suggested by the chapters in this volume. The intent is neither to display a complete list of possibilities nor to claim that these constitute the best research agenda. Instead, these are intended to start a dialogue about what such an agenda might include and how it might be formulated.

- What models of learning scientist, learning engineer, and instructor collaboration are efficient and effective?
- Which are the learner characteristics important to consider in personalizing instruction for motivation as well as for effectiveness?
- How do current roles (instructional designers, teachers, purchasing agents, managers, regulators) in developing and implanting online and blended courses need to change when learning-engineering practices are implemented, and what new roles are needed? What types of training are necessary for this shift?
What are effective ways in which students can provide ongoing feedback to instructors in synchronous online settings?

What measures and techniques can enable assessment of student skills such as systems thinking, collaboration, and problem solving in the context of deep, authentic, subject-area-knowledge assessments?

What high quality measures of non-cognitive aspects of the learner experience can be developed for efficient implementation with real-world, scalable environments?

What insights from cognitive neuroscience add new possible learner interventions to the learning engineering armamentarium for specific contexts that can be tested at scale?

**Developing Implementation Testbeds**

Integrating research into educational practice involves three steps (Carlile, 2004): Transfer from research to practice, Translation of research into practice, and Transformation of practice based on research. The initial stage, transfer, is likely insufficient because it reflects the traditional one-size-fits-all, scale-up model that has not led to sustainable improvements. Translation and transformation involve major shifts in policies and practices as mutual adaptation between local contexts and research-based innovations. In this process, risk-taking is essential. For example, it is crucial to see experiments that are informative failures as a success in advancing knowledge.

Researchers typically struggle to find even small numbers of classrooms and schools, as well as practitioners and policymakers, willing to engage in this type of risk-taking and experimentation. Therefore, educational test beds are needed that provide authentic contexts for transfer, translation, and transformation based on immersive learning. This requires substantial numbers of classrooms and schools willing to move beyond the traditional instructional and assessment models of practitioners, the conventional ideological and political standards of policymakers, and the usual rigorous practices and evidence standards of researchers. Finding venues willing to undertake immersive learning at scale is a substantial barrier to implementing the research agenda described in this volume.

**Concluding Thoughts**

The current state of learning-engineering work shows both mastery and appreciation of mastery to be in its infancy. Although there are now approaches to evidence-gathering and a body of learning science that can be drawn on, there is not yet an organized approach to integrating that science into principles and approaches that deliver predictable benefits at scale for the wide variety of learners and contexts out there. Yet, there are examples, now, of collaborations within online and hybrid learning environments between researchers and practitioners at scale showing what may be possible for the future.
References


INDEX

Page numbers in bold refer to information in tables, italics to figures.

Abdul Latif Jameel World Education Lab (J-WEL) 207–8
accountancy courses 72
action research 61
activity dashboards 80, 87
adaptive personalization 7, 110–11
ADDIE Model 58, 66n1
Agarwal, Anant 200–1
aggregate student learning (ASL) 150–1, 152, 159
aggregated data 56
algorithmic bias 28
Altman, I. 150, 156
analytics 18, 44–8; Google Analytics
   Academy 90–5; learning 2, 26–8, 29, 55;
   predictive 27–8, 82, 116–17, 221; see also
   big data; data analysis
Anderson, Lorin 22
Angelino, L. 145
application programming interfaces (API) 56, 77; Watson (IBM) 125–6, 127
Applied Online Course Design 74–7
applied research 61–2
apprenticeship 196, 197, 198, 199
apps 206–7
ARPANET 6
assessments 3, 124, 184, 214; autograder tool 121, 122; checkpoint 43; Digital Teaching Platform (DTP) 216, 221;
   formative 21, 62, 121, 221; mapping in
   Coursetune 74–9; norm-referenced-based tests 22; online tools 203;
   participation 72; personalization 114, 116
asset storage 60
Astin, A. W. 27
asynchronous courses 72–3, 84, 90, 145, 150, 219–21
attention spans 204–6
attrition rates 145–6
audio-visual media 19
autograde tool 121, 122
automated stage of mastery 175–6
Babson survey 36–7
backward design 175
behaviors 22; macro-behaviors 2;
   micro-behaviors 2, 37, 43, 46–7, 77–8, 89–92, 102–3, 109
Bell, Andrew 197–8
bias 28, 40
big data 1–2, 3, 26–8, 50, 151–3; Digital Teaching Platform (DTP) 218–19; see also analytics; data analysis
biology courses 73
Bird, Graeme 84
Blackboard 145, 146
Blase, K. 65
blended learning 1–2, 42, 199, 215; Harvard DCE 72–3; Kaplan 178; MIT 201; Starway 37
Bloom’s taxonomy 22, 176
Bluemix (IBM) 126, 135
Bonk, C. J. 152
Bouthillier, Larry 72–3
Bowen, W. G. 36
brain 28, 161, 208n6; attention spans 204–6; neurotransmitters 206, 207
Braun, Ian (virtual TA) 127, 130, 132, 135, 138
browsers (behavior) 99, 109
Bryk, Tony 41, 49
Bush, Vannever 20–1
Caliper 27
Canvas Network 146
Carnegie Foundation for the Advancement of Teaching 41
Carnegie Mellon University: KLI Framework 176; Open Learning Initiative (OLI) 41–4, 45
Cassidy Kimball (virtual TA) 130, 135, 136–9
change process: cyclic evaluation/ improvement 180–4; early adopters 171–4; evidence-based approach 174–8; incremental 38, 41, 50; role changes 188–91; scaling 166–7, 178–80, 184–7; transformational change 36–8, 49–51
chat channels 24, 158; online courses 146, 147–50, 152–3, 154; see also forums
chatbots 147, 150
Chautauqua Movement 4
Chen, Z. 144–5
Clark, Richard 20, 169
class group formation 153–6, 161–2; regrouping 156–7
classroom activities 215, 217; Digital Teaching Platform (DTP) 219
classroom based research 30
click-stream data 81, 82
cloud systems 56, 60
cognitive perspective 22, 167–8, 224; cognitive load theory 205; Digital Teaching Platform (DTP) 216, 217; interventions 169–70; task analysis 172
collaboration 31, 32, 85–9, 214–15; content development 58–61; course design 82–5; data intensive improvement model 41–7, 49–50; Harvard DCE (Division of Continuing Education) 80–5, 87–9; learning engineering 9–10, 39–40, 53–8, 63–4, 74, 221; organizational issues 48–50
collaborative data-intensive improvement research (CDIR) 9
communications models 21
communications tools 24–5, 60
communities of practice 178, 180, 219, 220, 222
competencies 7, 19, 168
completers 99, 103, 104–8
completion rates 109; Google Analytics Academy 96, 99, 100–3
computer aided instruction (CAI) 18, 23
computer programming education 198
Cone of Experience 20, 32n2
Consultation Model 5
content: asset storage 60; collaborative development 58–61; personalization 110–16; tagged 114
contextual inquiry 62–3
corporate training 6, 171; see also Kaplan, Inc.
correspondence courses 4, 5
costs: Higher Education 36, 37, 144–5; technological innovations 19, 55
CoSy 24
course design 8–9, 10, 88–9, 217–21; Coursetune 74–7, 78, 79, 86; partnership 82–5; see also design process; instructional design
course materials see content
Coursera 146, 200
Coursetune 74–7, 78, 79, 86
creativity 205, 207
Cuban, L. 6, 7
curriculum development 57–61, 215, 217
customization see personalization
Cyclops 24
Dale, Edgar 19–20
DALI (Deep Academic Learning Intelligence) 157–9; Omega Learning Map (ILM) 158–61
data analysis 56, 145, 151–3; aggregated data 56; analytics 18, 44–8; educational outcomes 38–41; effect of interventions 7; Google Analytics Academy 90–5; meta-analyses 29; see also analytics; big data
data collection 9, 27–8, 221–2; click-stream data 81, 82; Harvard DCE (Division of Continuing Education) 75–9; useraction events 77–8
data intensive improvement model 41–7
data mining, educational 2, 45–7, 145
DeBoer, J. 103, 114
decision making 1–2
declarative memory 159, 160
declarative stage of mastery 175–6
Dede, Chris 2, 3–4, 7, 8, 216–17, 221
deep learning 157–9, 162, 198
deep neural language networks (DNLN) 157, 158
demographics 76, 145–6; Google Analytics Academy 92–5
design process 63–4, 85, 189–91; see also
course design; instructional design
design science 25–6, 32n5
design thinking 18, 25–6, 29
Desire2Learn 152
diagnostic activities 113, 115, 116, 169–70
DiCerbo, K. E. 37, 38
differentiation 3, 109–10
digital assets 59–61
digital learning 194–5; chat channels 147–50, 152–3, 154; evaluation 48–9;
Higher Education (HE) 36–8; legacy systems 145, 146; system data 37, 46
Digital Teaching Platform (DTP) 12, 213, 215–21
direct instruction 196–7, 198
distance education 4–7, 217–18
Duncan, Liz (virtual TA) 130, 133–4, 135
Duolingo 206
dynamic regrouping 156–7
early adopters 167, 171–4
Ebbinghaus effect 206–7
Ecole 42 198, 208n4
e-commerce 92, 99–100, 103, 104, 117
education: changing nature of 194–5,
207–8; modes of 195–8; training in 170–1
educational data mining 2, 45–7, 145
educational outcomes see learning outcomes
educational software development 6
EDUCAUSE 37
edX 146, 201–2, 205
effect sizes 29, 32n6
effectiveness 2, 4, 213; evaluation of 21, 40, 48–9, 57, 63, 74, 84, 191, 219; of
learning engineering 7–8, 30, 215;
MOOCs 120; see also learning outcomes
efficiency 4, 7–8, 83, 215, 221; of teaching 2, 221–2 18F Guides 61
Elasticsearch 78
Electronic Information Exchange System (EIES) 23, 24–5
emotional states 169
episodic memory 130, 135, 159
Etesse, Chris 145
ethical issues 141
evaluation: change process 180–4; course materials 60–1; digital learning 48–9;
interventions 47, 62–4
evidence-based approaches 1, 7, 166, 170–1, 221–2; change process 174–8;
interventions 152; Kaplan 188–91; learning framework 172–3; scaling 11–12, 171–2, 192
expertise 39–40, 168–70, 171–2, 190;
acquisition of 175, 182; expertise reversal effect 110; transmission of 18, 172, 173, 182, 196
explicit personalization 110–11
extension schools 4
Facebook 146, 200
face-to-face learning 1, 23, 37, 80, 84, 217–18
faculty: and learners 144–5; and learning engineering 57–60, 64–5, 85–9
feedback: automated 19, 38, 72, 83, 125, 203; to designers 48, 50, 63, 178;
Digital Teaching Platform (DTP) 215, 217–18, 219–21; to learners 50, 87, 113, 171, 175–6; peer 38, 122
flight simulation 197
flipped classrooms 65, 72, 199, 201, 203
Ford Foundation 23
forgetting curve 206–7
forums 121, 122–5; virtual teaching assistants 127–35, 136–40; see also
chat channels
frequently asked questions 125, 130, 141

Gagne, Robert 22
game platforms 146, 147, 155
game-based learning 8
generative research 62–3, 65
Georgia Tech: Introduction to Computing 120–1; OMSCS 11, 122–7, 135, 141–2
Germany 197
Glaser, Robert 22
Goel, A. 124
Google Analytics, account properties 111
Google Analytics Academy 90–2; Advanced Google Analytics (AGA) course 90–1, 94, 99–103; business types 99–102, 104–6; completion rates 96, 99, 100–3; Custom Dimensions 91–2, 117n4; Google Analytics for Beginners (GAB) course 90–1, 94, 99–103, 117n7; personalization 109–16; user behaviors 104–8; user engagement 95–108, 109; user experience levels 97–8; user job roles 99–102, 107–8; user segmentation 92–5

Google Merchandise Store 91, 117n2

Google Tag Manager 91

Greer, D. G. 37

group formation 153–6, 161–2; regrouping 156–7

Guo, P. 205

Harvard 201; Harvard Extension School (HES) 71, 73; Harvard Summer School (HSS) 71

Harvard DCE (Division of Continuing Education) 10, 71–2; collaboration 80–5, 87–9; data analysis 75–9; online learning 72–5; organizational issues 79–81

Hattie, John 29, 32n6, 167

Heitmeyer, David 72–3, 84

HELIX Classroom 73

Higher Education (HE) 1–2, 36–8; costs 144–5

Hoxby, Caroline 155

human centered design 26, 58, 63

human genome project 151

hybrid courses 217–21

hybrid instruction see blended learning

Ian Braun (virtual TA) 127, 130, 132, 135

IBM: Bluemix 126, 135; Watson Engagement Manager 125–6, 127

implicit personalization 110–11

improvement science 38, 50; course design 10; iterative process 41, 180–4; learning outcomes 38–40

IMS Global Learning Consortium 27

institutional issues see organizational issues instructional design 28–9, 172; process 174–84, 222; role changes 189; see also course design; design process instructional media 18–23; comparison studies 20, 173–4; computer based 23–5 instrumented learning environments 26–8, 38

Integrated Postsecondary Education Data System (IPEDS) 37

interaction: learners with teachers 11, 120, 144, 156, 217, 219; in MOOCs 120; real-time 72, 75, 88, 145, 147, 153, 161

Internet 6–7; Internet of Things 194, 206 interventions 83, 84, 89, 214; cognitive perspective 169–70; evaluation of 47, 62–4; grouping students 151–3

James, William 159

Jill Watson (virtual TA) 135, 140, 141–2; Version 1.0 126–7, 128–9; Version 2.0 127, 130, 131, 133–4, 135; Version 3.0 130, 135, 136–9

job roles 223; Google Analytics Academy users 92–4, 99–102, 107–8

Joyner, D. 124

J-WEL (Abdul Latif Jameel World Education Lab) 207–8

Kaplan, Inc. 85, 166, 171–4; community of practice 178, 180; project-based training 179; Test Preparation Group (KTP) 173–4, 183; University 183–4, 187; Way for Learning checklist 172, 182–3

KBAI (Knowledge-Based Artificial Intelligence) 11, 121–7, 135, 141–2

Khan Academy 146

Kimball, Cassidy (virtual TA) 130, 135, 136–9

Kizilcec, R. F. 109

knowledge acquisition system 159–61

Knowledge Component Event Matrix 176

Koedinger, K. R. 176

Krumm, Andy 43

laboratory experiments 30

Lancaster, Joseph 198

Laney, Doug 1

learners see students

learning analytics 2, 26–8, 29, 55; see also analytics; big data; data analysis

learning assistance 125–41, 147; in MOOCs 120–1; technology 171; see also teaching assistants

learning environments 24–6, 145, 167; instrumentation 26–8, 38

learning management systems (LMS) 25, 55, 56, 122, 152; administrative functions 27, 37; system logfile data 37–8

learning outcomes 22, 38–41, 145, 161; data analysis 81; as evidence 1, 2;
improvement science 38–40; online environments 56, 151–2; under-represented minority (URM) students 45–7; see also effectiveness learning pathways 198–9, 203; online learning 199–200; personalization 112–14 learning process 88, 167–71, 204–7, 213; deep learning 157–9, 162, 198; evidence in system data 38; retrieval learning 206 learning styles 152 learning support see learning assistance learning technology adoption pattern 31–2 learning tools 122, 203; interoperability (LTI) 55 LiveEdu 147, 149 Liz Duncan (virtual TA) 130, 133–4, 135 Logstash 78 longitudinal research 63–4, 219 long-term memory 159–61, 167–8, 175–6, 206 LSATs (Law School Admission Tests) 173–4 LTI tools 55, 56 Lynda 146

machine learning (ML) 116–17 Madras system 198 Mager, R. 21 Mann, Horace 196 mapping, course design 74–7 Martin, S. 156, 157, 158, 159 mass lectures 144–5 mastery 19, 22, 224; learning engineering 182; routes to 7, 170; stages 175–7 mathematics courses 41–4, 45, 84 medicine 30, 31, 170 memex 21 memory 28, 159–61; Ebbinghaus effect 206–7; episodic 130, 135, 159; long-term 159–61, 167–8, 175–6, 206; short-term 159–61; working 159–61, 167–8, 173, 175–6 meta-analyses 29 military personnel training 20, 22, 23 mind wandering 205 MIT 195, 196, 208n1; bootcamps 202; Digital Learning Lab 203; Integrated Learning Initiative (MITili) 203; MicroMasters 201–2, 208n5; MITx 200, 202–3; Office of Digital Learning (ODL) 203; Office of Open Learning 200–4; Online Education Policy Initiative 7–8; Open Classrooms 198; OpenCourseWare 198, 200–1, 203; xPro 203 MOOCs (massive open online courses) 1, 7, 10, 28, 109–14, 121–2, 146; electrical engineering 200–1; teacher student interaction 11, 120–1; xMOOC (extended MOOC) 90; see also Google Analytics Academy; open learning Mood, T. A. 5 Moodle 152 motivation 109–10, 115, 124, 145, 168–70; see also learners mutual instruction 156, 197–8 nanotutors 124 Narayanan, V. G. 72 National Computer Systems (NCS) 24 NCME (National Council on Measurement in Education) 177 negative emotional states 169 networked improvement communities 41 neurotransmitters 206, 207 Next Generation Courseware Challenge 37 Next Generation Digital Learning Environment (NGDLE) 56 Niel, Xavier 198 non-cognitive learning 216, 217, 218, 224 Norman, Donald 26 norm-referenced-based tests 22 Nova NET 24–5

Olsen, L. A. 30 Omega Learning Map (ΩLM) 158–61 O’Neal, Cathy 27–8 online learning 1–2, 23–5, 72–5, 145–7, 217–21; Babson survey 36–7; big data 151–3; chat channels 146, 147–50, 152–3, 154, 158; class group formation 153–7, 161–2; course design 74–7; forums 121, 122–5; learner engagement 87, 205; learning pathways 199–200; personalized learning 4–7; tools 203 open learning: Carnegie Mellon OLI 41–4, 45; Google Analytics Academy 92, 109–10; MIT Office of Open Learning 200–4; personalization 92, 109–11, 117; see also MOOCs (massive open online courses) Open University, Britain 6, 25, 31 OpenCast 73; useraction events 77–8 organizational issues 166; collaboration 48–50; learning engineering 53–8, 64–6, 79–85
pain points 62, 63
participation 124–5, 144–5; assessment of 72; asynchronous 72–3, 75, 84, 90, 145, 150, 219–21; real-time 72, 75, 88, 145, 147, 153, 161
partnership see collaboration
Pearson Digital Learning 24
peers: feedback tools 38, 121, 122, 124;
peer effect models 153; peer to peer communications 146, 147, 217; peer to peer learning 156, 157, 197–8
Penn State University 4
personal data 150, 152, 153
personalization 2–7, 188, 214, 221, 222; adaptive 110–11; content 110–12, 114, 115–16; Digital Teaching Platform (DTP) 216; explicit 110–11; Google Analytics Academy 90–2, 103, 109–16; implicit 110–11; learning pathways 112–14, 162; online learning 4–7, 151–2; success metrics 114–16
Phelps-Coco, Adrienne 74
pilots see trials
Pinman, Isaac 4
PLATO (Programmed Logic for Automatic Teaching Operations) 23–5
predictive analytics 27–8, 82, 116–17, 221
Pressey, Sydney 19
privacy 2, 56, 76, 111, 126
procedural stage of mastery 175–6
programmed instructional materials 21
quality assurance 60–1, 62, 83
Quantway pathway 42
question channels 147
Quizlet 206–7
radio 5, 20
real-time participation 72, 75, 147; impact on learning 88, 145, 161
reliability 177, 184, 213
research 57; agenda 221–4; capabilities 62–4; comparison studies 20, 40, 124, 173–4; longitudinal research 63–4; paradigms 30–1, 40, 61–2, 87–9; reproducibility 29, 31
residential model 198–9
retrieval learning 206
robotics 194
Rogers, E. M. 58–9
role changes 188–91
Rosen, Andrew 180
Rudman, L. 173
Russia (Soviet) 5
Sacerdote, B. 153
Saettler, P. 18–19
safety 2
Sakai 152
Saxberg, B. 7, 8
scaling 115, 145, 166–7, 178–80, 213; change process 184–7; Digital Teaching Platform (DTP) 218–19; evidence-based approach 11–12, 192; Harvard DCE 84–5; MOOCs 109; personalization 162, 214; role changes 188–91; training films 20; user interfaces 150
Schon, D. 26
Scriyb 152–3, 154, 156, 161
security 2, 56
Seelye, Matthew 180
self-efficacy 169
semantic processing 127, 130, 135, 141, 159
sensory memory 159–61
short-term memory 159–61
Siemens, G. 56
silhouette 153
Simon, Herbert 17–18, 25, 31, 32n1, 166
simulation 196–7
Singapore University of Technology and Design (SUTD) 207
Sisko, Stacy (virtual TA) 127, 130, 131, 135
Skinner, B. F. 21
Slack 60
Snapchat 146
social engagement 156
social graphs 25
social learning 161
Social Learning Theory 157
social media 146, 147
Social Penetration Theory 150, 156
software development courses 72–3, 84
spaced learning 8
Spratt, Daniel 73
SRI 37, 41–4; Center for Technology in Learning (CTL) 38; see also Stanford University
Stacy Sisko (virtual TA) 127, 130, 131, 135
stakeholders 54, 61, 62–5, 223; changing roles 188–91; senior management 180, 181, 191
Stanford University 200; artificial intelligence MOOC 121; Machine Learning Course 28; see also SRI
State University of Iowa 5
Statway pathway 41–4, 45
Stockdale, S. 155–6
students: activity dashboards 80, 87;
notations by 79; attitude to research 64–5, 135, 140, 214; behavioral data 2, 22, 37, 43, 46–7, 77–8, 89–92, 102–3, 109, 150–3; demographics 92–5;
diversity 115; engagement 11, 42, 87, 95–108, 115, 124, 161; interaction with teachers 120, 144, 156, 217, 219;
navigation by 102, 110, 115; needs 112–13; pain points 62, 63; role changes 188; satisfaction with courses 26–7, 84, 145; silhouette 153; training DALI models 158; see also motivation
success metrics, personalization 114–16
Sullenberger, Sully 197
Sutherland, Andrew 207
Sweller, John 173
system logfile data 37–8, 39, 46–7
teachers 218; interaction with learners 120, 144, 156, 217, 219; role changes 188
teaching assistants: human 130; virtual 11, 125–42
teaching styles 7, 86–7, 152
teaching tools 216
Team-Based Learning 157
technology 5, 6, 171; cyclic approach 171; investment in 28
Technology Enabled Active Learning (TEAL) 202
television 5, 23
Tesla 194
testing effect 206
Thomdike, Edward 19
Ticknor, Anne 4
Tinto, V. 27
Tofel-Grebl, C. 168, 172
Towles, D. E. 157
training materials, co-ordination 23
transformational change 36–8, 49–51
trials 64, 173, 180, 183, 224; randomized control trials (RCT) 183–4, 185–7
Turoff, Murray 24
TUTOR programming language 24
Twitch 146
Twitter 146
Tyack, David 31–2, 36, 50
Tyler, Robert 20, 21
Udacity 122, 146, 200
Udemy 146–7, 148
under-represented minority (URM) students 45–7
universal memory bank 159–61
University of Chicago 4
University of Guelph 24
University of Illinois 23–4
University of Wisconsin 4
user centered design 26, 58, 63
user interfaces 55, 150; dashboards 80, 187; LiveEdu 149; Scriyb 154; Udemy 147, 148
user personas 63
useraction events, Opencast 77–8
users see students
validity 89, 177, 184
value for money 26
valuing 169
vampire learning theories 20, 32n3
Verduin, J. R. 5
Vincent, John H. 4
Virtual Teaching Assistants 11, 125–42
visual media: for instruction 5, 18–23, 173–4; live streaming 145–6, 147, 152, 156–7, 161–2; recorded lectures 77–8, 81, 122, 146
VORTAX 24
Vygotsky, Lev 197
Watson, Jill (virtual TA) 135, 140, 141–2; Version 1.0 126–7, 128–9; Version 2.0 127, 130, 131, 133–4, 135; Version 3.0 130, 135, 136–9
Watson Engagement Manager (IBM) 125–6, 127
The Well 24
Whole Earth 'Lectronic Link 24
whole student learning 150–1
word clouds 79, 81
work, changing nature of 171, 194–5, 196
worked examples 173–4
working memory 159–61, 167–8, 173, 175–6
World Wide Web, development of 6–7
writing courses 73
YouTube 50, 90, 146
Zone of Proximal Development 197