NEW AMERICA

EDUCATION POLICY

HIGHER EDUCATION

MANUELA EKOWO AND IRIS PALMER

PREDICTIVE ANALYTICS IN HIGHER EDUCATION

Five Guiding Practices for Ethical Use

MARCH 2017

About the Authors

Manuela Ekowo is a policy analyst with the Education Policy program at New America. She provides research and analysis on policies related to higher education including innovations in higher education delivery, the use of technology, and ensuring equitable outcomes for underrepresented students. Her writing has been featured in such media outlets as *EdSurge*, *Pacific Standard*, the *EvoLLLution*, and the *American Youth Policy Forum*.

Iris Palmer is a senior policy analyst with the Education Policy program at New America. She is a member of the higher education team, where she provides research and analysis on state policies related to higher education including performance based funding, state student financial aid, and state data systems. Palmer previously worked at the National Governors Association on postsecondary issues. There she helped states strengthen the connection between higher education and the workforce, support competency based systems, use data from effectiveness and efficiency metrics and improve licensure for veterans. Prior to joining NGA, she worked at HCM Strategists on the Lumina Foundation's initiative to develop innovative higher education models, including new technologies and competency-based approaches. Before joining HCM Strategists, Palmer worked at the U.S. Department of Education in all of the offices related to higher education: the Office of Vocational and Adult Education, the Office of Postsecondary Education, the Policy Office and the Office of the Undersecretary.

Acknowledgments

We would like to thank the Kresge Foundation, Lumina Foundation, and the Bill & Melinda Gates Foundation for their generous support of this work. The views expressed in this report are those of its authors and do not necessarily represent the views of the foundations, their officers, or employees.

About New America

New America is committed to renewing American politics, prosperity, and purpose in the Digital Age. We generate big ideas, bridge the gap between technology and policy, and curate broad public conversation. We combine the best of a policy research institute, technology laboratory, public forum, media platform, and a venture capital fund for ideas. We are a distinctive community of thinkers, writers, researchers, technologists, and community activists who believe deeply in the possibility of American renewal.

Find out more at **newamerica.org/our-story**.

About the Education Policy Program

New America's Education Policy program uses original research and policy analysis to solve the nation's critical education problems, serving as a trusted source of objective analysis and innovative ideas for policymakers, educators, and the public at large. We combine a steadfast concern for low-income and historically disadvantaged people with a belief that better information about education can vastly improve both the policies that govern educational institutions and the quality of learning itself. Our work encompasses the full range of educational opportunities, from early learning to primary and secondary education, college, and the workforce.

Our work is made possible through generous grants from the Alliance for Early Success; the Buffett Early Childhood Fund; the Foundation for Child Development; the Bill and Melinda Gates Foundation; the Evelyn and Walter Haas, Jr. Fund; the Heising-Simons Foundation; the William and Flora Hewlett Foundation; the Joyce Foundation; the George Kaiser Family Foundation; the W.K. Kellogg Foundation; the Kresge Foundation; Lumina Foundation; the McKnight Foundation; the Charles Stewart Mott Foundation; the David and Lucile Packard Foundation; the Pritzker Children's Initiative; the Smith Richardson Foundation; the W. Clement and Jessie V. Stone Foundation; and the Berkshire Taconic Community Foundation.

Find out more at **newamerica.org/education-policy**.

Contents

INTRODUCTION

Colleges are under increasing pressure to retain their students. Federal and state officials are demanding that those who enter their public institutions—especially students from underrepresented groups—earn a degree. Over two dozen states disburse some state funding on how many students an institution graduates, rather than how many it enrolls. Students and families are more anxious than ever before about crossing the degree finish line, as the financial burden of paying for college has increased significantly in recent years. And retaining students is becoming more crucial to the university bottom line. As recruiting and educating students becomes increasingly expensive, colleges hope to balance the resources they use to recruit students with revenue generated when those students are retained.

Because of these pressures, institutions have begun analyzing demographic and performance data to predict whether a student will enroll at an institution, stay on track in her courses, or require support so that she does not fall behind. Using data in this way is known as predictive analytics. Analyzing past student data to predict what current and prospective students might do has helped institutions meet their annual enrollment and revenue goals with more targeted recruiting and more strategic use of institutional aid. Predictive analytics has also allowed colleges to better tailor their advising services and personalize learning in order to improve student outcomes.

But while these are worthwhile efforts, it is crucial for institutions to use predictive analytics ethically. Without ethical practices, student data could be used to curtail academic success rather than help ensure it. For example, without a clear plan in place, an institution could use predictive analytics to justify using fewer resources to recruit low-income students because their chances of enrolling are less sure than for more affluent prospective students.

Last spring, New America interviewed more than 30 college administrators, experts, and education technology vendors, conducted an extensive review of existing literature, and visited a college campus (Georgia State University) to write **The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis**. That report, published in October 2016, examined how colleges are using predictive analytics and outlined the challenges schools face in ensuring that they are doing so ethically. Last fall we also convened an advisory council to discuss important considerations when using predictive analytics in higher education (see page 17 for a list of council members).

Our framework here aims to lay out some important questions to consider as administrators formulate how to use predictive analytics ethically. Examining the ethical use of data is an iterative process; colleges will continue to use student and institutional data in new and innovative ways and will therefore have to occasionally reassess whether their ethical standards address current data practices.

Using data ethically is complex, and no magic formula exists. This ethical framework is meant to start conversations on campus. It cannot address all possible issues surrounding the use—and potential abuse—of institutional data.

We recognize that student and institutional data may be used for various reasons and at various levels within an institution. For example, data on how students learn captured in an adaptive tool later used by faculty is likely governed differently than data collected in institutional enrollment processes used by admissions officers. There may be different rules about who has access to what information and whether students have the option to object to having their data collected or analyzed. In writing this guidance, we have tried to take these differences into account and believe that our recommendations are relevant for the variety of ways institutions may use student and institutional information for predictive analytics.

Colleges use data for predictive analytics in the following ways:

- **Early-Alert Systems.** In an early-alert system, flags are triggered based on academic and non-academic data from students that signal when they may need additional support. Academic interventions may include tutoring, meetings with an adviser, or assigning a coach or mentor to the student. For non-academic flags, colleges can deploy financial supports (i.e., emergency grants) or referrals to other supports (i.e., mental and medical health, child care, transportation, housing, and food).

- **Recommender Systems**. Recommender systems allow students to plan or map their degree, and integrate transfer credits or prior learning assessments into that plan. One common use for recommender systems is helping students choose courses to take next and/or choose a major based on data about their previous academic performance.

- **Adaptive Technologies.** Adaptive tools use data on how students learn to customize the learning environment for each individual student by identifying gaps in knowledge, skills, and abilities and adjusting content delivery to support deeper and more efficient learning.

- **Enrollment Management.** Enrollment managers use algorithms (computer-based rules) to decide how best to target recruitment efforts and distribute financial aid.

GUIDING PRACTICE 1: HAVE A VISION AND PLAN

Developing a vision and plan for data use will help steer the direction of a predictive analytics effort. Without such planning, predictive analytics may be used in a way that does more harm than good for students, leaves out key staff who should be included in the planning process, and/or fails to identify how success of this effort will be measured.

To develop a vision and plan, take the following steps:

Convene key staff to make important decisions.

In developing a plan, include key staff and stakeholders in decision making, and get their support. Including these individuals in the planning process can help ensure that you are using predictive analytics in a way that does not intentionally harm those whose data are being used and analyzed. Your team can include:

- head of advising/student success
- provost/chief academic officer
- head(s) of admissions, enrollment management, and financial aid
- head of student affairs
- head of institutional research
- head of communications, marketing, or public affairs
- key faculty
- other key staff (for example, chief diversity officer, privacy officer, information technology officer, data officer, digital learning officer, instructional designers, directors of centers for teaching and learning, career services, registrars, athletics, etc.)
- data scientists, modelers, and third-party vendors
- students

Ultimately, the staff and stakeholders included will depend on the way in which you plan to use data.

Consider the following three factors when developing the plan.

1. The purposes of predictive analytics

The plan should include the questions you hope to answer and the goals you aim to achieve. It should also explore the potential pitfalls of using student and institutional data for the purposes intended. The team should make sure that data will not be used for discriminatory purposes. For example, using predictive analytics to determine what amount of institutional aid might discourage low-income students from enrolling at your institution.

2. The unintended consequences of predictive analytics

The plan should also include a discussion about any possible unintended consequences and steps your institution and its partners (such as third-party vendors) can take to mitigate them. For example, using predictive analytics could lead to removing human judgment from decision making. This may result in decisions that typically require holistic review becoming partly or completely automated by data alone. Teams will also want to examine that they are not underestimating or overlooking how implicit bias—beliefs we may not be consciously aware of—may become more frequent with predictive analytics. In addition, they must guard against using biased algorithms that discriminate against certain students by limiting their access to college or to opportunities to pursue their interests once enrolled.

3. The outcomes to measure

The plan should also lay out the measurable outcomes you hope to achieve as a result of using predictive analytics.

Questions to Ask

Below are three questions to ask when creating a plan for using predictive analytics ethically in higher education.

- Have we set a goal/vision for using predictive analytics and/or adaptive technologies?
- Can our goals, methods, and measurable outcomes be explicitly stated in our institution's strategic plan and get support from key institutional officials? What would it take to make this happen? Is approval from the Institutional Review Board (IRB) necessary?
- Have we considered the unintended consequences predictive analytics may introduce? When drafting our vision and goals, have we made any assumptions about students or data?

GUIDING PRACTICE 2: BUILD A SUPPORTIVE INFRASTRUCTURE

A supportive infrastructure ensures the benefits of predictive analytics are understood and welcomed by campus stakeholders, and that processes and other supports are put in place to assist the data effort.

Communicate the benefits of using predictive analytics and create a climate where it can be embraced.

Predictive analytics uses student and institutional data to create change almost immediately. Many institutions may not be experienced with using data in this way, at this pace, and perhaps with such high stakes like ensuring students complete their degree in a timely manner. You should take the lead in communicating with campus leaders, staff, and students about why using predictive analytics is critical to institutional and student success. The head of communications and marketing could help in these efforts. Without a clear articulation of how using predictive analytics can benefit the campus, well-devised plans may fail to receive the support they need to be successful.

Develop robust change management processes.

With new tools often come new processes, reporting structures, people, and partners who bring new skills. This can, at best, create confusion for those charged with rolling out predictive analytics on a campus, and, at worst, chaos. Leaders convened to make important decisions about data use could also help ensure that processes are put into place to support the change taking place on campus.

Assess institutional capacity.

Assess your school's capacity to use predictive analytics. Having the appropriate technology, data infrastructure, talent, services, financial resources, and data analysis skills are essential. Maintaining a sound infrastructure can help ensure that cleaning, sharing, and using large amounts of data for making decisions institution-wide can be carried out smoothly and that different data systems can "speak" to one another.[1] Experts in information technology, student data laws, and staff with experience drafting contracts with vendors would help ensure the success of the project.

Questions to Ask

Below are two questions to ask to determine whether your institution has the infrastructure needed to embrace the changes that predictive analytics can bring.

- Have we communicated with stakeholders about why using predictive analytics in our decision making is important? Do we have the right messengers to build support for these new policies?
- Do we have the processes, structures, tools, infrastructure, people, and knowledge required to support the successful use of predictive analytics?

GUIDING PRACTICE 3: WORK TO ENSURE PROPER USE OF DATA

Predictive models (showing how different data points are related) and algorithms need data to build predictive tools that will support enrollment efforts or help students make academic progress. To build and use these tools ethically, consider the quality of your data and data interpretation, as well as issues around privacy and security.

Ensure data are complete and of high enough quality to answer targeted questions.[2]

Data about students and institutional processes should not only be accurate but also comprehensive. Consider for example, developing an adaptive learning tool for a developmental science course based on eight semesters' worth of information captured in the app about how students attempted questions or worked their way through course content—on students who tested out of the developmental course. The information collected may be accurate, but it is not comprehensive because the app was not tested on the population of students for whom the tool is designed.

Comprehensiveness also means considering all relevant data about the students who are being examined. For example, consider an early-alert system that flags at-risk students solely on their past grades and demographics. The system may have accurate data, but could potentially be missing important information such as students' current performance that may be key to determining whether they are struggling academically.

Beyond being accurate and comprehensive, quality data is also timely, derived using consistent tools and processes, and is well defined.

Ensure data are accurately interpreted.

Include staff members who are knowledgeable about your institution's data and can accurately interpret predictive models derived from this information. It is essential that those analyzing the data take context into consideration. It is also important to train faculty so that they can easily interpret dashboards that explain how students using adaptive tools are faring in their courses. Lastly, look for ways to ensure that data used solely for reporting purposes is sound even though they may also be included in data sets that are used for predictive analytics. If institutional researchers are responsible for both compiling data sets for reporting purposes as well as for conducting analysis for predictive analytics projects, the integrity of data for reporting should not come into question because information is being used on campus in innovative ways. Put simply, predictive analytics should not diminish the quality of data your institution is required to report to remain in compliance for federal funding.

Guarantee data privacy.

Communicate with students, staff, and others whose data are collected about their rights, including the methods used to obtain consent to use the data for predictive analytics and how long the information will be stored.[3] Make students and staff aware that their data are going to be used for predictive analytics and get consent to use highly sensitive information like health records.

Be vigilant that data are well protected so that the information does not get into the hands of those who intend to misuse it. It is especially important to protect the data privacy of vulnerable student groups, such as high school students who are minors and enrolled in dual-enrollment programs, undocumented students, and students with disabilities.[4]

In addition, make school policies on ownership of and access to student and institutional data clear. For example:

1. Advisers may have access to analyses of their students' performance data, and allow students to see these only and in a way that encourages progress.

2. Faculty may be allowed access to only selective information to enable interventions for students in their classes, but they can access all learning analytics data from learning management systems (LMS), adaptive technologies, or other digital tools required for class.

3. The president and provost may have access to only high-level analyses of students and department-level data that is devoid of personally identifiable information.

4. Enrollment managers and financial aid officers may have exclusive access to predictive analytics of prospective students for recruiting purposes.

5. If students will not be able to see a predictive tool, they should know what their rights are to access their data.

Monitor data security.

Security threats occur without notice.[5] As colleges collect and store more data on students and staff, and more devices that store data on the teaching and learning process are used in classrooms, security becomes an ever more pressing issue. For this reason, schools need to be vigilant about assessing data privacy and security. Monitoring threats and risks should be a regular undertaking. Data security requires you and your vendors to have security protocols that adhere to student privacy laws and meet industry best practices.[6]

To keep institutional data secure, involve your information technology (IT) department. Information security and privacy officers help keep institutional data safe. Providing regular training to IT and other staff about keeping these data secure should be a top priority.

Questions to Ask

Below are four questions to ask to determine whether your institution is using and maintaining data properly.

- How do we ensure that the data collected and the collection methodology are appropriate for the questions we intend to answer or the tools we plan to build?
- What role will students' current performance data play in predictive models? What role will qualitative data play?
- How will we get consent to collect student and staff data and analyze it, or is consent not necessary? Is awareness and transparency about how data are used sufficient?
- How will we safeguard student and institutional data and work to ensure that they are not used—internally and externally—for non-educational purposes?7

GUIDING PRACTICE 4: DESIGN PREDICTIVE MODELS AND ALGORITHMS THAT AVOID BIAS

Predictive models and algorithms can help determine the interventions an institution uses to support students or meet recruiting goals. Therefore, it is crucial that predictive models and algorithms are, at the very least, created to reduce rather than amplify bias and are tested for their accuracy. You should also ensure that models and algorithms are created in consort with vendors who can commit to designing them in a way that does not intentionally codify bias and so that they are able to be tested for veracity.

Design predictive models and algorithms so that they produce desirable outcomes.[8]

It is crucial to address bias in predictive models, ensure the statistical significance of predictions beyond race, ethnicity, and socioeconomic status, and forbid the use of algorithms that produce discriminatory results. An algorithm should never be designed to pigeonhole any one group.

Therefore, design or know how predictive models and algorithms are created in order to ensure desirable outcomes as determined by their vision and plan. Failing to take this approach may lead to inadvertent discrimination. For example, imagine a course recommender system with a predictive model that finds that low-income students or students of color are likely to struggle in college-level-math courses and recommends that these students be kept out of Science, Technology, Engineering, and Math (STEM) courses and programs. To avoid this self-defeating approach, consider using predictive analytics to help ensure that students from these groups who are interested in going into STEM are given appropriate support if they start falling off track.

Test and be transparent about predictive models.

Before predictive models can be used to develop algorithms, test them for accuracy, perhaps by an external evaluator. Predictive models should also be updated or refreshed to reflect new campus realities and goals. You may also want to limit the variables used in predictive models to those that can be easily explained and you should work to ensure algorithms can be understood by those who will be impacted by them. Such practices foster transparency and makes it easier to hold individuals accountable for creating poorly designed models or algorithms that produce discriminatory outcomes.

Choose vendors wisely.

Most colleges rely on an outside vendor to help them build models and predictive tools. To ensure models and algorithms are sound, transparent, and free from bias, you must be intimately involved with or knowledgeable about how predictive models and algorithms are built. Partnering with third-party vendors may make this harder.

Some vendors are transparent about their models and algorithms, and allow colleges to have a hands-on approach in the design process, or even let institutions take the lead.[9] Not all vendors, however, take this approach. Many consider their models and algorithms proprietary, meaning institutions are not involved in the design process or are deliberately kept out. You should make transparency a key criterion when choosing to work with any vendor.

Questions to Ask

Below are three questions to ask to determine whether your institution is positioned to address bias in predictive models and algorithms, ensure open and transparent designs, and work with a cooperative vendor.

- How will we address bias in predictive models and algorithms? Will we discontinue their use, or consider how to minimize the impact gap of interventions for different student groups?
- Do we understand how the algorithms work and are improved through machine learning, when a computer—on its own—learns from data and algorithms it is fed to create new algorithms?
- What are our standards for working with vendors? Have we asked the right questions when selecting them, such as:
    - How will your company and/or product be a fit with our institution?
    - How does your company ensure tools and services are developed and used transparently?
    - How will you hold yourself accountable for student outcomes?

GUIDING PRACTICE 5: MEET INSTITUTIONAL GOALS AND IMPROVE STUDENT OUTCOMES BY INTERVENING WITH CARE

How your institution acts as a result of what it learns from predictive analytics is where the rubber meets the road. Students will experience these actions or interventions firsthand, even if they do not see or understand how the algorithmic-based decisions are made. Despite the use of technology, humans primarily still have to deliver interventions. Therefore, it is important that interventions are thought about in the context of other supports offered at your institution and are disseminated with carefully communicated messages. Staff deploying interventions should be trained on how to intervene appropriately, and you should test the effectiveness of interventions once deployed.

Communicate to staff and students about the change in intervention practices.

Adding predictive analytics to the student success toolbox may spark a culture change as interventions informed by data become central to your institution. To get the campus to embrace this change, it is important to communicate how faculty, staff, and students will benefit from using interventions that are informed by predictive analytics, and allow them to guide the change as well. For example, faculty and advisors may be used to making decisions on how to intervene (or not to) based on their instincts about what a student needs, rather than looking at student-generated data to guide them. Offer professional development opportunities to faculty and staff and information sessions to students to communicate the benefits brought by using students' data to inform interventions. Training could also address how resource-constrained institutions should make decisions about who to help, starting with students with the highest need first.[10]

Embed predictive-driven interventions into other student success efforts.

Despite being a powerful tool, predictive analytics is still only one part of a suite of tools—like first-year orientation programs—that can ensure student and institutional success. Look for opportunities to leverage predictive analytics in ways that further advance other activities so that all student success efforts are connected and build upon one another. For example, it might be wise to help a student flagged as at-risk in an early-alert system see the benefits of taking advantage of learning communities or other supports offered. A one-off predictive analytics project that is not integrated or positioned to play off your broader efforts to help students succeed will likely not see the impact it could have if it was leveraged appropriately.

Recognize that predictive-driven interventions can do harm if not used with care.

Even when institutional data, predictive models, algorithms, institutional practices, and training are as good as they can be, mistakes can be made when acting on information. This is why interventions used in response to predictive analytics should be carefully calibrated to avoid harming students. Interventions used in response to data generated by students and predictive models can range from targeting a particular student for increased outreach based on his predicted chances of enrolling, requiring a meeting with an adviser based on the recommendation of an early-alert system, to changing the type of practice problem a student is assigned based on an adaptive technology system.

However, these tools should not be used without examining their potentially negative effects. Algorithms used for strategic enrollment management, early-alerts, recommender systems, and adaptive technologies require that colleges understand where they can do more harm than good. Viewing students from a wellness or asset mindset rather than an illness or deficit mindset may help ensure students are not harmed. This approach values all students as full of potential. In addition, it leaves room to consider institution-specific characteristics or barriers that have an impact on a student's risk of dropping out.[11] Finally, it will be wise to determine how individuals will be sanctioned for misusing or mishandling student and institutional data, as well as how to rebuild trust after a harmful incident has occurred.

Predictive tools can be used with care in the following ways:

Early-Alert Systems

Because interventions are deployed based on flags, they must be carried out with care to ensure that suggested meetings or external resources do not unintentionally communicate to students that they do not belong in college. Starting with small pilots to gather student feedback and determine the impact of early-alert systems and various supports may ensure success once they are implemented institution-wide.

Recommender Systems

Make sure that these systems are not simply suggesting majors that conform to historical norms like those which have discouraged women from STEM majors in the past, for example. These systems should expand rather than constrain student choice. Piloting these systems can allow your institution to examine if disparate outcomes for different groups result; piloting also allows for recommender systems to be recalibrated before institution-wide implementation.

Adaptive Technologies

Because the adaptations based on how a student behaves are seamless, students may be unaware of other learning paths they could have taken. Faculty should look at students' learning analytics dashboards to ensure that every student is making progress in the most optimal way possible and check that these technologies are truly accelerating the learning process.

No matter how effective the tool, however, adaptive technologies cannot improve poorly designed courses or make up for poor teaching.[12] Therefore, continue to encourage faculty to take advantage of professional development opportunities and consult with support staff who can help them improve their pedagogy.

Piloting a few courses or sections with adaptive technologies and assessing data on impact may help to understand how best to use the tool to improve student outcomes before adopting the adaptive tools institution-wide.[13]

Enrollment Management

Because the algorithms used in enrollment management are based on a student's likelihood of enrolling in the college and determined by historical college enrollment data, they can discourage an institution from enrolling students who do not fit the mold.

To gauge whether your institution's enrollment processes enable access for historically underserved groups, enrollment managers could examine over time whether the use of predictive analytics results in less diverse classes before committing to use this information indefinitely.

Carefully communicate when deploying interventions.

The messages you send should not demoralize students, and dissemination strategies should ensure that students are able to access interventions with relative ease.

Craft messages in the right way.

Studies have found that the way we deliver messages is key to changing behavior and attitude. For instance, an experiment in England found that people were much more likely to pay their delinquent taxes if they received reminder messages that addressed them in personal language and told them that most other people were paying.[14] Psychologists have also found that interventions that communicate that intelligence is something that can be improved with hard work make a large difference in the achievement of underrepresented groups in college.[15]

Given that framing a message in a certain way can have such a profound effect on behavior and attitudes, delivering a message in the wrong way can be extremely detrimental. Therefore, you should train yourselves and staff on how to carefully craft messages that are most likely to create positive changes in student behavior and mindset. Pilot these messages on a select number of students to gather feedback and information on impact before communicating them to students you wish to target. Lastly, do not underestimate the power of positive messages.[16] A school needs to strike a balance between communicating disapproval to students for things they do wrong and praise for things they do right.

Ensure interventions are accessible to target populations.

When designing an intervention, think about how it will be deployed and if the population you are targeting will be able to take advantage of it. For instance, if you are trying to encourage students who work off campus to see an adviser, the adviser needs to be available outside of traditional business hours and thus able to work with these students' schedules. Underrepresented populations generally do not have access to the same resources as students who come from more affluent households, and colleges should make sure that resources can be accessed in many different ways.

Train staff on implicit bias and the limits of data.

Staff should be trained on how implicit bias and the limitations of data can impact how they intervene with targeted students. Personal biases and an overreliance on institutional data can negatively affect the students they hope to serve. With the proper training, staff should eagerly embrace their obligation to use student and institutional data to produce positive results for students.

Combat implicit bias.

Even people who say they are not biased tend to fail implicit bias tests that force them to make split-second decisions when working with members of particular groups.[17] No matter how carefully interventions are designed, implicit bias can affect how faculty and staff deploy interventions for students and can result in negative outcomes.[18] And implicit bias may be heightened with predictive systems because analytics may serve to "confirm" bias, or make implicit bias even more invisible. Data can empower individuals to assume they no longer have to address their implicit biases because they are being led by numbers rather than by their beliefs, acknowledged and unacknowledged.

Fortunately, there are ways to recognize and address implicit bias. One experiment showed that when people were assessed for implicit bias and armed with strategies to combat it, they experienced a sharp reduction relatively quickly.[19] Ensure that all staff with access to institutional data and who deploy interventions reliant on data are trained on how to combat attitudes and beliefs they may not be consciously aware they have. Viewing all students as of intrinsic value can go a long way.

Understand data's limits.

Staff using data generated by students, predictive models, and algorithms to deploy interventions should know what that information does and does not tell them. Consider a scenario where a predictive model showed that a low-income student was at risk of dropping out. Staff should know this does not mean her socioeconomic status caused her to be at-risk, but that being poor is highly correlated with being at risk of dropping out. Predictive models are not certainties and should not be treated as such.

Training can also help faculty and staff understand that using institutional data responsibly means they should never allow it to supplant human judgment.[20] As institutions become more data-informed, staff may evolve into "student success scientists," confident with using institutional data to improve student outcomes. This use of data, however, should never lead to disregarding the role an individual plays when intervening.

Train students to use their own data.

Staff may also wish to train students to use their own data to guide their experiences on campus.[21] For example, students can use data they generate in adaptive learning tools to understand the conditions under which they learn best.

Evaluate and test interventions.

Do not declare an intervention successful until it has been tested and evaluated for its effectiveness.

What interventions work when, for whom, and why?

Test the efficacy of interventions you are using. Such testing can uncover whether these interventions have differential impacts across different groups and allow recalibration as necessary. Efficacy research could also help reveal whether the interventions are having any unintended consequences. For example, what happens when an intervention is accidentally provided to a student who does not need it? To test for efficacy, identify a well-matched control group and monitor student performance in both that group and the treatment group over at least one semester. Testing for efficacy should be a particular focus for interventions that lack an extensive empirical research base. Consider seeking approval from an Institutional Review Board (IRB) to conduct the research. IRBs provide another check on the ethics of the experiment and the intervention.

Test tools based on vendor claims before committing to them long-term.

Insist that vendors partner with independent researchers to validate the effectiveness of their tools and services. Whether tools are effective is a particular concern for adaptive technologies. Many third-party vendors claim their products are adaptive and will accelerate student achievement despite having little external validations for these claims.[22] As the field of technology-enabled student learning continues to develop, new tools and validation of claims should go hand in hand. Early-alert and recommender systems should be tested for effectiveness as well. External research can help confirm that models are using factors that go beyond those typically characterized with at-risk status (for example low socioeconomic income or first-generation status) to make predictions. These studies can also validate that the flags the vendor has identified are key indicators of risk and that addressing them leads to student success.

Questions to Ask

Below are some questions to help guide careful intervention.

- Have faculty, staff, and students embraced the necessary culture change brought on by predictive analytics? Are they trained to use new tools and intervention strategies to reap their benefits?
- How are we training staff to understand that results of predictive analytics are only snapshots of a student's experience at a given time and can change?
- Have staff been trained on how to balance decisions based on data with human judgment?
- What are the best practices for sharing labels like "at-risk," "low-risk," and "high-risk" from early-alert systems with students and staff?
- Are communication channels accessible to students?
- How should we ensure that using predictive analytics in enrollment and student success efforts does not encourage faculty and staff to profile students based on race, gender, age, and socioeconomic status?[23]
- Have we internally and externally validated the effectiveness of the predictive tools we are using and of the interventions we are deploying?

Colleges and their partners have a lot to consider to ensure that student and institutional data are used responsibly. Institutions and even students may be eager for these data to be used in new ways to promote student success. However, excitement about new tools and methods should not overshadow the need to make sure predictive tools are deployed in a purposeful and secure manner; have the right supports and infrastructure to take hold on a campus; are built with quality data; do not further entrench inequitable structures; and can be tested for and produce evidence of effectiveness. Predictive analytics are already changing how institutions recruit and support students. As use of these tools become second-nature, addressing their ethical use will become even more important.

Advisory Council

In fall 2016, we convened a group of experts in predictive analytics, learning analytics, and adaptive technologies and representatives from institutions using these tools. We met at our New America office in Washington, DC to discuss an early draft of this ethical framework. We are indebted to our advisory council members for their invaluable insights.

Members

Dror Ben-Naim, CEO and founder, Smart Sparrow

Sherry Bennett, associate vice provost for data analytics, University of Maryland University College

Kevin Carey, director of Education Policy, New America

Sylvia Cini, director, special projects (the NC-CBE Project), Central Piedmont Community College

James Cousins, senior statistical analyst, Rapid Insight

Scot Henley, director of strategic partnerships, Rapid Insight

Elisha Jarrett, associate director, University Advisement Center, Georgia State University

Amy Laitinen, director of higher education in Education Policy, New America

Mark Milliron, co-founder and chief learning officer, Civitas Learning

Denise Nadasen, independent research consultant, former associate vice provost for institutional research, University of Maryland University College

George Siemens, professor, executive director of LINK Research Lab, University of Texas–Arlington

Sharon Slade, senior lecturer, The Open University Business School

Bryan Terry, vice chancellor for enrollment management, University of North Carolina–Greensboro

Notes

[1] James Wiley, *Learning Analytics: What is it and Why Should You Care?* webinar, Eduventures, 2016, 20 min. mark, https://attendee.gotowebinar.com/recording/1715852044791268867.

[2] *Online Learning Analytics: An Improvement Cycle*, webinar, Online Learning Consortium and EdSurge Higher Ed, 2016, 50 min. mark, https://sas.elluminate.com/site/external/recording/playback/link/meeting.jnlp?suid=M.CC1B0F2D511E701400B2A1EFBECA55&sid=251&mkt_tok=eyJpljoiWkRJNE9EZ3lNVGxpTm1JMSIsInQiOiJyc1gxWVVweFcrOGVOK0NiRFBEa2tpaGRhXC9kaGVtakFjUEtRWVBaNHdaVnNsekNUQThkcU5qRFo2TTZoS3RTY2NBdjB5TFc3Zm5CRmJjMFZBdHJPNGMyVjZJb01MN2lPNnRPZ2hsXC80Y1lRPSJ9.

[3] Sharon Slade and Paul Prinsloo, "Learning Analytics: Ethical Issues and Dilemmas," *American Behavioral Scientist* 57, no. 10 (2013): 1509–1528, http://oro.open.ac.uk/36594/2/ECE12B6B.pdf; Kamala D. Harris, Ready for School: Recommendations for the Ed Tech Industry to Protect the Privacy of Student Data (Sacramento, CA: California Department of Justice, November 2016), 15, https://oag.ca.gov/sites/all/files/agweb/pdfs/cybersecurity/ready-for-school-1116.pdf.

[4] For more, see Niall Sclater, "Accessibility Considerations for Learning Analytics," *Jisc*, December 14, 2016, https://analytics.jiscinvolve.org/wp/2016/12/14/accessibility-considerations-for-learning-analytics/.

[5] For more, see "How Higher Ed Is Coping as a Cybersecurity Target in 2016," *EdTech*, March 8, 2016, http://www.edtechmagazine.com/higher/article/2016/03/how-higher-ed-coping-cybersecurity-target-2016?utm_source=Sailthru&utm_medium=email&utm_campaign=Issue:%202016-03-10%20Higher%20Ed%20Education%20Dive%20Newsletter%20%5Bissue:5189%5D&utm_term=Education%20Dive:%20Higher%20Ed.

[6] For more, see Privacy Technical Assistance Center, *Data Security Checklist*, U.S. Department of Education, July 2015, http://ptac.ed.gov/sites/default/files/Data%20Security%20Checklist.pdf; Kamala D. Harris, *Ready for School: Recommendations for the Ed Tech Industry to Protect the Privacy of Student Data* (Sacramento, CA: California Department of Justice, November 2016), 14–15, https://oag.ca.gov/sites/all/files/agweb/pdfs/cybersecurity/ready-for-school-1116.pdf.

[7] For more, see Manuela Ekowo and Iris Palmer, *The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis* (Washington, DC: New America, October 2016), 16, https://www.newamerica.org/education-policy/policy-papers/promise-and-peril-predictive-analytics-higher-education/.

[8] See Principle 7 of *Policy on Ethical use of Student Data for Learning Analytics*, The Open University, September 2014, 6, http://www.open.ac.uk/students/charter/sites/www.open.ac.uk.students.charter/files/files/ecms/web-content/ethical-use-of-student-data-policy.pdf.

[9] For more, see Manuela Ekowo and Iris Palmer, *The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis* (Washington, DC: New America, October 2016), 16, https://www.newamerica.org/education-policy/policy-papers/promise-and-peril-predictive-analytics-higher-education/.

[10] For more, see Paul Prinsloo and Sharon Slade, "Educational Triage in Open Distance Learning: Walking a Moral Tightrope," *International Review of Research in Open and Distance Learning* 15, no. 4 (2014), http://oro.open.ac.uk/40903/.

[11] For more, see Vanessa Scholes, "The Ethics of Using Learning Analytics to Categorize Students on Risk," *Educational Technology Research and Development* 64, no. 5 (2016): 939–955, http://link.springer.com/article/10.1007/s11423-016-9458-1?view=classic.

[12] Dror Ben-Naim, "What Makes a Smart Course 'Smart'?" *EdSurge News*, January 13, 2017, https://www.edsurge.com/news/2017-01-12-what-makes-a-smart-course-smart.

[13] Garrett Zimmer, "The Edtech World is a 'Swamp of Gimmicks'—and Here's How We Can Drain It," *EdSurge News*, December 3, 2016, https://www.edsurge.com/news/2016-12-03-the-edtech-world-is-a-swamp-of-gimmicks-and-here-s-how-we-can-drain-it.

[14] *Applying Behavioural Insights to Reduce Fraud, Error and Debt* (London, UK: Cabinet Office Behavioural Insights Team, February 2012), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60539/BIT_FraudErrorDebt_accessible.pdf.

[15] Paul Tough, "Who Gets to Graduate?" *New York Times*, May 15, 2014, http://www.nytimes.com/2014/05/18/magazine/who-gets-to-graduate.html.

[16] *From Throwing Stones to Creating Ripples: Ramapo's Approach to Student Success*, webinar, 28 min. mark, https://vimeo.com/188005355.

[17] "Understanding Implicit Bias," The Kirwan Institute for the Study of Race and Ethnicity at Ohio State University, http://kirwaninstitute.osu.edu/research/understanding-implicit-bias/.

[18] For more, see B*reaking Down Bias in Admissions: The How-to Guide to Preventing Admissions Bias at Your School* (Toronto, Ontario: Kira Talent, 2016), http://start.kiratalent.com/breaking-down-admissions-bias/.

[19] Patricia G. Devine, Patrick S. Forscher, Anthony J. Austin, and William T. L. Cox, "Long-Term Reduction in Implicit Race Bias: A Prejudice Habit-Breaking Intervention," *Journal of Experimental Social Psychology* 48, no. 6 (2012): 1267–1276, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3603687/.

[20] For more, see Data Quality Campaign and the Consortium for School Networking, *10 Foundational Principles for Using and Safeguarding Students' Personal Information*, 2014, http://studentdataprinciples.org/the-principles/.

[21] For more, see Alyssa Wise, "Data-Informed Learning Environments," *Educause Review*, October 17, 2016, http://er.educause.edu/articles/2016/10/data-informed-learning-environments.

[22] Phil Hill, "Marketing Claims From Adaptive Learning Vendors As Barrier To Adoption," *e-Literate*, August 29, 2016, http://mfeldstein.com/marketing-claims-adaptive-learning-barrier-adoption/?utm_source=e-Literate+Newsletter&utm_medium=email&utm_campaign=6d303c1ff1-RSS_EMAIL_CAMPAIGN&utm_term=0_deab6fbf84-6d303c1ff1-40283229.

[23] *Online Learning Analytics: An Improvement Cycle*, webinar, Online Learning Consortium and EdSurge Higher Ed, 2016, 47 min. mark, https://sas.elluminate.com/site/external/recording/playback/link/meeting.jnlp?suid=M.CC1B0F2D511E701400B2A1EFBECA55&sid=251&mkt_tok=eyJpljoiWkRJNE9EZ3lNVGxpTm1JMSIsInQiOiJyc1gxWVVweFcrOGVOK0NiRFBEa2tpaGRhXC9kaGVtakFjUEtRWVBaNHdaVnNsekNUQThkcU5qRFo2TTZoS3RTY2NBdjB5TFc3Zm5CRmJjMFZBdHJPNGMyVjZJb01MN2lPNnRPZ2hsXC80Y1lRPSJ9.