Current Issues in Education
Mary Lou Fulton Teachers College • Arizona State University
PO Box 37100, Phoenix, AZ 85069, USA

Arizona State University
Mary Lou Fulton Teachers College + ASU ShapingEDU

Volume 22, Issue 1 January 7, 2021 ISSN 1099-839X

# Shaping the Futures of Learning in the Digital Age

# Catalyzing a Culture of Care and Innovation Through Prescriptive Analytics and Impact Prediction to Create Full-Cycle Learning

David Kil, Chief Data Scientist, Civitas Learning
Angela Baldasare, Senior Principal Strategic Consultant, Civitas Learning
Mark Milliron, SVP & Executive Dean of the Teachers College, Western Governors University

Abstract: Student success, both during and after college, is central to the mission of higher education. Within the higher-education and, more specifically, the student-success context, the core raison d'être of machine learning (ML) is to help institutions achieve their social mission in an efficient and effective manner. While there should be synergy among people, processes, and ML, this synergy is not often realized because ML algorithms do not yet connect the dots on fully understanding and strategically fostering student success. Transitioning from risk to impact prediction is a catalyst for institutional transformation, which can lead to continuous learning and student-success process innovation. This paper explores how ML can complement and facilitate organizational transformation in promoting a culture of care and innovation through virtuous full-cycle learning.

Keywords: *Machine learning, data-informed decision making, impact prediction, prediction-based propensity score matching, real-world data, real-world evidence, prescriptive analytics, student success knowledge base, influence diagram, change management, impact elasticity*

Citation : Kil, D., Baldasare, A., & Milliron, M. (2021). Catalyzing a culture of care and innovation through prescriptive analytics and impact prediction to create full-cycle learning. *Current Issues in Education*, *22*(1). Retrieved from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/1903
This submission is part of a special issue, *Shaping the Futures of Learning in the Digital Age*, guest-edited by Sean Leahy, Samantha Becker, Ben Scragg, and Kim Flintoff.

Accepted: 10/27/2020

## Introduction

In the past 15 years, higher education has undergone a shift in the focus of institutional data use related to student success, moving from a *culture of accountability* to a *culture of care*. Whereas in the culture of accountability, the focus was on reporting and performance metrics, the culture of care has embraced the use of data to inform proactive practices that support students in their educational journey, improving experiences and outcomes. In the pursuit of improved reporting metrics and student-experience optimization, there has been an increased use of predictive analytics and other data-informed approaches to impact student success in higher education. Gagliardi and Turk (2017) discuss the importance of creating a data-enabled executive focusing on improving student success outcomes. Romero, Ventura, Pechenizkiy, and Baker (2011) discuss various data mining techniques used in education in the learning space, including case studies using both structured and unstructured data. Mandinach and Jackson (2012) explore ingredients in a culture of data-driven decision making and examine Easton's cycle of inquiry consisting of problem identification using data, solutioning, monitoring progress, and using research to examine impact. Dowd (2005) discusses a culture of evidence in institutional decision making, explaining various forms of benchmarking practices in higher education.

A common primitive underpinning of all these efforts to improve institutional effectiveness and to move toward an informed culture of care is the knowledge base of what's working for whom, when, and with what. In healthcare, for example, many researchers and companies are building knowledge bases on evidence-based medicine and precision medicine to provide more effective and personalized treatment to patients (Chow et al., 2018). This knowledge base is, in most institutions, fragmented and incomplete, largely undocumented, often anecdotal or evidenced by descriptive data rather than rigorous evaluation methods. Higher-education leaders want to be data-informed in their student success efforts but lack a bridge that links evidence-based student-success knowledge to current actions that can be taken to specific students with a high probability of impact. A student-success knowledge base would identify the types of evidence-based programs, practices, policies, or interventions that have been demonstrated to have the largest impact on the success of specific student populations. Unfortunately, today's ML applications are not geared toward building an understanding of efficacy, instead focusing almost exclusively on risk prediction.

In healthcare, Kil, Shin, and Pottschmidt (2004) highlight that predicted risk is not equal to impactability or statistically significant and replicable student success outcomes. An initial pilot targeting patients with diabetes based on predicted risk scores showed promising results. This intervention focused on lifestyle coaching leveraging behavioral science. When the pilot was expanded to all patients prioritized by prediction scores, the same impact analysis showed negative outcomes despite the overall targeted population being much sicker. A thorough, scientifically rigorous, drill-down impact analysis revealed a mismatch between intervention and sub-population characteristics. This finding led to differentiating risk prediction from the potential for positive impact or impact elasticity. It also facilitated the use of social psychological factors derived from the Patient Activation Measure to personalize interventions further (Hibbard et al., 2004). In the end, a portfolio optimization approach was developed to assign patients to various patient-care programs, each personalized to the unique impact elasticity profile of a group of patients.

Creation of a student success knowledge base similar to this portfolio of patient-care programs is not a trivial task. There are multiple challenges associated with the creation.

- *Conflict between randomized experiments and observational studies*: There is an academic bias towards randomized controlled trials (RCT). Black (1996) posits that RCTs and observational studies can complement each other. He examines the strengths and weaknesses of randomized experiments and observational methods while suggesting that hypotheses be derived from retrospective observational studies and be confirmed through randomized experiments for higher success rates in prospective clinical trials. The biggest weakness of observational studies, according to Black, is the potential for confounding factors being unevenly distributed between intervention groups. Could data-driven ML approaches help us infer important confounders from data so they can be used in the matching process to address the biggest shortcoming?
- *Population heterogeneity in the real world coupled with sparsity of results from RCTs*: RCTs, while important, are no panacea in replicating results in the presence of heterogeneities of treatment and population in the real world. It is well known that drugs approved based on RCT results from relatively homogeneous clinical trial populations resulted in unexpected side effects in heterogeneous real-world populations, thus requiring pharmacovigilance using real-world health data (WHO, 2002). The National Academies of Science, Engineering, and Medicine (2019) held a workshop to explore how to maximize the impact on medical product development of real-world evidence from observational studies. Franklin, Glynn, Martin, and Schneeweiss (2019) also discuss potential use cases of real-world data for regulatory decision making.
- *Sparsity of high-fidelity time-scale metrics to chart student success progress and responses to heterogeneous treatments*: The longer the time span between intervention and outcome, the more noise the measurement process needs to deal with. In healthcare, there are a lot of -omics efforts in immunomics, genomics, radiomics, metabolomics, and related Internet-of-things (IoT) sensor technologies to understand and untangle the mysteries of how human bodies respond to various dynamic treatments (Caron et al., 2007). Jones and Agusti (2006) talk about the need for new markers or surrogate endpoints in the management of chronic obstructive pulmonary disease in conjunction with observable patient symptoms. The key idea here is to use fast surrogate endpoints highly correlated with primary endpoints to strengthen the discovery of causality between treatment and outcomes consisting of multiple time-scale metrics.
- *Lack of a unifying framework around intervention opportunities and impact results:* Oreopoulos and Petronijevic (2019) report that they failed to replicate earlier promising results from scalable online and text-message interventions to improve college achievement. However, student characteristics may differ across the interventions in terms of relative magnitudes of student success drivers and intentionality in intervention design that accommodates such potential differences. Too frequently, intervention results are viewed in isolation without examining differences in student success drivers and intervention design intentionality, personalized to target population characteristics. Yeager and Walton (2011) explore how wise psychological interventions should be designed

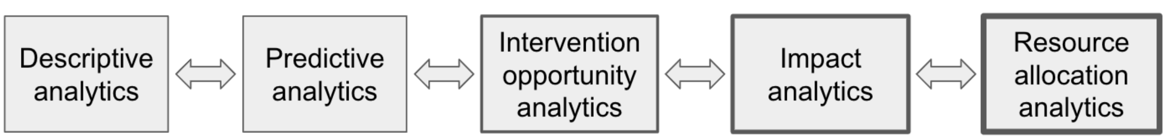
with the intention of having a long-lasting impact by addressing the underlying social psychological impediments with many nuances.
- *Lack of consensus on consistent methodologies in measuring impact results across time and space in the presence of noise and confounders in observational studies*: There appears to be a lack of consensus among academic researchers on the best ML-based impact analysis algorithms in the presence of real-world constraints and challenges despite recommendations from What Works Clearinghouse (WWC, 2019). The most frequent argument against ML-based approaches in causal inference is dependence on the practitioner's skill level in building ML models (King & Nielson, 2019). However, more and more robust ML modeling algorithms are being made available through ML libraries offered by most cloud analytics vendors, leading to ML modeling democratization.

Despite these challenges, the time is ripe for a connected ML approach that goes beyond predictive modeling. The authors' goal is to reinforce the importance of connected ML in catalyzing institutional transformation. Connected ML is defined as a set of complementary ML algorithms with a focus on improving student success, not just on predicting student risk. This paper is organized as follows. The building blocks of student success problem solving are first presented. Next we discuss how the proposed ML framework can catalyze institutional transformation that involves people and processes by empowering a data-informed culture of care. Finally key concepts are reinforced through real-world examples and recommendations.

## Method

Figure 1 shows the analytics building blocks that power ML applications focused on solving the problem of student success (Baer et al., 2019). While descriptive and predictive analytics solutions are designed to help explain who is at risk and why, they just represent the tip of the iceberg.

**Figure 1**
The five synergistic building blocks of ML analytics.




### Descriptive and Predictive Analytics

Predictive models assign success probabilities to students based on student and institutional variables. Going beyond the role of traditional predictive models, not only should ML recommend intervention strategies, but it also needs to gather and provide evidence that the recommended intervention strategies will bear fruit in a cost-effective, scalable way. Another important need is to create and sustain a rewarding user experience for the key stakeholders by helping them see how their actions are helping students do better. Nir Eyal (2014) mentions the importance of variable intrinsic reward while Mihaly Csikszentmihalyi (2009) talks about how to create the optimal user experience called *flow* by immersing the user in a deeply engaged and satisfying state.

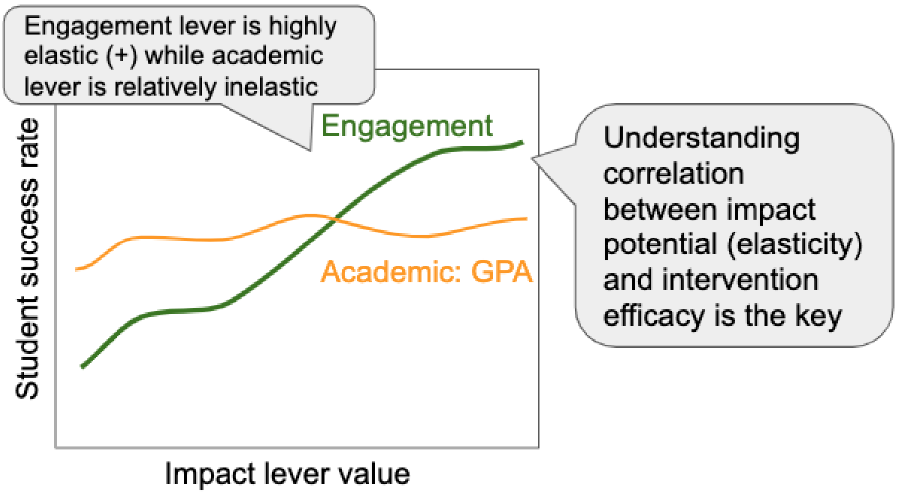
This paper posits that, given the social psychological and motivational profile of stakeholders in higher education, the optimal user experience comes from empowering them to maximize human intelligence in designing intentional interventions and knowing how they are contributing to student success. ML can help stakeholders stay engaged by transforming their roles from operational to empowered through intentional intervention design and execution with timely feedback on how much difference they are making and what they can do differently to further improve student success.

Those higher-education leaders who are engaged innovators deploy interventions in a data-informed way, testing, learning, and applying that learning in an iterative cycle of improvement. A range of data is necessary to employ this iterative improvement strategy. Both descriptive and predictive analytics can be used to provide a comprehensive dashboard view of student and stakeholder health statistics and trends, where health can encompass student-success and stakeholder-engagement metrics.

**Prescriptive Intervention Opportunity Analytics**

The key tenet behind prescriptive analytics is understanding the relationship between impactable levers, or simply impact levers, and student success rates, which is called elasticity, borrowing a term from economics, such as price elasticity of demand. The same concept is directly applicable to student success science. For example, academic performance can be inelastic for certain groups of students, such as those contemplating on early transfer or those with logistical challenges. In these situations, engagement matters far more. For other student segments, the opposite can be true. It all depends on the characteristics of students in heterogeneous student populations. Figure 2 illustrates the basic concept of impact elasticity. The steeper the magnitude of the slope, the more elastic. For this example, programs designed to improve student engagement through peer mentoring or sense-of-belonging nudges by faculty can be more effective.

**Figure 2**
The concept of impact elasticity



Student success rates can be defined in different time scales, spanning multiple dimensions, such as academic performance, engagement, thriving quotient, learning metrics, persistence, completion, job success, and donations. The longer the time duration from treatment to impact, the greater the amount of performance ambiguity for treatment. This is why elasticities should be measured with respect to short-term metrics or surrogate endpoints, such as persistence, LMS engagement, grades-in-progress, successful course completion, term GPA, or course-based net promoter scores.

Impact levers should be derived from time-series student records. Let’s explore student enrollment behavior. From data, students who enroll early tend to persist higher than those who enroll at the last minute, often getting into sections at unpopular times or being forced to take non-optimal courses that may not count towards graduation. The magnitude of the elasticity metric by student segment can provide insights into how to nudge the right students at the right time to be more proactive in enrollment or in seeking assistance with enrollment barriers. Nudging can be more effective by understanding the key differences between students who register early and those who do not (descriptive analytics) as well as investigating and addressing in intentional intervention design key impediments to proactive re-enrollment, such as financial aid, registration blockers (satisfactory academic progress, parking tickets, etc.), or personal issues. Moreover, the registration timeliness nudging is more precisely applied when student segments whose behavior-to-outcomes predictions are less elastic (e.g., strong students registering late when the institution cancelled their original course) are identified and removed from treatment protocols. They clearly do not need the reminder nudge even though a traditional risk prediction trigger would have been activated.

In summary, prescriptive analytics is a set of ML techniques to evaluate and rank order various static and dynamic groups of students based on N and impactability. This is the first step in improving student success in an efficient manner.

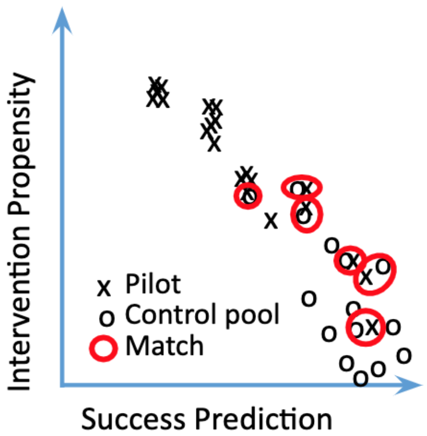
**Impact Analytics**

ML-based impact analysis forms the core of the student-success knowledge base creation process given the relative lack of RCTs in real-world, heterogeneous observational settings. While there are many algorithms for causal inference, here we focus on observational studies relying on RWD given the reality of how most student-success programs operate.

Rosenbaum and Rubin (1983) proposed that, in observational studies, pilot and control samples should be matched in a single-dimensional projection of known covariates (x) about the samples or students in this case. This projection or data compression is typically performed using ML, which is known as propensity score modeling. Propensity score is defined as $p(z = 1|x)$, where z is the treatment variable and $x$ is the student covariate vector. That is, a propensity score model is built using covariates of patients ($x$) in the pilot ($z = 1$) and control ($z = 0$) pools by training a model to learn differences between pilot and control students in the covariate vector space.

The next key question is on covariate selection as there is typically no clear guideline. Through experimentation, Kil et al. (2004) found that constraining covariates to the best features of a model designed to predict an outcome metric produced consistent, high-quality impact results. This finding led to a two-dimensional matching algorithm in the prediction- and propensity-score space. Figure 3 shows a pictorial depiction of prediction-based propensity score matching (PPSM).

**Figure 3**
The core concept of prediction-based propensity score matching.




Brookhart et al. (2006) also found that using the same approach resulted in the highest statistical power. McCall, Cromwell, Urato, and Rabiner (2008) similarly used prediction scores in evaluating impact results of Medicare health savings demonstration programs even when they were relying on cluster-based randomized controlled trials. Leacy and Stewart (2014) found similar advantages of prognostic-propensity score matching. Recently, there have been more validations in ML-based algorithms for causal inference (Hill & Su, 2013; Wager & Athey, 2018).

Furthermore, the judicious use of ecological momentary assessment and feature engineering on time-series student data can overcome the key objections of observational studies by inferring non-academic factors from data. Feature engineering algorithms, especially digital signal processing, are often less appreciated in ML due to a recent proliferation of deep learning networks. However, deriving features from linked-event data inspired by subject matter experts can be very helpful in impact analysis by incorporating into the matching process inferred non-academic factors, such as response to adversity, enrollment behavior, and engagement, to ensure that they are evenly distributed between pilot and control.

Many higher-education institutions conduct impact analyses of their student success programs using PPSM (Milliron et al., 2017). Based on impact analysis results and intervention insights, they identify how to improve intervention efficacy further, make those changes, and then “rinse and repeat,” resulting in full-cycle learning.

### Evidence-based Student Success Knowledge Base Creation

The student success knowledge base should be constructed from a scientifically rigorous impact analysis of each student-success program in a portfolio of student success programs. At a minimum, impact results should be a function of student segments and program operational parameters. The strength of evidence can be assigned based on design intentionality, treatment specificity, efficacy results with p value/confidence interval, and replicability that accommodates population heterogeneity. This knowledge base with evidence strength is of paramount importance in resource allocation optimization -- unless decision makers know what works for whom by how much, it is impossible to allocate resources optimally.

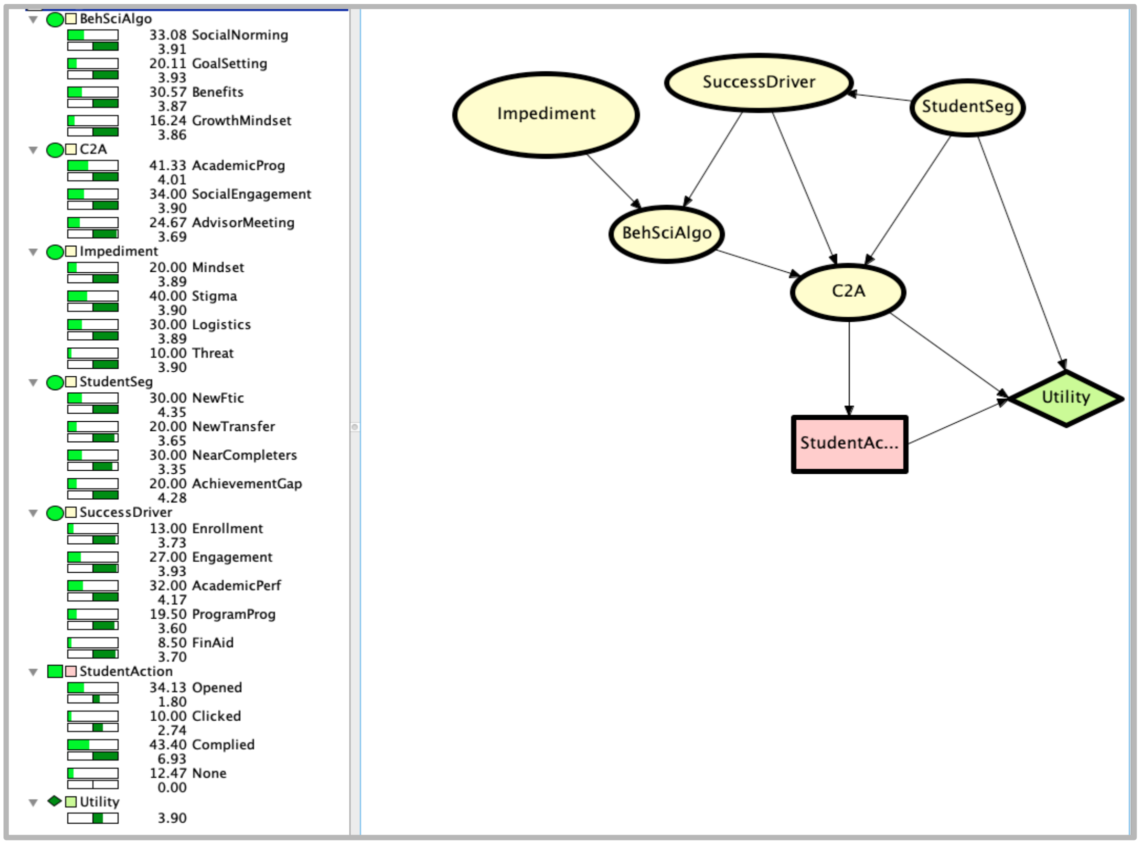
Each institution has a portfolio of student success programs and interventions designed to improve various influenceable factors associated with student success. These factors can be grouped into academic, non-academic, and financial aid categories. More specifically, we can identify impact levers that can be influenced through interventions in the areas of engagement, enrollment behavior, academic performance, academic progress, and financial aid. PPSM can be used to measure impact results of each intervention for the overall population and various drill-down student segments under different operating conditions. This ability to drill-down and see differential results for sub-segments of the student population creates the opportunity to optimize interventions by matching interventions to those for whom they will be most effective. Having this knowledge can facilitate the identification of further opportunities for continuous process improvement and portfolio optimization.

While there are multiple methods for resource allocation optimization borrowing ideas from financial portfolio optimization, one popular ML approach is an influence diagram (Owens et al., 1997). The fundamental ideas are summarized succinctly below:

1. The building blocks are three types of vertices and edges that connect them.
2. The first type of vertex represents a probabilistic state of variable important for student success, such as an impediment to success or student success drivers that depend on student segments. They are represented as circles in Figure 4.
3. The Bayesian network describes a causal relationship among the vertices that are associated with student success, such as student segments, student success drivers (impact elasticity levers), impediments (social psychological blockers), and intentional intervention design (IID). IID consists of call to action (C2A) and behavioral science algorithms used in content design to address impediments and success drivers in order to maximize compliance with prescribed call to action.
4. The utility vertex (diamond in Figure 4) is constructed from and analogous to the evidence-based student success knowledge base. The utility value can be mean persistence lift or its Sharpe ratio version, which is defined as mean over standard deviation. Standard deviation can be derived from confidence interval.
5. The decision or action vertex (rectangle in Figure 4) represents a student's action in response to delivered intervention.
6. An edge represents a direction of conditional probability relationship. For example, student success drivers depend on student segments.

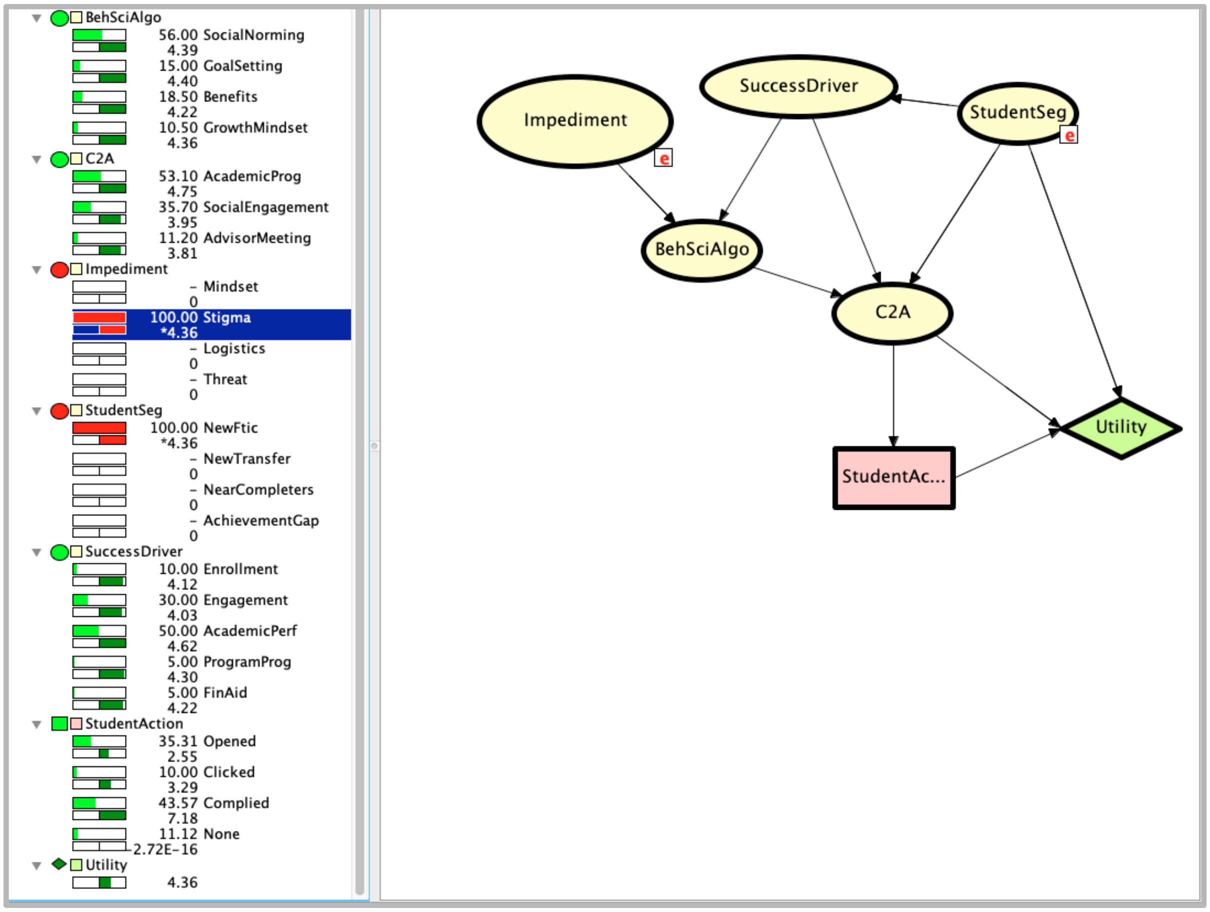
Figure 4 shows a student-success influence diagram. One of the nicer features of the influence diagram framework is that the *a priori* and conditional probabilities can be derived from data or entered by subject matter experts. They can be refined continuously as more data is gathered. A main use case for the influence diagram can be thought of as an air traffic controller, precisely engaging students with the right programs where they are likely to benefit the most based on heterogeneities in student characteristics in terms of impact elasticities, treatment, and timing of treatment based on external stimuli. This framework promotes the design, test, and implementation of the best-of-breed intervention programs that engage the right students at the right time, leading to student success portfolio optimization at an institutional level. An influence diagram can be used to decide on optimal strategies for each student segment based on (1) conditional probability relationships inferred from data and/or instantiated by subject matter experts and (2) instantiated evidence for vertices, i.e., known nodes. Here, utility is a function of student segment and call to action subject to their action or response to intervention.

**Figure 4**
Student-Success Influence Diagram




When targeting new, struggling, first-time-in-college (FTIC) students whose main social psychological blocker is stigma inferred from their good academic performance in high school, those evidences can be entered into the network, which will propagate and update state probabilities in each remaining vertex as shown in Figure 5. For example, social norming becomes more important (33% to 56%) in the behavioral science algorithm node while academic program is the call to action with the highest probability (41% to 53%) and utility (4.01% to 4.75% predicted lift in persistence). Furthermore, utilities for the uninstantiated node values (below the probability bar) change by propagating the entered evidences through the network.

**Figure 5**
After two pieces of evidence are entered into the Impediment and StudentSeg vertices, probabilities and utilities change across the board, reflecting the new evidence, compared to Figure 4.




Such information is crucial in designing intentional interventions where appropriate behavioral science algorithms are leveraged to address the key social psychological blockers and to give students an incentive to comply with the call to action that has the highest probability to improve student success. Given the likely success drivers, it looks like an academic program has the highest utility score of 4.75% persistence lift, followed by a social engagement program with 3.95% persistence lift utility. The utility vertex can be instantiated from impact analysis results. Once the top success drivers are known from prescriptive intervention opportunity analysis, they can be entered into the network, thereby influencing intervention design based on which call to action will have the highest probability of impact.

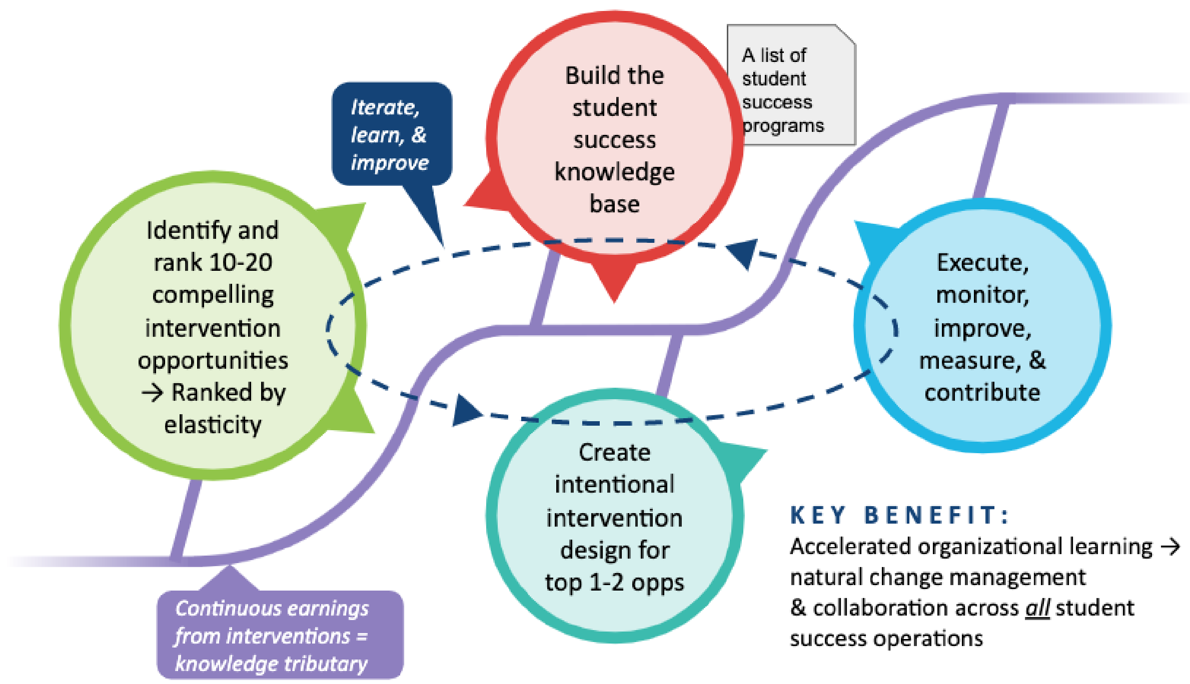
## Discussion: Putting Together Building Blocks

The ultimate goal is to leverage data and evidence on student success in catalyzing institutional transformation, where students and stakeholders are the key beneficiaries of the transformation. Unfortunately, on-ground realities from our recent user study feature these comments from educators trying to lead student success innovation:

- It is very hard for me to get timely answers to basic questions that should help me make daily decisions.
- I struggle to understand where my team and I are spending time on day to day and how much difference we are making towards student success.
- I do not know how to find the best opportunities for student success.
- I don't know which student success programs are working.
- It is not obvious what actions I should be taking. Besides, I am so busy with handling inbound student traffic that I have little time for proactive, intentional student success interventions.
- We have silos and a high degree of organizational inertia. I find it difficult to manage cross-team planning and execution to transition from analysis to action.

To address these pain points, it is very important to have a fully-connected ML analytics suite in Figure 1. Based on the authors' collective experience, success begets more success. Figure 6 shows the necessary ingredients in creating full-cycle learning for organic organizational transformation.

**Figure 6**
Creation of a virtuous cycle in an organization




The first task is to identify and rank prescriptive intervention opportunities for various static and dynamic student segments. Rule-induction algorithms can be used to find compelling intervention opportunities with appropriate guidance from subject matter experts.

The second task is to compile a list of student success programs with taxonomy, organized around impact levers. This step is crucial in constructing a portfolio of student success programs with an eye towards resource allocation optimization and impact prediction-informed investment decisions. Scientifically rigorous impact analysis can help an institution create its own knowledge base. A network effect can be especially helpful when a comparable group of institutions undertakes the same behavior to accelerate the pace of learning. Figure 7 shows an example of such a knowledge base for new and early-term students compiled from approximately 1500 programs submitted by more than 60 institutions, analyzed through PPSM.

**Figure 7**

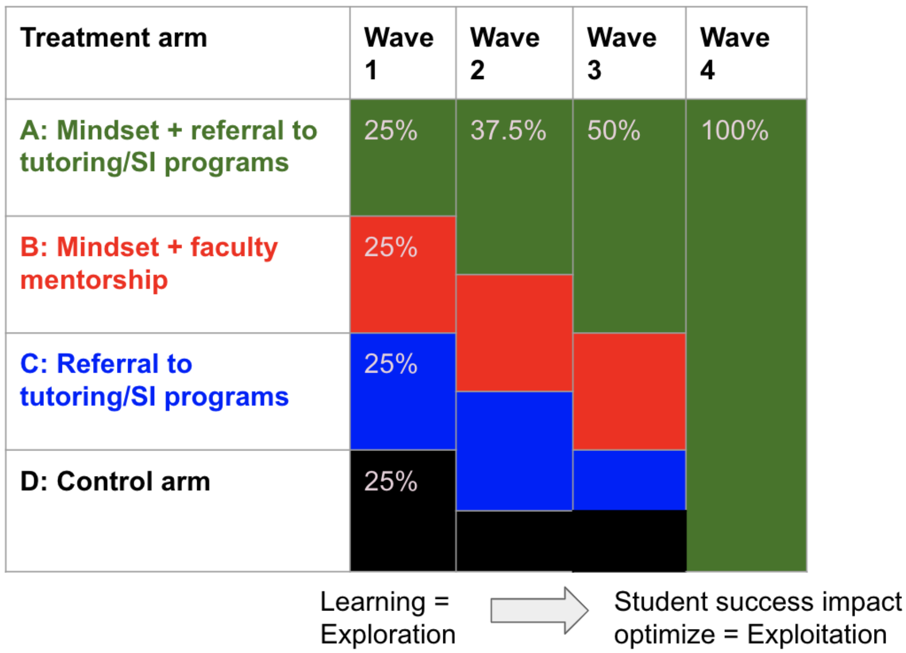
A table of a subset of student success programs along with their efficacies for new (left) and returning students. This information exists for a large number of meaningful student segments for various types of institutions. There are more detailed metrics to assist in resource allocation and investment optimization.

| subpopulation | program_type1 | program_type2 | Overall lift in persistence |
|---|---|---|---|
| 0_terms_completed | academic | computer lab | 1.11% |
| | | course combination | 3.29% |
| | | course redesign | -1.20% |
| | | deved | 3.08% |
| | | faculty training | 5.01% |
| | | fye | 0.21% |
| | | si | 6.07% |
| | | tutoring | 5.20% |
| | employment | career planning | 2.22% |
| | | employer course | 1.55% |
| | fin aid | scholarship | 5.03% |
| | holistic advising | advisor meeting | 7.08% |
| | | conversation | 3.38% |
| | non-academic | engagement | 5.93% |
| | | mentorship | 5.24% |
| | | volunteer | 1.07% |
| | student experience | campus housing | 0.72% |
| | | event attendance | 3.62% |
| | | fye | 2.48% |
| | | learning community | 6.36% |
| | | recreation center | 2.07% |
| | | student org | 6.64% |

| | | | |
|---|---|---|---|
| 1_3_terms_completed | academic | computer lab | 1.49% |
| | | course combination | 1.06% |
| | | course recommendation | 1.75% |
| | | course redesign | -2.82% |
| | | deved | -0.85% |
| | | faculty training | 4.46% |
| | | fye | 2.98% |
| | | si | 2.62% |
| | | tutoring | 3.37% |
| | employment | career planning | 1.55% |
| | | employer course | -1.44% |
| | fin aid | scholarship | 2.77% |
| | holistic advising | advisor meeting | 6.25% |
| | | alert | 2.26% |
| | | conversation | 1.41% |
| | non-academic | engagement | 5.71% |
| | | volunteer | 2.37% |
| | student experience | campus housing | 0.82% |
| | | event | 1.68% |
| | | event attendance | 3.38% |
| | | fye | 4.14% |
| | | greek | 5.08% |
| | | learning community | 3.86% |
| | | recreation center | 1.02% |
| | | student support services | 2.84% |
| | | wellness center | 4.25% |

Another option to accelerate the pace of learning is multi-armed bandit intervention design (Vermoral & Mohri, 2005). Here the study population is split into multiple arms consisting of different treatments and control, where control is the business-as-usual (BAU) treatment. N-wise randomization, going beyond the usual pairwise randomization, is helpful in maximizing the pace of learning by virtue of evaluating N-1 treatments concurrently. Multi-armed bandit design has exploration and exploitation phases. During the exploration phase, the impacts of N-1 treatments are evaluated in comparison to the BAU treatment. During the exploitation phase, more students are assigned to more effective treatments, thereby improving the overall impact results. Figure 8 shows the multi-armed design framework.

***Figure 8***
Multi-armed bandit intervention design framework to accelerate the pace of learning in student success




The third step is intentional intervention design. Walton (2014) explores wise psychological intervention as a means of addressing key social psychological blockers. Yeager et al. (2019) show the results of cost-effective scalable behavioral science algorithm-based interventions on large numbers of students. An influence diagram in Figures 4-5 can provide a Bayesian network framework to help with intentional intervention design. The key here is to give strong incentives to students so that they are more likely to comply with recommended calls to action highly personalized to the most elastic impact lever.

The final step is execution, followed by rinse and repeat. Multi-armed bandit intervention design is highly appropriate to accelerate the speed of learning (Agrawal & Goyal, 2012). People in higher education are highly motivated by seeing student success. Even more motivating is the knowledge that they are playing an important role in cracking the mysteries of student success science.

## Change Management Based on Evidence for Continuous Improvement

The process of building and activating a student success knowledge base will require not only the methodological building blocks described above, but also commitment and leadership. Despite accreditors' mandates for both program and co-curricular assessment in higher education, mantras about being data-informed, and significant institutional investments in data infrastructure and analytics capacity, there often remains a shortage of impact measurement and a gap between the proliferation of potentially useful data and meaningful action informed by that data. It takes purposeful change management and alignment of people, processes, and technology to create a data-informed student-success culture. Some of the most common challenges that need attention in transforming an institution in this way include:

- *Intervention fatigue.* More is not necessarily better with interventions, particularly when resources are limited and the window of opportunity with students is short. Yet there is enormous pressure to continue adding interventions to make gains in student success metrics. Creating an inventory of student success programs is an important first step in establishing clarity of purpose and design, identifying duplicative efforts, planning for measurement, and beginning to parse out effects.
- *Fear of measurement.* The onus is on leaders to create an institutional climate where measurement of interventions is seen as an opportunity to learn and improve rather than an existential threat (Lakos & Phipps, 2004). In our healthcare experience, the single most effective strategy to improve patient care was the benchmarking of patient-care programs, where rigorous impact analysis results were used to collaborate on how to best improve patient care from a systems perspective, leading participants to a positive-sum mentality.
- *Lack of a positive-sum mentality*: While rigorous measurement is required for optimal resource allocation, it's not necessary to hack away at every intervention or program that shows a smaller impact. In our study of over 1500 student success initiatives across more than 60 institutions of all types and sizes, we found that 40% had little to no overall impact. However, each of those "ineffective" programs was impactful for approximately 16% of the student segments in drill-down impact results. Institutions can use data like these to optimize those interventions by targeting them toward students for whom they are effective, while also perhaps making programmatic adjustments to increase effectiveness for other student segments. Improving existing interventions reduces the need to add new interventions and is, therefore, an antidote to intervention fatigue.
- *Siloed efforts and territoriality.* Recognizing student success as a shared responsibility is an important part of creating a culture of care. Realizing that shared responsibility, however, requires that key functional players from across units coordinate, collaborate, and even share resources. Collaborative development and use of a student success knowledge base are a solid foundation for cross-functional teams, empowering them to identify the presenting issues of students they collectively serve, match them to the indicated interventions, and coordinate service delivery to maximize impact without duplication of efforts.
- *Putting programs before students*. Every program or initiative on campus is someone's "baby" and sometimes the needs of the program overshadow the needs of students. "Best-practicism," the blind adoption of popularized strategies without evidence of effectiveness accounting for heterogeneities in population and environment, is another manifestation of a program-centric rather than a student-centric approach tailored to an institution. A student success knowledge base ensures that the opportunity to positively impact students is the driving force of interventions.

From our consulting engagements with numerous higher-ed institutions, we have seen these practice patterns emerge as catalysts for successful stakeholder engagement and institutional transformation.

- *Start small, scale fast*: Early success begets more success. Identify low-hanging fruits and work on them with a clear measurement plan. Use anecdotes and quotes from participating students to attract more similar students who can benefit. Communicate measurement results with the broader campus community to raise awareness of resources and increase appropriate referrals in addition to campaigns and targeted efforts.

- *Invest in behavioral science in crafting alert messages*: Many institutions use human- or auto-generated alerts, Too frequently, alert messages sent to students are transactional, incorporating very little of behavioral science. Our internal impact analyses of 49 alert-triggered interventions show that impact results are largely negative with 44 out of 49 interventions generating statistically significant negative results. An example of what gets sent to students when they miss N consecutive classes is: "You missed N consecutive classes. You are in danger of failing." From a behavioral science or any perspective, that is a terrible thing to say to students when they are most vulnerable or when early-alert systems do not have a complete picture of what is going on. Paunesku et al (2019) found similar results on the opposite side when they administered mind-set interventions to academically underachieving students.
- *Leverage conflict:* Conflicts can be a great catalyst for institutional change. Honest conversations around educational and operational goals can turn conflicts into opportunities for hard conversations and deeper collaborations--particularly with vital issues such as closing achievement gaps and diversity, equity, and inclusion outcomes.
- *Positive sum, not zero sum*: Everyone wins from student success. In healthcare, benchmarking was the most effective way to collaborate and improve as long as everyone embraced system-level performance improvement as the goal regardless of where a department was.
- *Human relationship building*: We have seen great success whenever interventions included human relationship building (Colver, 2018). Invest in teaching soft skills to students and those who work with students every day. Faculty have a large influence on student success. Their timely and empathetically grounded nudges can go a long way in helping students feel that they belong and that they can thrive.

## Conclusion

The world of education is going through a major transformation. Hao (2019) discusses a grand experiment in China, where Artificial Intelligence (AI) tutors and human teachers work together to accelerate learning and usher in a new era of learning analytics. Baer, Hagman, and Kil (2020) describe how AI and human intelligence can complement each other in preventing an AI winter of disillusionment and creating a virtuous cycle of student success from both cross-industry and pragmatic perspectives with case studies.

Machine learning and human creativity can complement each other. While machine learning can discover and recommend interesting patterns and opportunities for student success interventions, student success people need to be willing partners in maximizing synergy. Indeed, the future of this work will be a thoughtful interplay of data strategy, domain expertise, and design thinking all aimed at helping more, and more diverse students learn well and finish strong.

## References

Agrawal, S., & Goyal, N. (2012). Analysis of Thompson sampling for the multi-armed bandit problem. *Conference on Learning Theory*, *23*(39), 1–25.

Baer, L., Kil, K., & Hagman, A. (2019). Sherlock Holmes redux: Putting the pieces together. In L. Bare & C. Carmean (Eds.), *An analytics handbook: Moving from evidence to impact*. Society for College and University Planning.

Baer, L., Hagman A., & Kil, D. (2020). Preventing a winter of disillusionment: Artificial intelligence and human intelligence in student success. *Educause Review*, *55*(1), 47–54

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal*, *312*(5), 1215–1218. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2350940/pdf/bmj00541-0043.pdf

Brookhart, M. A., et al (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, *163*(12), 1149–1156.

Caron, M., Choquet-Kastylevsky, G., & Joubert-Caron, R. (2007). Cancer immunomics using autoantibody signatures for biomarker discovery. *Molecular & Cellular Proteomics, 6*(7), 1115–1122.

Chow, N., Gallo, L., & Busse, J.W. (2018). Evidence-based medicine and precision medicine: complementary approaches to clinical decision-making. *Precision Clinical Medicine, 1*(2), 60–64.

Colver, M. (2018). *The lifecycle of sustainable analytics: From data collection to change management.* Office of Student Analytics, Utah State University.

Csikszentmihalyi, M. (2009). *Flow: The psychology of optimal experience*. Harper Collins.

Dowd, A. (2005). *Data don't drive: Building a practitioner-driven culture of inquiry to assess community college performance.* [Report]. Lumina Foundation for Education Research. https://files.eric.ed.gov/fulltext/ED499777.pdf

Eyal, N. (2014). *Hooked: How to build habit-forming products*. Penguin Group.

Franklin, J. M., Glynn, R.J ., Martin, D., & Schneeweiss, S. (2019). Evaluating the use of nonrandomized real-world data analyses for regulatory decision making. *Clinical Pharmacology & Therapeutics*, *105*(4), 867–877.

Gagliardo, J., & Turk, J. (2017). *The data-enabled executive: Using analytics for student success and sustainability*. American Council on Education.

Hibbard, J., Stockard, J., Mahoney, E., & Tusler, M. (2004). Development of the patient activation measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Services Research*, *39*(4p1), 1005–1026.

Hao, K. (2019, Aug 2). China has started a grand experiment in AI education. It could reshape how the world runs. *MIT Technology Review.* https://medium.com/mit-technology-review/china-has-started-a-grand-experiment-in-ai-education-it-could-reshape-how-the-world-learns-c771f5980deb

Hill J. & Su, Y. (2013). Assessing lack of common support in causal inference using Bayesian nonparametrics: Implications for evaluating the effect of breastfeeding on children's cognitive outcomes. *Annals of Applied Statistics*, *7*(3), 1386–1420.

Jones P., & Agusti, A. (2006). Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *European Respiratory Journal*, *27*(4), 822–832.

Kil, D., Shin, F., & Pottschmidt, D. (2004). *Propensity score primer with extensions to propensity score shaping and dynamic impact analysis.* Paper presented at the Disease Management Association of America Conference, Orlando, FL.

King G., & Nielsen, R. (2019) Why propensity scores should not be used for matching. *Political Analysis, 27*(4). https://doi.org/10.1017/pan.2019.11

Lakos, A., & Phipps, S. (2004). Creating a culture of assessment: A catalyst for organizational change. *Libraries and the Academy*, *4*(3), 345–361. https://escholarship.org/uc/item/0843106w

Leacy, F., & Stuart, E. (2014). On the use of propensity and prognostic scores in estimation of the average treatment effect on the treated: A simulation approach. *Statistics in Medicine, 33*(20), 3488–3508.

Mandinach, E., & Jackson, S. (2012). *Transforming teaching and learning through data-driven decision making.* Corwin and SAGE.

McCall, N., Cromwell, J., Urato, C., & Rabiner, D. (2008). *Evaluation of phase 1 of the Medicare Health Support Pilot Program under traditional fee-for-service medicare: 18-month interim analysis.* Report to Congress, CMS Contract No. 500-00-0022.

Milliron, M., Kil, D., Malcolm, L., & Gee, G. (2017). From innovation to impact: How higher education can evaluate innovation's impact and more precisely scale student support. *Planning for Higher Education Journal*, 45N4.

National Academies of Sciences, Engineering, and Medicine. (2019). *Examining the impact of real-world evidence on medical product development*: *Proceedings of a workshop series.* The National Academies Press.

Oreopoulos, P., & Petronijevic, U. (2019). *The remarkable unresponsiveness of college students to nudging and what we can learn from it.* Working Paper 26059. NBER.

Owens, D., Shachter, R., & Nease, R. (1997). Representation and analysis of medical decision problems with influence diagrams. *Medical Decision Making*, *17,* 241–252

Paunesku, D., et al (2019). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784–793

Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. (2011). *Handbook of educational data mining*. CRC Press.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55

Vermorel, J., & Mohri, M. (2005). Multi-armed bandit algorithms and empirical evaluation. *ECML*, LNAI 3720, 437–448

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Associatiion, 113*(523), 1228–1242

Walton, G. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*(1), 73–82.

WHO. (2002). *The importance of pharmacovigilance: Self-monitoring of medicinal products.* Author.

WWC (2019, Nov). *What Works Clearinghouse standards handbook version 4.1*. http://ies.ed.gov.

Yeager, D., & Walton, G. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301

Yeager D., et al, (2019, Aug). A national experiment reveals where a growth mindset improves achievement. *Nature*. https://doi.org/10.1038/s41586-019-1466-y

## Author Notes

**David Kil**
Chief Data Scientist, Civitas Learning
david.kil@civitaslearning.com

**Angela Baldasare**
Senior Principal Strategic Consultant, Civitas Learning
angela.baldasare@civitaslearning.com

**Mark Milliron**
SVP & Executive Dean of the Teachers College, Western Governors University
mark.milliron@wgu.edu

## Guest Editor Notes

**Sean M. Leahy, PhD**
Arizona State University, Director of Technology Initiatives
sean.m.leahy@asu.edu

**Samantha Adams Becker**
Arizona State University, Executive Director, Creative & Communications, University Technology Office; Community Director, ShapingEDU
sam.becker@asu.edu

**Ben Scragg, MA, MBA**
Arizona State University, Director of Design Initiatives
bscragg@asu.edu

**Kim Flintoff**
Peter Carnley ACS, TIDES Coordinator
kflintoff@pcacs.wa.edu.au

Current Issues in Education
Mary Lou Fulton Teachers College • Arizona State University
PO Box 37100, Phoenix, AZ 85069, USA

Volume 22, Issue 1 January 7, 2021 ISSN 1099-839X