RETHINKING REINVENTING RELEARNING

Please Be eAdvised

Netflix meets Google meets academia. How **data mining** is reshaping the college experience.

Likes: tattoos, photography, education.

Swiped ID card, Starbucks, 2:01 p.m. and 4:44 p.m.

Psych major. She's got mail: Take statistics now.

Likes: tattoos, photography, education.

PHOTO ILLUSTRATION BY THE NEW YORK TIMES; PHOTO BY ARIZONA STATE UNIVERSITY

25 math lessons to go before Monday. Get cracking!

Swiped ID card, Starbucks, 2:02 p.m. and 4:43 p.m. Are you friends?

Likes: tattoos, photography, education.

By Marc Parry

CAMPUSES are places of intuition and serendipity: a professor senses confusion on a student's face and repeats his point; a student majors in psychology after a roommate takes a course; two freshmen meet on the quad and eventually become husband and wife. Now imagine hard data substituting for happenstance.

As Katye Allisone, a freshman at Arizona State University, hunkers down in a computer lab for an 8:35 a.m. math class, the Web-based course watches her back. Answers, scores, pace, click paths — it hoovers up information, like Google. But rather than personalizing search results, this data shapes Ms. Allisone's class according to her understanding of the material.

With 72,000 students, A.S.U. is both the country's largest public university and a hotbed of data-driven experiments. One core effort is a degree-monitoring system that keeps tabs on how students are doing in their majors. Stray off-course and a student may have to switch fields.

And while not exactly matchmaking, Arizona State takes an interest in students' social lives, too. Its Facebook app mines profiles to suggest friends. One classmate shares eight things in common with Ms. Allisone, who "likes" education, photography and tattoos. Researchers are even trying to figure out social ties from anonymized data culled from swipes of ID cards around the Tempe campus.

This is college life, quantified.

Data mining hinges on one reality about life on the Web: what you do there leaves behind a trail of digital breadcrumbs. Companies scoop those up to tailor services, like the matchmaking of eHarmony or the book recommendations of Amazon. Now colleges, eager to get students out the door more efficiently, are awakening to the opportunities of so-called Big Data.

The new breed of software can predict how well students will do before they even set foot in the classroom. It recommends courses, Netflix-style, based on students' academic records.

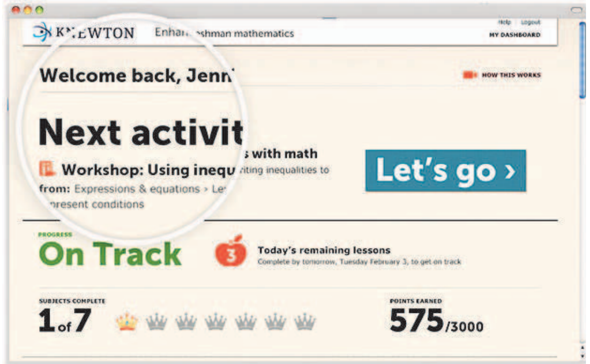
Data diggers hope to improve an education system in which professors often fly blind. That's a particular problem in introductory-level courses, says Carol A. Twigg, president of the National Center for Academic Transformation. "The typical class, the professor rattles on in front of the class," she says. "They give a midterm exam. Half the kids fail. Half the kids drop out. And they have no idea what's going on with their students."

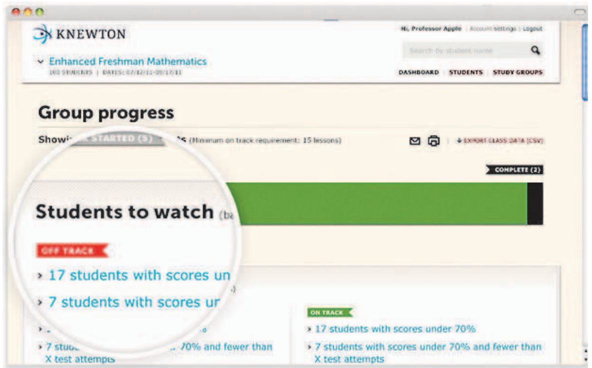
As more of this technology comes online, it raises new tensions. What role does a professor play when an algorithm recommends the next lesson? If colleges can predict failure, should they steer students away from challenges? When paths are so tailored, do campuses cease to be places of exploration?

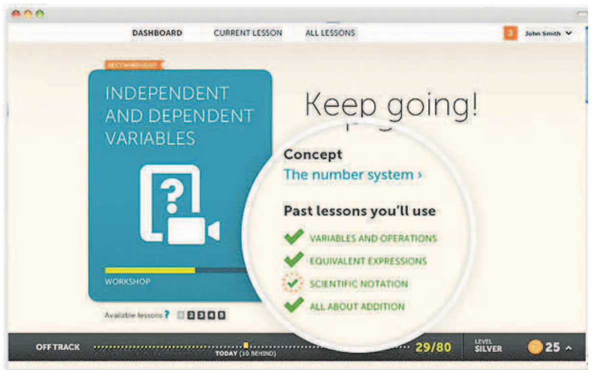
"We don't want to turn into just eHarmony," says Michael Zimmer, assistant professor in the School of Information Studies at the University of Wisconsin, Milwaukee, where he studies ethical dimensions of new technology. "I'm worried that we're taking both the richness and the serendipitous aspect of courses and professors and majors — and all the things that are supposed to be university life — and instead translating it into 18 variables that spit out, 'This is your best fit. So go over here.' "

ALERT! YOU ARE OFF-TRACK

Ever since childhood, Rikki Eriven has felt certain of the career that would fit her best: working with animals. Specifically, large animals. The soft-spoken freshman smiles as she recalls the episode of "Animal Planet" that kindled this interest, the one about zoo specialists who







What's the next math lesson? Who's falling behind? The software knows all.

treat rhinos, hippos and giraffes. So when Ms. Eriven arrived at Arizona State last fall, she put her plan in motion by picking biological sciences as her major.

But things didn't go according to plan. She felt overwhelmed. She dropped a class. She did poorly in biology (after experiencing problems, she says, with the clicker device used to answer multiple-choice questions in class). Ms. Eriven began seeing ominous alerts in her e-mail inbox and online student portal. "Off-track," they warned. "It told me that I had to seek eAdvising," she says. "And I was, like, eAdvising?"

Yes, eAdvising. Universities see such technology as one answer to a big challenge. On average, only 31 percent of students at public colleges get their bachelor's degree within four years, and 56 percent graduate within six years. Such statistics have come under greater scrutiny as parents and politicians demand accountability from colleges. Tennessee, for example, doles out higher education dollars in part by measuring how effective an institution is at graduating students.

Yet some students show up with ambitions that bear no relation to their skills. Or parents push them into a major that doesn't interest them. Or they feel like shoppers in a cereal aisle, confounded by the choices.

At Arizona State, which has more than 250 majors, the old system let students explore without much structure. A student could major in engineering to please his parents, only to pack his schedule with "Chinese Thought" and music, says Elizabeth D. Capaldi, the provost. No longer. Technology has redrawn the road map.

Under Arizona State's eAdvisor system — in use from 2008-9 and based on a similar effort at the University of Florida — students must pick a major freshman year and follow a plan that lays out when to take key courses. (Students can still study broadly, by choosing from five "exploratory" majors, like "arts and humanities" or "science and engineering," and staying in them for 45 credits.) If they fail to sign up for a key course or do well enough, the computer cracks a whip, marking them "off-track." Wander off-track two semesters in a row, and a student may have to change majors.

If that sounds harsh, there's a rationale: One way to ensure that students will reach the finish line is to quickly figure out if they've selected a suitable track. So the A.S.U. system front-loads key courses. For example, to succeed in psychology, a student must perform well in statistics.

"Kids who major in psych put that off, because they don't want to take statistics," Ms. Capaldi says. "They want to know: Does their boyfriend love them? Are they nuts? They take all those courses, then they hit statistics and they say: 'Oh, God, I can't do this. I can't do experimental design.' And so they're in the wrong major. By putting those courses first, you can see if a student is going to succeed in that major early." Arizona State's retention rate rose to 84 percent from 77 percent in recent years, a change Ms. Capaldi credits largely to eAdvisor.

For students who run off-track, the outcome can sting. Ms. Eriven was shocked to learn she would have to change her major after the system flagged her. She cried, called her mother, and recalibrated her plans. In a meeting with an adviser, she detailed her interests. She likes science. She is family-oriented, interested in music, and good at writing. The adviser suggested a few possible majors, including psychology, family and human development, and creative writing.

Writing. It would involve only a couple of classes each semester. She could still take science and, hopefully, switch back to biology. So that's what she chose. "I didn't really have, like, a backup plan," Ms. Eriven says.

But what if you could rewind this story and shape a student's path before reaching such a crossroads?

YOU WILL PASS (OR NOT)

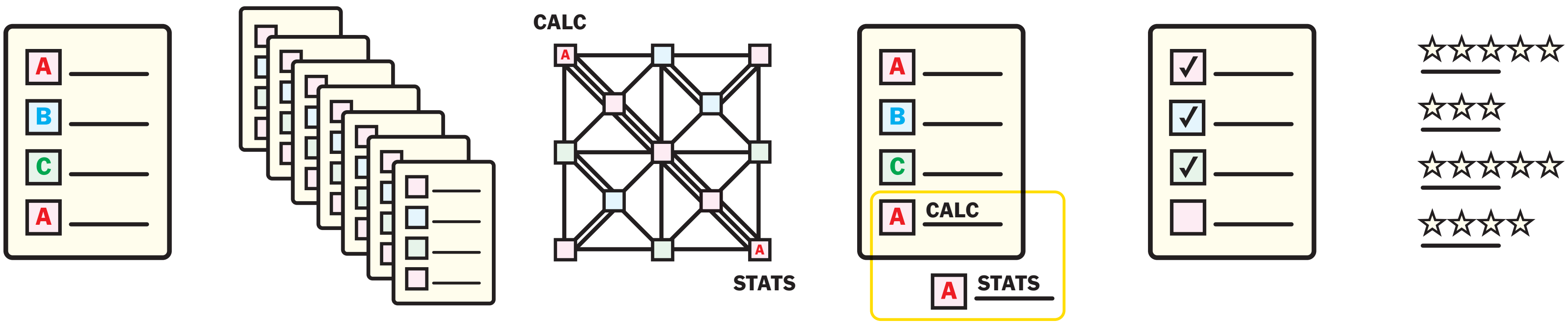
When Adam Lange began working full time at Rio Salado College in 2008, he was still an undergraduate at nearby Arizona State, a 22-year-old computer science major with a budding obsession with data. Over time, that obsession would shape the learning experience for thousands of students — and drive his fiancée bonkers.

Mr. Lange's idea of fun is converting his home into a surveillance lab. He outfitted his cat Sammy, who has an eating disorder, with a device that is read by a scanner every time the cat cranes his neck over the bowl. Mr. Lange monitors the logs and feeds Sammy a treat if he hasn't eaten. He also rigged a webcam next to his fish tank, logging the coordinates of his Betta fish several times a second to find out what common paths it takes and how far it travels (90 feet in one hour!). At Rio Salado, a com-

This article is a collaboration between The New York Times and The Chronicle of Higher Education, a daily source of news and opinion for professors, administrators and others interested in academe. Marc Parry is a technology reporter for The Chronicle.

Advising by Algorithm

At Austin Peay State University in Tennessee, a program called "Degree Compass" provides students with a customized list of course recommendations based on degree requirements as well as predicted grades. Here is how the lists are generated.



Your records ▶
The system has access to your previous college grades, your high school grade-point average and all of your standardized test data, including ACT and SAT scores.

Grade database ▶
Software searches transcripts of students who have taken classes or tests in common with you. The database includes more than 500,000 grades.

Grade correlations ▶
Every grade in the database is connected to grades in other courses. The system finds correlations. Grades in calculus, for example, have a strong correlation with grades in statistics.

Predicted grades ▶
To predict your success in a given course, the software combines your grades in correlated classes with other grades and scores you have received. The more correlations, the more weight in the prediction.

Requirements ▶
The system knows the requirements for your major and the core curriculum, and knows which of these you have already fulfilled.

Recommendations
Courses are given a star rating based on your predicted grade as well as their importance to requirements.

Source: Tristan Denley, Austin Peay State University

THE NEW YORK TIMES

munity college with about 70,000 students, 43,000 of them online, Mr. Lange got excited about the behavioral data students leave behind: the vast wake of clicks captured by software that runs Web courses. Records of when they logged in, opened a syllabus, turned in homework — all of it just sitting there. Could you mine it to model patterns of students who succeeded in the past? Use that to identify current ones likely to fail? And then help those students? Many educators are now asking similar questions.

Mr. Lange and his colleagues had found that by the eighth day of class they could predict, with 70 percent accuracy, whether a student would score a "C" or better. Mr. Lange built a system, in 2009, that sent professors frequently updated alerts about how well each student was predicted to do, based on their course performance and online behavior.

To Mr. Lange, the underlying math doesn't differ much from what he might deploy in his fish espionage. Say the Betta makes two consecutive movements side to side, and then swims upward 85 percent of the time. In the future, if the fish moves left and then right, Mr. Lange can say with confidence that he'll then swim up. Similarly, Rio Salado knows from its database that students who handed in late assignments and didn't log in frequently often fail or withdraw from a course. So the software is more likely to throw up a red flag for current students with those characteristics.

> When students' paths are so tailored, campuses may cease to be places of exploration.

"There's a predictability about the fish," says Mr. Lange, now 26 and working for Ellucian, a higher-education software company. "The same concept applies to students."

Still, once you identify students in need of extra assistance, how do you help them?

Rio Salado has experimented with various intervention strategies, so far with mixed results. And in a cautionary tale about technical glitches, the college began sharing grade predictions with students last summer, hoping to prod those lagging behind to step up, but had to shut the alerts down in the spring. Revisions to courses had skewed calculations, and some predictions were found to be inaccurate over a period of about five days. An internal analysis found no surge in the number of students dropping classes. An improved system is promised for fall.

YOU MAY ALSO LIKE . . .

Austin Peay State, a midsize university about 45 minutes northwest of Nashville, takes the algorithmic approach to higher education one step further. Before students register for classes, a robot adviser assesses their profiles and nudges them to pick courses in which they're likely to succeed.

The project is the work of Tristan Denley, a programmer turned math professor turned provost. Mr. Denley's software borrows a page from Netflix. It melds each student's transcript with thousands of past students' grades and standardized test scores to make suggestions for every student. When students log into their online portal, they see 10 "Course Suggestions for You" ranked on a five-star scale. For, say, a health and human performance major, kinesiology might get five stars (as the next class needed for her major). Physics might also top the list (to satisfy a science requirement in the core curriculum).

Behind those recommendations is a complex algorithm, but the basics are simple enough. Degree requirements figure in the calculations. So do classes that can be used in many programs, like freshman writing. And it bumps up courses for which a student might have a talent, by mining their records — grades, high school grade-point average, ACT scores — and those of others who walked this path before.

"We're steering students toward the classes where they are predicted to make better grades," Mr. Denley says. The predictions, he adds, are within about half a letter grade, on average.

The prediction process is far more subtle than getting a suggestion to watch "Goodfellas" because you liked "The Godfather." Take the hypothetical health major encouraged to take physics. The software sifts through a database of hundreds of thousands of grades other students have received. It analyzes the historical data to figure out how much weight to assign each piece of the health major's own academic record in forecasting how she will do in a particular course. Success in math is heavily predictive of success in physics, for example. So if her transcript and ACT score indicate a history of doing well in math, physics would likely be recommended over biology, though both satisfy the same core science requirement.

Mr. Denley points to a spate of recent books by behavioral economists, all with a common theme: People find it difficult to make wise choices when there are many options and little information. The same goes for college students trying to construct a schedule, he says. They know they must take a social science class, but they don't know the implications of taking political science vs. psychology vs. economics. They choose based on course description or to avoid having to wake up for an 8 a.m. class on Monday. Every year, students in Tennessee lose their state scholarships because they fall a hair short of the G.P.A. cutoff, Mr. Denley says, a financial swing that "massively changes their likelihood of graduating."

"When students do indeed take the courses that are recommended to them, they actually do substantially better," he says. And take them they do. Last fall, 45 percent of classes on student schedules were from top-10 recommendations, 57 percent from their top 15. Though these systems are in their infancy, the concept is taking hold. Three other Tennessee colleges have adopted Mr. Denley's software, and some institutions outside the state are developing their own spins on the idea.

Some express concerns about deferring such important decisions to algorithms, which have already come to dictate — and limit — so much of what we see and do online. Mr. Zimmer, the Milwaukee information-studies professor, sees the value in preventing students from going down paths that may frustrate them or cause them to quit college. But as higher education gets more efficient, he fears the loss of the unanticipated discovery.

"It's the same as if you're worried about whether or not Google or Amazon are going to present you with alternative topics, or only the topics that fit your history," he says. "We hope the role of a university is to make sure people are exposed to diverse things and challenged."

IN THE CLASS: DIRECTION THROUGH DATA

At Arizona State, algorithms figure in course content, too. Thousands of A.S.U. students now take math courses through a system that mines performance and behavioral data, building a profile on each user and delivering recommendations about what learning activity they should do next. The system, created by a start-up company called Knewton, gave the university a fresh way of addressing the continuous problem of students being unprepared for college math. But it also offers a glimpse into what many more students will experience as teaching increasingly shifts from textbooks and lectures that feed the same structure of information to a class of 300, regardless of individual expertise, to machines that study their users' learning patterns and adapt to them.

That excites some educators. George Siemens, a data-mining expert at the Canadian distance-learning university Athabasca, calls the traditional approach an inefficient model "that generates a fair degree of dropouts."

Knewton dismantles that model. Ms. Allisone's 8:35 a.m. class is not a lecture. Although students are supposed to show up at a fixed time, and an instructor is there to work with them, the action is on screen. Knewton allows Ms. Allisone to skip past some concepts she gets, like factors and multiples. When she struggles with inverting linear functions, the software provides more online tutoring. Two students who complete the same lesson might see different recommendations as to what to do next, based on their proficiency.

As the company develops and works with more data and content — major universities like University of Nevada, Las Vegas, are adopting its technology, as is the publishing giant Pearson — it will tailor instruction more finely. What time of day does a student best learn math? What materials and delivery styles most engage the student? Say you have the same concept explained in a video, in a textbook-like format and in Socratic steps. Knewton will associate a student's "engagement metrics" with those styles and use that to help determine the next step.

But what sounds flashy may be based, at least in part, on flawed assumptions, warns Richard E. Clark, professor of educational psychology and technology at the University of Southern California. He says there is no evidence that there are "visual" learners who benefit from video over text, as Knewton implies. Studies, he says, have shown that "learning styles" are not effective for shaping instruction.

The broader problem with data mining, as Mr. Clark sees it, is that it is seldom done right. Data analysts often make "questionable assumptions" about the meaning of keystrokes, he says. They assume students who are spending the most time on some learning material are most interested in that content, for example. "That assumption may be true when people choose to watch Netflix movies but is not at all the case in many university courses where few choices are available," Mr. Clark says.

Meanwhile, dismantling old models leaves both professors and students adjusting to new roles.

Suzanne Galayda, an Arizona State math instructor, finds it takes longer to penetrate the wall of computer screens and build rapport with students. In her remedial class, they start off feeling uncomfortable asking questions. But even as software elbows her off center stage, it also helps her play her part with far more information — so much data about what students do, and when, that it sometimes surprises them.

"Students don't realize that we're watching them in these classes," she says.

Ms. Galayda can monitor their progress. In her cubicle on a recent Monday, she sees the intimacies of students' study routines — or lack of one — from the last activity they worked on to how many tries they made at each end-of-lesson quiz. For one crammer, the system registers 57 attempts on multiple quizzes in seven days. Pulling back to the big picture, a chart shows 15 students falling behind (in red) and 17 on schedule (in green).

On Wednesday, Ms. Galayda rubs her hands with satisfaction. The bar is mostly green. Mostly. When class meets, she taps her nails on the hard drive of Carolina Beltran's computer. "You were working on it at 4 a.m.," the instructor tells the student.

"Yeah, I mean, like, I sleep. My sleeping schedule is weird," Ms. Beltran stammers.

Arizona State's initial results look promising. Of the more than 2,000 students who took the Knewton-based remedial course this past year, 75 percent completed it, up from an average of 64 percent in recent years.

In Ms. Galayda's experience, students "either love it or hate it."

Ms. Allisone raves. "I learned more in this semester than I have in a year in high school," she says. She praises the clarity and concision of the system's instructional videos, contrasting that with the many teachers who "have issues communicating correctly."

But another freshman, a health sciences major who requested anonymity because she did poorly for two semesters, recalls a downward slide that began when she started falling a couple of lessons behind. That scared her at first, until she talked to her peers. Some were six lessons behind. Twelve, even. How bad could two be? She didn't sweat it. As she juggled social life, work and other classes, math fell through the cracks. She ended up having to retake the course, a case study in the danger of giving self-paced classes to freshmen.

"I like lecture better," she says. "I'm not used to teaching myself. So it was a huge adjustment."

THE SOCIAL NETWORK

These experiments are only the beginning. Colleges will likely dig deeper into the data at their disposal, touching more and more aspects of student life. Already, some researchers are eyeing the next frontier: social life.

Research shows that social ties can be critical to academic success. If students are more integrated into campus life, they're more likely to stay in school. If a friend drops out, they're more likely to as well.

"If the university could model, at a high level, the social network of the college, that would be a very useful data layer," says Matt Pittinsky, who co-founded Blackboard, a company that provides a platform for online classes, and later became an assistant research professor in the sociology program at Arizona State. A university might reach out to a student "who is not showing evidence of social integration," Mr. Pittinsky says, pointing out extracurricular activities and communities that might tie them more deeply to the institution.

Working with computer scientists, Mr. Pittinsky started an academic research project that tiptoes toward a better understanding of social connections. The research team's raw material: anonymous logs from swipes made with Arizona State ID cards. When students use these cards, be it to buy food on campus or access the fitness center, the transaction gets recorded. The question that struck Mr. Pittinsky was whether or not you could infer social ties from those trails.

Say two students swipe within 5 or 10 seconds of each other at different times of day in different contexts. Are they more likely to be friends? And can you predict attrition by pinpointing changes in how a student uses a campus? Say someone goes to Starbucks at 2 p.m. every day before 2:15 p.m. class. Then stops. "If that happens three weeks in a row," Mr. Pittinsky says, "and we're not seeing log-ins into Blackboard, and maybe you've made a request at the registrar to have your transcript sent somewhere, there ought to be an adviser with a really big red flashing light saying, reach out to this student."

The prospect of card-swipe surveillance discomforts Mr. Zimmer. He worries authorities might misuse location data to do things like track foreign students or instigators of a student protest.

But the broader issue of privacy hangs over even less Orwellian efforts to collect and monitor personal data. In his own courses, Mr. Zimmer includes a disclaimer on his syllabus disclosing what he can see through Milwaukee's online-learning platform, including "the dates and times individual students access the system, what pages a student has viewed, the duration of visits, and the IP address of the computer used to access the course Web site."

For his part, Mr. Pittinsky stresses that the card-swipe research is "very focused on the ability to protect anonymity."

As for students, they've never been too fond of adults meddling on Facebook, let alone getting all Big Brother with card swipes. "Creeping on us" is how Ms. Allisone describes the card-swipe project. Ms. Allisone has managed to keep one aspect of her life — she hopes to transfer — from any "creeping." But that, too, may change.

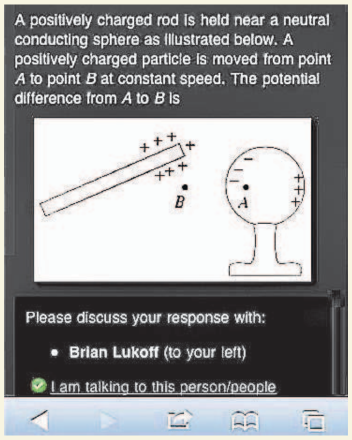
Arizona State monitors requests for transcripts to be sent elsewhere, according to Ms. Capaldi, the provost. "Which," she says, "is kind of sneaky." ■

Group Dynamics

ERIC MAZUR, a Harvard physics professor, has long worked to supplant lectures with more interactive classes. Students, he found, assimilate new material better by working on conceptual problems in class and debating their conclusions with peers. But they tend to pair up with the same friends, which can be unproductive.

Mr. Mazur and his colleagues came up with a novel solution: take students out of the matchmaking. Their software, called Learning Catalytics and now in use at various campuses, is intended to force students to defend their ideas by matching them with classroom partners who have different opinions.

When Merri Su Ruhmann sits down in a graduate seminar on student development theories at the University of Texas, Austin, she "checks in" to her seat on a map of the classroom displayed on her iPad. Then the lecturer, Cassandre Alvarado, poses questions in Learning Catalytics. If there is enough divergence in answers, she clicks a button on her laptop and students are automatically grouped. Ms. Ruhmann obeys her prompt: *Please discuss your response with Jessica Khalaf behind you.*



"It forces them to either have certainty, and to really defend their idea, or it gives them that moment of cognitive uncertainty, which is really powerful for learning," Ms. Alvarado says.

The responses can be educational for Ms. Alvarado, too. At times, she has planned to fly through what seemed like easy questions, only to discover students had major gaps in understanding. "I have data now," she says. "Not just a feeling." *MARC PARRY*

BEN SKLAR FOR THE NEW YORK TIMES

MATCHMAKING University of Texas students are grouped based on their responses to questions, then must defend their answers.