

How Social Science Research Can Improve Teaching

Gary King, *Harvard University*

Maya Sen, *University of Rochester*

ABSTRACT We marshal discoveries about human behavior and learning from social science research and show how these can be used to improve teaching and learning. The discoveries are easily stated as three social science generalizations: (1) social connections motivate, (2) teaching teaches the teacher, and (3) instant feedback improves learning. We show how to apply these generalizations via innovations in modern information technology inside, outside, and across university classrooms. We also give concrete examples of these ideas from innovations we have experimented with in our own teaching.

Humans have theorized about how to teach for thousands of years and update the substance of what we teach almost every year. Yet generations have passed without any major improvements in the procedures and style of teaching in our classrooms. If your great-great-grandparents went to college, they probably sat in a classroom with all the other students facing forward, trying to look attentive, while the professor professed. If you are a professor at a university today, you probably lecture to the same sea of students, all still trying to look like they are paying attention. To be sure, you may use some newer technologies (such as electricity, radio, TV, whiteboards, and PowerPoint slides), you may have added a few group activities, and you perhaps teach a seminar with lots of discussion. But if your ancestors were to walk into a classroom today, they would know where to sit, what to do, and how to act. Our methods of teaching have changed very little.

Education researchers, often in and around schools of education, have written volumes about how to improve teaching and learning. Many of these ideas are extremely promising but too few are firmly established by rigorous empirical research, replicated in different areas (Whitehurst 2010). The problem is not the researchers; the problem is the almost unique (and probably underappreciated) difficulty of doing research in this area. Methodologically, we have a large number of students, but the unit of analysis for a teaching intervention is the professor or class. Thus, any one professor intervening in his or her own classroom has an $n = 1$

study. Although intervening in your own classroom is easy, getting a reasonable sample size with the right unit of analysis is almost impossible and rarely done: imagine the difficulty of explaining to (say) 50 of your colleagues that they and their classes will be assigned (randomly or otherwise) to treatment and control groups to test a hypothesis. For one example, in 75 years of education research on the effects of a variable as important as class size, only one fully randomized large-scale study has ever been done (Chingos 2013)!

We need to look for more opportunities for this type of research within education.¹ But even without it, substantial progress is now possible thanks to research from other fields—fields where large-scale randomized trials, and many other types of rigorous research designs, are possible. Recent developments in social science research mean we know more about how human beings think and learn, which, we show, can be marshaled to improve our teaching. In addition, technology has progressed far past that used in most current classrooms. Although no technology used by itself has any necessary effect on learning (and new teaching technology can often be a distraction), some new technologies make it easier to take advantage of social science insights to improve teaching. Finally, unprecedented societal forces outside the university—including for-profit universities, massive open online courses, commercial and not-for-profit ventures, and the web—are now conspiring to overturn centuries of stable funding models (King and Sen 2013). In every other area of society, one either adjusts to forces like these or gets run over. It is time for those inside universities to pay attention and to use their unique advantage—research—to improve teaching.

In this article, we discuss how social science knowledge and technological innovations can help us teach better. We do this by distilling three principles from social science research: (1) social connections motivate, (2) teaching teaches the teacher, and (3) instant feedback improves learning. We find evidence for these

Gary King is the Albert J. Weatherhead III University Professor and Director of the Institute for Quantitative Social Science at Harvard University. His homepage is at <http://gking.harvard.edu>, and he can be reached at king@harvard.edu.

Maya Sen is an assistant professor of political science at the University of Rochester. Her home page is at <http://mayasen.org>, and she can be reached at msen@ur.rochester.edu.

principles in research from social and cognitive psychology, public health, economics, sociology, and political science. To show how these principles can be used in teaching, we draw from our experience developing and using several inter-related technologies in teaching our class, Harvard's Gov2001: Advanced Quantitative Political Methodology (see <http://j.mp/G2001>). We illustrate these outside, inside, and across classrooms in Section 2.

We conclude in Section 3 by discussing the growing movement in natural and physical science departments to devote some of their own faculty positions and other resources to education research. That scientists are now participating in what is essentially social science research is gratifying, but social scientists have the same needs, our own teaching issues, and more knowledge of the area that we can bring to bear on common problems. It is time we make a contribution, both in terms of research, which we take a step toward in this article, and resources.

1. SOCIAL SCIENCE LEARNING PRINCIPLES

Thanks to advances in the methods of causal inference, huge increases in data collection, and improved theoretical understandings, we have a better understanding of how and why people learn and behave than ever before. From this massive literature, we extract three principles that can be applied to improve teaching.

Principle 1: Social Connections Motivate.

Coaxing individuals to take actions that benefit themselves—such as losing weight, exercising, and not smoking—is often extremely difficult. But getting them to take actions that involve social interaction or that benefit the community—such as recycling or joining the PTA—is often far easier. Social scientists have learned how to use this insight by making individual activities into social activities, thereby increasing the effectiveness of individual-level interventions. For example, the large “get out the vote” literature shows a tiny effect of all types of individual citi-

zen contacts, such as phone calls, in-person visits, or mailings. But studies that add a social component—such as explaining to a respondent which of a person's neighbors have already voted—can increase a person's propensity to vote by as much as 8 percentage points (Gerber, Green, and Larimer 2008).

The same insight applies more widely: we tend to lose and gain weight when our friends do (Christakis and Fowler 2007; VanderWeele 2011). We drink less, exercise more, and smoke less when our friends and associates do (Christakis and Fowler 2008; Rosenquist et al. 2010). Social networks influence what we eat (Pachucki, Jacques, and Christakis 2011), how happy we are (Fowler and Christakis 2008), the probability we end up lonely or depressed (Cacioppo, Fowler, and Christakis 2009), where we live (DiPrete et al. 2011), the kind of health habits we take up (Centola 2010), and whether our marriages persist or end (McDermott, Fowler, and Christakis 2009). Social connections motivate recycling (Burn 1991), influence the importance of attending religious services (Lim and Putnam 2010), and affect many other behaviors and attitudes.

Social connections affect so many aspects of our lives that our argument that they can also be applied to education and learning should be no surprise. It is not only for efficiency that a group of students are all taught together in the same classroom or that elementary schools spend so much time trying to integrate students socially into the class environment. Some research in education provides evidence for this point directly in the context of traditional higher education (Garrison, Anderson, and Archer 1999; Summers and Svinicki 2007) and online education (Barab and Duffy 2000; Dawson 2006; DeSchryver et al. 2009; Graff 2003; Rovai 2003; Shea 2006), where community building has been shown to be of particular importance due to relatively infrequent social interactions.

Principle 2: Teaching Teaches the Teacher.

Of course, social connections can also distract students, detract from a common purpose, and derail lectures. Finding ways of using this powerful tool in a productive way is crucial. We discuss specific implementations in Section 2.

Social psychologists have demonstrated that under normal circumstances, we “mind wander” (i.e., think about subjects other than those in which we are nominally participating) almost half of all our waking hours (Killingsworth and Gilbert 2010; Morse 2012). Although the literature does not include measures of mind wandering while watching university lectures, it is doubtful that the rate is any lower. People also tend to be less happy when mind wandering, which cannot possibly help students learn, to say nothing about teaching evaluations.

So how do we get students to pay more attention? One strategy is to use the fact that social interactions eliminate about half of this effect: when engaged in conversation with others, people's minds wander only about a quarter of the time (Morse 2012). If we can turn the students into teachers—arranging for them to explain what they have learned to others, having them ask questions, debate, persuade, and otherwise engage the subject matter

Of course, social connections can also distract students, detract from a common purpose, and derail lectures. Finding ways of using this powerful tool in a productive way is crucial.

socially—we can capture a great deal more of their attention than would otherwise be possible.

Almost anyone who has taught understands this fact: study a subject yourself and you can learn a great deal. But teach that same subject to someone else and you understand it far better than you ever realized. The person you are teaching will also learn, even if not as much as you learned. That “teaching teaches the teacher” has been demonstrated empirically in many studies (Chi et al. 1994; VanLehn et al. 2007). We believe it is explained, in part, by the difficulty of mind wandering instead of being engaged socially, by being forced to organize thoughts in a more productive way, and by active rather than passive engagement. We give some examples on how to harness this principle.

Principle 3: Instant Feedback Improves Learning.

Suppose you are an athlete practicing to make the Olympic diving team and you arrive for practice on a Saturday. How much would you improve if the coach watched you silently all day and

then gave you a summary of how you did after practice was over? You would learn some, but you would learn a lot more if, as is typical, you received detailed feedback immediately after every dive.

It is the same story with university education: economics, psychological, medical, and educational research demonstrates convincingly that immediate and frequent feedback improves learning (Dihoff, Brosvic and Epstein 2003, 2004; Dubner and Levitt 2006; Hattie and Timperley 2007; Hodder et al. 1989). The more chances you have to try and fail, the quicker you will master the skill. Implementing this advice involves frequent evaluation: like in science in general, students learn more when they have the chance to be proven wrong. This involves eliminating waiting periods before questions can be answered, understanding the limits of their knowledge, and encouraging students to ask questions as soon as they hit a stumbling block. Requiring them to wait until office hours, section, or the next class should not be part of the drill.

2. IMPLEMENTING THE PRINCIPLES

We now give some ways of combining the social science principles outlined in Section 1 with innovations in information technology. We do so outside (Section 2.1), inside (Section 2.2), and across (Section 2.3) classrooms. The technologies we describe are those we have developed or tried ourselves, but they represent only a few of the possible applications of the principles.

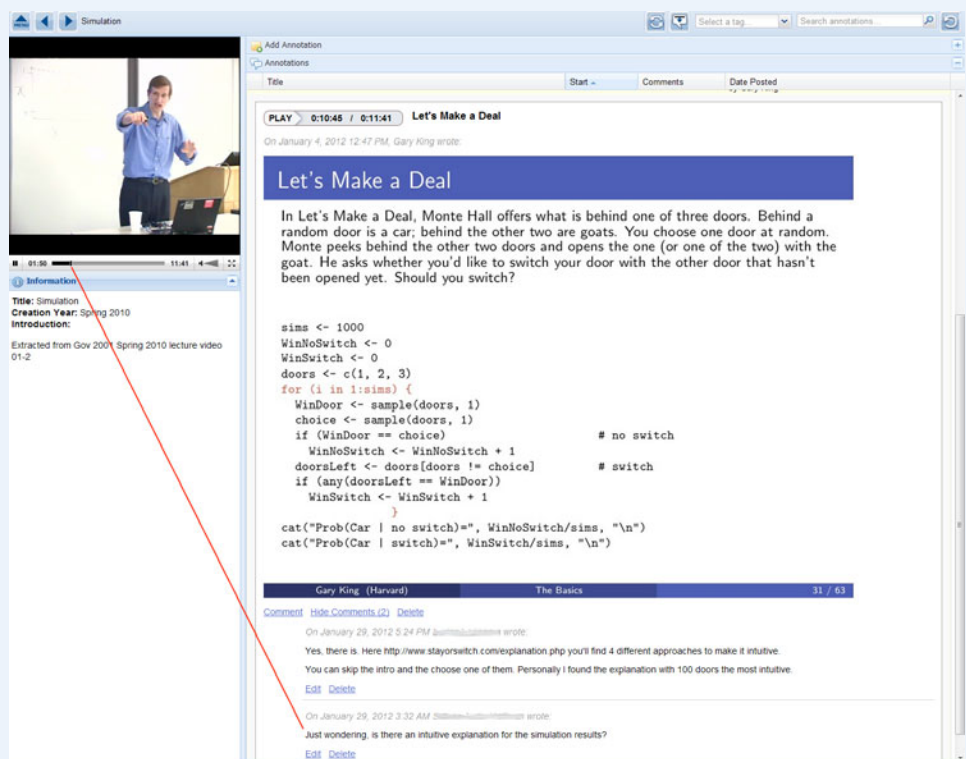
2.1. Outside the Classroom

Here we give examples of three innovations, each of which takes advantage of the principles described earlier. In all cases, we seek to make the class and its social connections continue throughout the week until the next classroom experience.

Making Lectures into Interactive Homework. Putting a university lecture together incurs significant start-up costs for instructors: getting the material together, writing slides, preparing the final presentation, and more. The good news is that after the lecture is written, the marginal costs associated with repeated presentations are low. The bad news is that small yearly improvements result in the same lectures being presented over and over again, or the lectures improving but learning not so much. This situation, combined with the fact that lectures today are often videotaped, disincentivizes students to come to class, pay attention, and learn.

As an alternative, we assign portions of the lecture videos as homework, using an open-source collaborative video technology

Figure 1
Collaborative Video Annotation



The lecture video on the left, slides on the top right, and discussion forum at the bottom right can all be resized. The solid line illustrates, for the purpose of this article, the connection between a point on the timeline and the comments. (Color online)

that we helped develop. This system, which was created by the Harvard University Academic Technology Department, has been released open-source and can now be used by instructors around the world.² (Commercial analogues exist as well—e.g., echo360 and others.) This video Collaborative Annotation Tool (CAT) has at least three benefits. First, with CAT, students can hit “rewind” as often as they like. Because social connections motivate, students rarely stop the professor to ask questions in class, even when it would be beneficial: students do not want to be seen by their peers as not paying attention, not understanding the material, or disrespecting the professor or other students, and so they sit quietly, trying to appear attentive. Because so much time is spent mind wandering, a live lecture can sometimes be described as little more than a sequence of missed opportunities. Collaborative annotation outside of class can help change this.

Second, if rewind does not help, a student can stop the CAT playback and annotate the timeline of the video or one of the associated slides (that turn in sync with the video) with a question or comment. That is, the students can literally pause the video to write in clarifying or substantive questions that correspond with what the instructor is saying at *that exact moment* (see figure 1). Other students, motivated by their social connections, then help clarify, as can the teaching staff. In our experience, a lively, Facebook-style discussion about the material then often develops (often during those late-night hours that federal regulations require faculty to be sleeping). Because, in our experience, students are highly motivated to provide feedback to their peers in near real

Figure 2
Collaborative Text Annotation

22 2 Conceptualizing uncertainty and inference

distinguish between these two cases, I refer to the hypothetical parameter value as $\hat{\theta}$ and the single unobserved true value as θ . In the next chapter, I will introduce $\hat{\theta}$ as a point estimator for θ , based on the maximum of the likelihood with respect to $\hat{\theta}$. $\hat{\theta}$ is a number in a single experiment, but a random variable across hypothetical experiments.

The likelihood that a hypothetical model (summarized by the hypothetical parameter value $\hat{\theta}$) produced the data we observe, given \mathcal{M}^* , is denoted $L(\hat{\theta}|y, \mathcal{M}^*)$, where \mathcal{M}^* may again be suppressed since it appears in all subsequent expressions. The likelihood axiom then defines this concept as follows:

$$L(\hat{\theta}|y, \mathcal{M}^*) \equiv L(\hat{\theta}|y) \tag{2.5}$$

$$= k(y) \Pr(y|\hat{\theta})$$

$$\propto \Pr(y|\hat{\theta}).$$

In the second line of this equation, $k(y)$ is an unknown function of the data; since it is not a function of $\hat{\theta}$, it is treated as an unknown positive constant. In the third line, “ \propto ” means “is proportional to.” The third line is only a more convenient way of writing the second without the constant. For a given set of observed data, $k(y)$ remains the same over all possible hypothetical values of $\hat{\theta}$. However, $k(y)$ is a function of y and therefore may change as y changes. The likelihood $L(\hat{\theta}|y)$ is similar to the concept of inverse probability in that it permits one to measure and compare the uncertainty one has about alternative hypothetical values of $\hat{\theta}$. However, the unknown value $k(y)$ ensures that likelihood is a relative rather than an absolute measure of uncertainty. This likelihood axiom is but one way to make the measure explicitly relative. Indeed, one could use any monotonic function of $\Pr(y|\hat{\theta})$.⁵ The choice represented in Equation (2.5) is arbitrary, just as is the choice of making the scale of probability range between 0 and 1. The advantage of likelihood is that it can be calculated from a traditional probability, whereas inverse probability cannot be calculated in any way.

If the data are continuous rather than discrete, the likelihood is calculated in the same way, except that the underlying probability distribution is now a density. Hence, a more general way to write the formula is as follows:

$$L(\hat{\theta}|y, \mathcal{M}^*) \equiv L(\hat{\theta}|y) = k(y) f(y|\hat{\theta}) \tag{2.6}$$

$$\propto f(y|\hat{\theta})$$

(a) (b)

1 thread on page 2
2 I'm trying to understand what "under essentially the same conditions" means. When we say

1 thread on page 3
2 If the meaning of "model" is the same here as it was in the first chapter, this appears to

3 threads on page 4
3 According to an earlier statement in this book, that exact probability should be known to
2 Likelihood is a measure of relative uncertainty, but relative to what? If it is the relati
2 I'm a bit uncomfortable about this apparent interchangeability of "M" as a model and "M" a

2 threads on page 6
1 Why is this always our null hypothesis then? Does it make sense to have more sophisticated
2 how was this equation derived?

1 thread on page 7
1 Yes! I am always answering this question while teaching stat to others.

1 thread on page 8
1 I don't see why it's necessarily impossible. For example, you could know that in the past

4 threads on page 9
1 Shouldn't it be "is NOT a number?"
1 I'm not really clear on what $k(y)$ "is", besides some unknowable constant. Is it $\Pr(-\theta$
1 I have a conflict between my understanding (or misunderstanding...) of likelihood inferenc
2 What if it's not monotonic?

Shouldn't it be "is NOT a number?" ... 29 Jan, 08:27PM
Ready Act

I think so. ... 29 Jan, 01:22AM
Ready Act

It's a number, because once you've run the single experiment, it just "is" whatever value came from that experiment. It's no longer a variable, it's a constant thing that you know - a number. ... 29 Jan, 03:08AM
Ready Act

Maybe my English deceives me, but it seems like, it should be either:
Ready Act

"is NOT a number, but a variable" or
"is a number, but NOT a variable".
So, do you mean, Jesse, that it's the latter? ... 29 Jan, 11:24AM
Ready Act

Thanks for your post, Jesse. The structure of the sentence confused me, too, Larisa. I now think it means that there are two cases: in the case of a single experiment, it is a number. In the case of multiple experiments (not sure why they're "hypothetical"), it's not a single number, but a random variable. ... 29 Jan, 02:11PM
Ready Act

I think they're hypothetical because you don't necessarily run them. You just think of theta, but as something that could vary if you ran your experiment several times. Eg. in a linear regression, you get only one estimate of beta, your effect, but you also use statistical theory to calculate its standard error, which measures the hypothetical variation it would show if you ran the regression on several samples from the same population. ... 29 Jan, 06:50PM
Ready Act

A screen shot of NB software, with a solid line added to highlight the connection between the assigned text (a) and discussion forum (b). (Color online)

time, teaching in this way teaches the students who serve as teachers and learning is greatly enhanced.

Finally, collaborative video annotation can improve the classroom experience. First, it enables the instructor to take and encourage questions even at the cost of not getting through the planned material; collaborative video annotation makes it easy to assign the portions of the lecture for which there was no time during class. Second, students come to class much more prepared. Of course, we also assign written material for students to consume, but social science evidence indicates that seeing the same material via different modalities enhances learning (Mayer 2003). (Practicing what we teach, we have posted a video explaining many of the points in this article at j.mp/HUteach.)

Making Reading Interactive. Suppose you are assigned a chapter to read and you cannot understand one of the key points on the second page. In a traditional class, you would be expected to meet with a teaching assistant at office hours or in the scheduled section meeting. In either case, that could be days from now; if you wait, you will lose sight of what you were reading and probably will not have time to complete the assignment. Alternatively, you could skip this key point, pretend you understand it, make some confused assumption about it, and keep reading. Either option violates all three social science teaching principles.

Instead, our practice is to follow analogous techniques for reading assignments as we do with videos. To do this, we put all class readings in a collaborative text annotation system. We used NB,³ an online annotation system that was created by faculty and graduate students at the Computer Science and Artificial Intelligence Laboratory at MIT. (A modified screen shot of NB is in figure 2.) NB and other such systems enable students to highlight passages in the text they do not understand and ask questions in a separate text field. Other students, or the professor or teaching assistants, can then see the questions posted and instantly respond—again 24/7. On any given night, even if class is not scheduled to meet for another week, students are able to receive fast feedback on their questions (in a Facebook-style discussion forum). Because teaching teaches the teacher, considerable pedagogical benefits go not only to those who get their questions answered but also to those answering questions. And because social connections motivate, students give more time and attention to the readings and the class than they would otherwise.

E-mail Lists to Create Community. Many university courses today have a “class e-mail list” that instructors use to disseminate logistical information to students. (Our e-mail list was created by the Institute for Quantitative Social Science at Harvard for this class, but instructors can quickly create e-mail or

Figure 3

Querying the Class E-mail

The Gov2001 Archives

You can get [more information about this list](#).

likelihood or log likelihood

About 112 results (0.15 seconds)

powered by Google™ Custom Search

[\[gov2001-l\] poisson log-likelihood](#)
<https://lists.gking.harvard.edu/pipermail/gov2001-l/001928.html>
[gov2001-l] poisson log-likelihood. Jens Hainmueller jhainmueller. Sun Apr 13 16:42:59 EDT 2008.
Previous message: [gov2001-l] poisson log-likelihood; Next ...

[\[gov2001-l\] seeing the log-likelihood iterations](#)
<https://lists.gking.harvard.edu/pipermail/gov2001-l/000800.html>
[gov2001-l] seeing the log-likelihood iterations. Miya Woolfolk woolfolk. Wed Mar 25 18:02:37 EDT 2009.
Previous message: [gov2001-l] seeing the ...

[\[gov2001-l\] seeing the log-likelihood iterations](#)
<https://lists.gking.harvard.edu/pipermail/gov2001-l/000799.html>
[gov2001-l] seeing the log-likelihood iterations. Patrick Lam plam. Wed Mar 25 17:49:38 EDT 2009.
Previous message: [gov2001-l] seeing the log-likelihood ...

[\[gov2001-l\] Assignment 4. Binomial model and log-likelihood](#)
<https://lists.gking.harvard.edu/pipermail/gov2001-l/000859.html>
Mar 2, 2009 ... Binomial model and log-likelihood Olona. Remember to apply the log function to the entire Likelihood function. Your terms not only reduce from ...

(a)

[gov2001-l] likelihood vs loglikelihood

Kosuke Imai kimai at fas.harvard.edu

Wed Mar 5 09:26:56 EST 2003

- Previous message: [\[gov2001-l\] likelihood vs loglikelihood](#)
- Next message: [\[gov2001-l\] likelihood vs loglikelihood](#)
- Messages sorted by: [\[date\]](#) [\[thread\]](#) [\[subject\]](#) [\[author\]](#)

That's right. This is because log is monotone transformation:
 $\log(x) > \log(y)$ if $x > y$.

You should *always* take log. This is a rule whenever you are calculating the density function. The logscale is much more stable numerically. Think about the Gamma function, which is equal to $\Gamma(x+1) = x!$ if x is an integer. You definitely want to take a log if x is large.

Another important reason is that the law of large numbers and central limit theorem apply to the loglikelihood function, not to the likelihood function. Do you see how?

Kosuke

On Wed, 5 Mar 2003, Phillip Y. Lipscy wrote:

```
> For 2b, if we use the product term of the negative binomial distribution to  
> estimate the likelihood rather than the loglikelihood, am I right to assume that  
> we should get the same result for the maximum point? i.e. we use loglikelihood  
> instead of likelihood to get rid of the product term, but our results should not  
> change?  
>  
> Thanks,  
> Phillip.
```

(b)

Class e-mail archives (a) and getting an answer (b). (Color online)

discussion lists in many other ways, for example via Google Groups or services such as Piazza.com.) We go further and encourage students to ask questions of the entire e-mail list, instead of just the teaching staff, and ask those students who know the answer to make a contribution by responding. This speeds feedback, helps some students get the benefit of being teachers, and motivates them with social connections. Also, we eliminate any reason to wait until “business hours” to contact, or receive a response from, the teaching staff. To enhance social connections, students can include noncourse related information when it can help build camaraderie in the class; this may even include job opportunities, relevant papers, conferences that might be of interest, and class social events.

In addition, we recently discovered that the class’s e-mail list was available going back for over a decade. We then turned this information into a searchable knowledge base, as well as a community in its own right, by making the e-mail lists searchable. Students now have access to more than 10,000 class e-mails covering many topics and providing instantaneous answers to hundreds of key questions. (Figure 3 displays a screen shot of a sample search (a) and a sample answer (b).) In addition, we obtained permission from Harvard’s General Counsel to make available not only the questions and answers in the archive, but also the author of each e-mail. Thus, not only does the archive provide instant feedback, but it allows students a glimpse into a remarkable network of students who have taken this class in previous years. A tremendously motivating feature of this innovation is finding a question similar to yours by a student who now happens to be a tenured professor at a major university, partner in a law firm, or leader of a major corporation.

We also encourage the time-spanning nature of the class community by building and regularly posting to a Facebook group exclusive to class alumni. We use this Facebook group to communicate job opportunities, data problems, methodological advice, and other information.

2.2. Inside the Classroom

So if all this activity is going on outside the classroom, what is the point of going to class? The innovations in Section 2.1, if used properly and along with other innovations, can improve the classroom experience itself and greatly increase the amount of information learned overall.

Understanding Confusions. Using the innovations we introduce outside of classroom, instructors can learn exactly what students have the most trouble with and use that information to make the classroom experience far more powerful. Thus, for video and text annotation, and for students querying our e-mail database, we collect ongoing data about the topics students discuss, the kinds of questions they ask, and how they answer others’ questions. Before each class, we automatically construct and study heat maps of the readings and assigned lecture timelines, colored by the intensity of annotations. By additionally soliciting students’ feedback, we can piece together before walking into class (1) what students think or say they are confused about and (2) what students are *actually* confused about, judged by direct evaluations.

How exactly do we use this knowledge to increase what students learn? We do this in two ways.

Informed Lecturers. First, the social science learning principles also apply to us as teachers. The instant feedback provided to us on what students are having trouble with, provided by these technological innovations, ought to substantially improve teaching compared to end-of-semester student evaluations or even midterms or final exams. We use this information to focus more time on material we now know students have stumbled over or find confusing. If students have seen a video presentation from a lecture in a previous year, we develop a new way to approach the material for the current year. We also stop and prompt students for questions during parts of the lecture about which we now know they will have difficulty with, or we can ask questions ourselves to generate discussion.

Computer-Assisted Peer Instruction. Second, because students have learned far more outside of class than is typical, and because our lectures are more effectively directed to what they do not understand, we spend less time presenting traditional lectures. This is a substantial benefit because—although lectures may generate a kind of “collective effervescence” that people resonate with, much like they do with concerts, sporting events, or religious rituals (Konvalinka et al. 2011) and that possibly increases cooperation (Wiltermuth and Heath 2009) and further engagement—lectures also include minimal feedback for the instructor, minimal feedback for the students, minimal social connections among the students, and little opportunity for students to learn by teaching.

Thus, we spend a portion of the class via a version of “computer-assisted peer instruction” (CAPI). Peer instruction was introduced by Eric Mazur (1997; see also Crouch and Mazur 2001, and Fagen, Crouch, and Mazur 2002) and has seen widespread use (it is related to another similar protocol called “team-based learning”; see Sweet and Michaelsen 2012).

First, we use an automated system we helped develop called “Learning Catalytics” that implements CAPI and that students sign into when they come to class.⁴ (Instead of prohibiting smartphones in class, we require them or, alternatively, a laptop, tablet, or some other web-enabled device.) We then deliver to their device (and, optionally, also to the screen in the front of the room) a difficult conceptual question. We then give the students a minute or two without discussion to reflect on the question and to indicate their answer on their device. The question can be one of many types—multiple choice, a freehand drawing, a mathematical expression, a directional vector, unstructured text, a map highlight, a drawing, or others. We construct the question out of the most difficult parts of the week’s assignments, so that, ideally, only about 20% of students initially get the answer correct.

Next, Learning Catalytics automatically puts students into groups of two to five in preparation for a discussion about the question (s). We use an automated and analytical approach to select

we automatically assign each student a seat where we want them to sit. We then avoid transaction costs for the remaining CAPI questions and choose groups that do not require students to move.)

Next, we ask the students to try to persuade the other members of their group of the veracity of their answers. Because social connections motivate, we often get highly animated discussions. (Over the course of the semester, we use different groupings so students get to know more students than just the friends they came in with.) We allow the ensuing discussion to continue for approximately two to seven minutes, permitting the time to vary according to the complexity of the question. During this time, the teaching staff move among the groups as participants or just listening in and learning about the students’ misunderstandings and difficulties. Because teaching teaches the teacher, having the students try to persuade their classmates substantially improves their understanding of the subject matter. This is even true for those who got the right answer the first time.

We then deliver the same question to each student’s device again and have them answer it. A minute or two later we project on the screen in front of the classroom (and to the students’ devices) a summary of the answers before and after discussion, which gives the students immediate feedback. For multiple-choice questions, we use overlapping histograms. For freehand drawings, we superimpose all the drawings on top of one another (using alpha-transparency). For equations, we automatically check for algebraically equivalent versions. For free text, we cluster responses. When it works best—which, like in survey research, is primarily a function of us asking sufficiently clear questions—the proportion of correct student answers increases from 20% to more than 80%.

Figure 4 gives an example of a multiple-choice question delivered to a student’s phone (a) and the instructor’s view (c). After the first round, a personal message is delivered to each student’s phone or other device that tells them who to discuss their question with (see note in phone (a)). A seating chart appears at the top right for the instructor, coded with letters for each answer and

When it works best—which, like in survey research, is primarily a function of us asking sufficiently clear questions—the proportion of correct student answers increases from 20% to more than 80%.

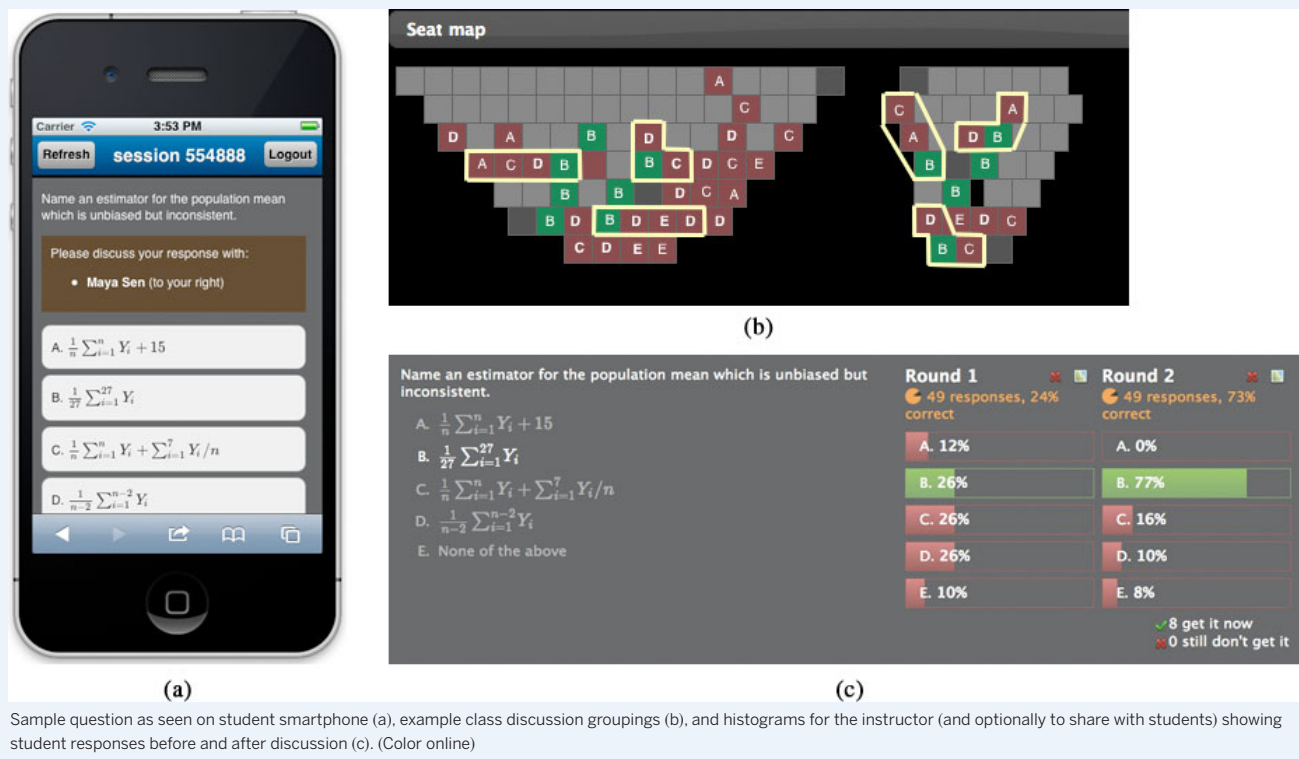
students into groups so that the conversation will be maximally productive. This system is continually updated, but for predictors we begin with data collected to characterize each student at the start of the semester and add each student’s initial answer to the question just asked, their answers to all previous CAPI questions and answers, their experience in the system, and how productive previous CAPI discussions they participated in were. Finally, data from thousands of other similar students in hundreds of other classrooms taking similar courses can be used as well.

Once grouped, the system delivers to each student’s device instructions regarding which other students to talk with and (optionally) where to move their seat to have the discussion. (Most instructors spend time and effort trying to convince students to fill in seats up front; as an alternative, we can let students sit where they like when they walk in, but on the first CAPI question

green for correct (b); the grouping is also shown. The instructor can also see, and optionally can show to the class, histograms of student answers before and after discussion (panel c). Finally, students are given the option of indicating whether they “get it now” or “still don’t get it” (see current tally at bottom right of panel (c)) to provide instant feedback to the instructor.

We intersperse CAPI questions at different points in the lecture. To ensure that students are in a participatory mood during the lecture, we usually begin class with a CAPI question. We also use other questions at the most difficult points in the lecture. Indeed, many who use CAPI do not lecture at all, thus completely “flipping the classroom” as the practice is sometimes called, but with computer assistance. Students are told that answers to these questions do not count to their grade (except for their participation).

Figure 4
Multiple Choice Question Delivered to a Student's Smartphone



The last time we taught with this technology, we asked 14 questions with quantifiably correct answers. Among these, the median increase in the percent correct was a substantial 31 percentage points. We also collected data from a sample of courses using Learning Catalytics across disciplines from other instructors, for classes in which we had no part and at colleges and universities all over the country. These data include 275 distinct questions asked in classes that were part of 19 separate courses, in four different disciplines (statistics, physics, math, and biology). Each of these questions was asked of a classroom of students, followed by an application of our automated grouping technology, peer instruction, and a repeat of the same question. Figure 5 summarizes these results with the median percentage point increase (black dot) and the 10th and 90th percentile values (vertical lines at the ends of each horizontal bar). As is evident from figure 5, the learning from CAPI, without any instructor intervention, is considerable—similar to the results we found in our class. (Furthermore, in our experience, students who learn in this way during CAPI retain the information for far longer than for any other teaching method we have tried.)

CAPI can also be used for subjects with no definitive answer, such as in philosophy, which encourages students to hone their arguments and debate skills; the difference is that we do not necessarily expect a particular directional change in the percent giving each answer, but it seems to do a good job at helping students tune their debating skills and to more deeply understand the issues at hand.

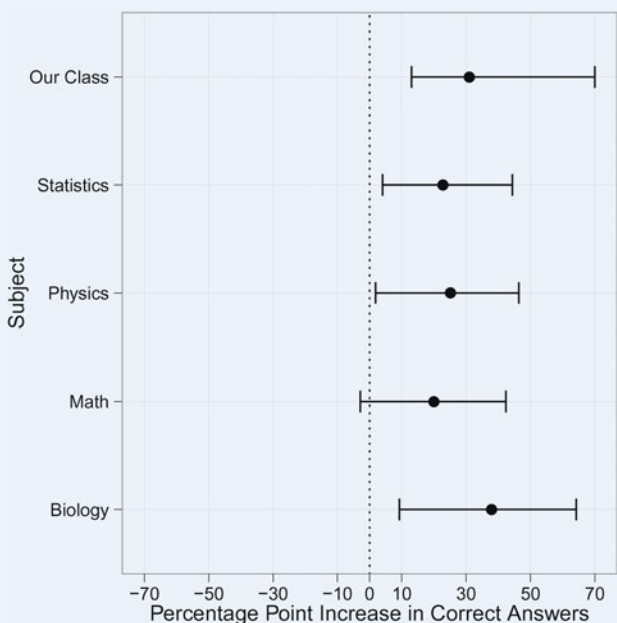
2.3. Across classrooms

During the last few decades, social scientists have been highly successful in convincing the world outside of academia about

the value of large-scale observational and experimental data collection and analysis. After all, this type of quantitative social science has already remade existing companies and established new industries; led to a huge increase in the expressive capacity of human beings; and had a role in reinventing medicine, friendship networks, political campaigns, public health, legal analysis, policing, economics, sports, public policy, and program evaluation, among many other areas. In recent years, what is effectively quantitative social science is much of what is now known to the general public as “big data.” There is no reason why those of us in social science departments who are responsible for creating, applying, and popularizing the innovations that made these changes possible should not also turn this productive machinery to improve our teaching.

Unfortunately, using data collection to improve teaching and learning beyond a single classroom is rare, at least aside from end-of-semester student ratings. Although we have greatly increased the amount of data collected about the classroom we control, many of us need to work together with university officials to implement big data strategies for education. With appropriate protections for individual student privacy and federal regulations, we should do what most businesses do now and instrument as many aspects of the university as possible. The results may be substantial. For example, instead of students receiving ad hoc, idiosyncratic advice from a few other students they happen to know regarding what to major in, what classes to take, and what careers to pursue, good data collection and analytics can give students systematic advice from tens of thousands of previous students they would never have time to know. Students can study many more paths through a college education and see which

Figure 5
Median Percentage Point Increase



Percentage point increase in the percent of students giving the correct answer from before to after computer-assisted peer instruction. Medians and 10th and 90th percentiles averaged by discipline.

suit them, understand what hurdles stand in their way, what roadblocks they should avoid, and which choices will confront them. We can use the instructional staff more efficiently, and help us learn from each other what works, what does not, and what only works well for some instructors. Instead of instructors experimenting by changing what they do in their classroom and never evaluating it because of the absence of a proper control group, they can learn by observational studies—if we make the effort to collect the data and apply the methods we do for research to our teaching as well.

University officials, faculty committees, and staff need to take on board the overwhelming impact social science research has on every area it touches and how it can revolutionize university operations to improve teaching and learning as well.

CONCLUDING REMARKS

In recent years, rigorous education-related research has taken root within physics (Mazur 1997; Deslauriers, Schelew, and Wieman 2011), chemistry (Golde, McCreary, and Koeske 2006), computer science (Porter et al. 2011), medicine and nursing (Ende 1983; Hodder et al. 1989; Rao and DiCarlo 2000), and other areas. In addition, numerous science departments now have dedicated research groups, faculty lines, postdocs, and other staff who specialize in education research adapted to their disciplinary areas—for example, the physics education groups at the University of Arizona, the University of Colorado, Boulder, Harvard University, Kansas State University, the University of Maryland, the Ohio State University, and others; the chemistry education groups at Iowa State University, Purdue University, and Cambridge University; the computer science education groups at Bowdoin College, Duke University, and Villanova University, and the medical education research and evaluation group at Stanford University, among others.

These groups are studying an aspect of human behavior—that is, social science research. It is gratifying to see another area where we have had an influence, but social scientists are, of course, especially well situated to make major contributions to these emerging literatures. We should accept the challenge and encourage our colleagues to join in, systematize social science knowledge, harvest useful social science generalizations for teaching, develop new technologies and innovations that improve our teaching and our students’ learning, and contribute our valuable faculty lines. The result likely will be that classrooms will be filled with better educated and knowledgeable students, albeit trying very hard to look attentive.

ACKNOWLEDGMENTS AND DISCLAIMER

For advice, suggestions, comments, and help with various technologies, we thank Ferdi Alimadhi, Tim Brenner, Matthew Chingos, Merce Crosas, Phil Desenne, Leslie Finger, Dave Kane, David Karger, Konstantin Kashin, Brian Lukoff, EJ Marmonti, Eric Mazur, Jen Pan, Molly Roberts, Julie Schell, Brandon Stewart, Dustin Tingley, and Sacha Zyto. Gary King is cofounder of Learning Catalytics (recently acquired by Pearson). ■

NOTES

1. About 80% of children with some types of cancer take part in randomized experiments that will help only the next child diagnosed. As hard as it would be, we ought to be able to find a way to convince faculty to participate in teaching experiments too.
2. See <http://harvard-atg.github.com/Catool>. A modified screen shot of this video annotation tool is in figure 1.
3. See <http://nb.mit.edu>.
4. See <http://LearningCatalytics.com>.

REFERENCES

Barab, Sasha A., and Thomas M. Duffy. 2000. “From Practice Fields to Communities of Practice.” *Theoretical Foundations of Learning Environments* 1: 25–55.

Burn, Shawn M. 1991. “Social Psychology and the Stimulation of Recycling Behaviors: The Block Leader Approach.” *Journal of Applied Social Psychology* 21 (8): 611–29.

Cacioppo, John T., James H. Fowler, and Nicholas A. Christakis. 2009. “Alone in the Crowd: The Structure and Spread of Loneliness in a Large Social Network.” *Journal of Personality and Social Psychology* 97 (6): 977.

Centola, Damon. 2010. “The Spread of Behavior in an Online Social Network Experiment.” *Science* 329 (5996): 1194–197.

Chi, Michelene T., Nicolas DeLeeuw, Mei-Hung Chiu, and Christian Lavancher. 1994. “Eliciting Self-Explanations Improves Understanding.” *Cognitive Science* 18 (3): 439–77.

Chingos, Matthew M. 2013. “Class Size and Student Outcomes: Research and Policy Implications.” *Journal of Policy Analysis and Management* 32 (2): 911–38.

Christakis, Nicholas A., and James H. Fowler. 2007. “The Spread of Obesity in a Large Social Network over 32 years.” *New England Journal of Medicine* 357 (4): 370–79.

Christakis, Nicholas A., and James H. Fowler. 2008. “The Collective Dynamics of Smoking in a Large Social Network.” *New England Journal of Medicine* 358 (21): 2249–258.

Crouch, Catherine H., and Eric Mazur. 2001. “Peer Instruction: Ten Years of Experience and Results.” *American Journal of Physics* 69: 970.

Dawson, Shane. 2006. “A Study of the Relationship between Student Communication Interaction and Sense of Community.” *The Internet and Higher Education* 9 (3): 153–62.

DeSchryver, M., P. Mishra, M. Koehler, and A. P. Francis. 2009. Moodle vs. Facebook: Does Using Facebook for Discussions in an Online Course Enhance Perceived Social Presence and Student Interaction? In *Proceedings of Society for Information Technology and Teacher Education International Conference*, Chesapeake, VA, pp. 329–336.

Deslauriers, Louis, Ellen Schelew, and Carl Wieman. 2011. “Improved Learning in a Large-Enrollment Physics Class.” *Science* 332 (6031): 862.

- Dihoff, Roberta E., Gary M. Brosvic, and Michael L. Epstein. 2003. "The Role of Feedback during Academic Testing: The Delay Retention Effect Revisited." *Psychological Record* 53 (4): 533–48.
- Dihoff, Roberta E., Gary M. Brosvic, and Michael L. Epstein. 2004. "Provision of Feedback during Preparation for Academic Testing: Learning Is Enhanced by Immediate but not Delayed Feedback." *Psychological Record* 54 (2): 207–32.
- DiPrete, Thomas A., Andrew Gelman, Tyler McCormick, Julien Teitler, and Tian Zheng. 2011. "Segregation in Social Networks Based on Acquaintanceship and Trust." *American Journal of Sociology* 116 (4): 1234–83.
- Dubner, Stephen J., and Steven D. Levitt. 2006. "Freakonomics: A Star Is Made." *New York Times Magazine*, May 7.
- Ende, Jack. 1983. "Feedback in Clinical Medical Education." *The Journal of the American Medical Association* 250 (6): 777.
- Fagen, Adam P., Catherine H. Crouch, and Eric Mazur. 2002. "Peer Instruction: Results from a Range of Classrooms." *Physics Teacher* 40 (4): 206–09.
- Fowler, James H., and Nicholas A. Christakis. 2008. "The Dynamic Spread of Happiness in a Large Social Network." *BMJ: British Medical Journal* 337: a2338.
- Garrison, D. Randy, Terry Anderson, and Walter Archer. 1999. "Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education." *The Internet and Higher Education* 2 (2–3): 87–105.
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (01): 33–48.
- Golde, Michael F., Christine L. McCreary, and Randi Koeske. 2006. "Peer Instruction in the General Chemistry Laboratory: Assessment of Student Learning." *Journal of Chemical Education* 83 (5): 804.
- Graff, Martin. 2003. "Individual Differences in Sense of Classroom Community in a Blended Learning Environment." *Journal of Educational Media* 28 (2–3): 203–10.
- Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81.
- Hodder, R. V., R. N. Rivington, L. E. Calcutt, and I. R. Hart. 1989. "The Effectiveness of Immediate Feedback during the Objective Structured Clinical Examination." *Medical Education* 23 (2): 184–88.
- Killingsworth, Matthew A., and Daniel T. Gilbert. 2010. "A Wandering Mind Is an Unhappy Mind." *Science* 330 (6006): 932–32.
- King, Gary, and Maya Sen. 2013. "The Troubled Future of Colleges and Universities." *PS: Political Science and Politics* 46 (1): 81–89.
- Konvalinka, I., D. Xygalatas, J. Bulbulia, U. Schjødt, E. M. Jegindø, S. Wallot, G. Van Orden, and A. Roepstorff. 2011. "Synchronized Arousal between Performers and Related Spectators in a Fire-Walking Ritual." *Proceedings of the National Academy of Sciences* 108 (20): 8514.
- Lim, Chaeyoon, and Robert D. Putnam. 2010. "Religion, Social Networks, and Life Satisfaction." *American Sociological Review* 75 (6): 914–33.
- Mayer, Richard E. 2003. "The Promise of Multimedia Learning: Using the Same Instructional Design Methods across Different Media." *Learning and Instruction* 13 (2): 125–39.
- Mazur, Eric. 1997. *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall.
- McDermott, Rose, James H. Fowler, and Nicholas A. Christakis. 2009. "Breaking Up Is Hard to Do, Unless Everyone Else Is Doing It Too: Social Network Effects on Divorce in a Longitudinal Sample Followed for 32 Years." Available at SSRN: <http://ssrn.com/abstract=1490708> or <http://dx.doi.org/10.2139/ssrn.1490708>.
- Morse, Gardiner. 2012. "The Science behind the Smile: An Interview with Daniel Gilbert by Gardiner Morse." *Harvard Business Review*, January-February: 84–90.
- Pachucki, Mark A., Paul F. Jacques, and Nicholas A. Christakis. 2011. "Social Network Concordance in Food Choice among Spouses, Friends, and Siblings." *American Journal of Public Health* 101 (11): 2170.
- Porter, Leo, Cynthia Bailey Lee, Beth Simon, and Daniel Zingaro. 2011. Peer Instruction: Do Students Really Learn from Peer Discussion in Computing? In *Proceedings of the Seventh International Workshop on Computing Education Research*, 45–52. ACM.
- Rao, Sumangala P., and Stephen E. DiCarlo. 2000. "Peer Instruction Improves Performance on Quizzes." *Advances in Physiology Education* 24 (1): 51–55.
- Rosenquist, J. Niels, Joanne Murabito, James H. Fowler, and Nicholas A. Christakis. 2010. "The Spread of Alcohol Consumption Behavior in a Large Social Network." *Annals of Internal Medicine* 152 (7): 426.
- Rovai, Alfred P. 2003. "The Relationships of Communicator Style, Personality-Based Learning Style, and Classroom Community among Online Graduate Students." *The Internet and Higher Education* 6 (4): 347–63.
- Shea, Peter. 2006. "A Study of Students' Sense of Learning Community in Online Environments." *Journal of Asynchronous Learning Networks* 10 (1): 35–44.
- Summers, Jessica J., and Marilla D. Svinicki. 2007. "Investigating Classroom Community in Higher Education." *Learning and Individual Differences* 17 (1): 55–67.
- Sweet, M., and L. K. Michaelsen. 2012. *Team-Based Learning in the Social Sciences and Humanities: Group Work That Works to Generate Critical Thinking and Engagement*. Sterling, VA: Stylus Publishing.
- VanderWeele, Tyler J. 2011. "Sensitivity Analysis for Contagion Effects in Social Networks." *Sociological Methods & Research* 40 (2): 240–55.
- VanLehn, Kurt, Arthur C. Graesser, G. Tanner Jackson, Pamela Jordan, Andrew Olney, and Carolyn P. Rosé. 2007. "When Are Tutorial Dialogues More Effective than Reading?" *Cognitive Science* 31 (1): 3–62.
- Whitehurst, Grover (Russ). 2010. "Education Research: Past, Present, and Future." *Policy Perspectives*, 1–12. <https://www.wested.org/cs/we/views/rs/1006>.
- Wiltermuth, Scott S., and Chip Heath. 2009. "Synchrony and Cooperation." *Psychological Science* 20 (1): 1.