Convocation 2014 – Jeff MacKie-Mason remarks

September 3, 2014

I. Attention:

In our mission, we are explicit about our commitment to impact: we want to build a better world. We say we will do this by “creating and sharing knowledge.” My central focus in today’s convocation is impact through creating knowledge.

If we’re in the business of knowledge creation, we need to think for a moment about what we mean by knowing something. One of the most common, and important things we want to know is whether A causes B. For example, if a health librarian provides credibility scores for health information on the Internet, will users experience better health? May sound like a good idea, but before we spend a lot of time and money implementing it, is it correct?

Here’s an example:

As you can clearly see, Tyler Vigen, using reliable data from the US Federal Aviation Administration and the US National Science Foundation, has shown that the number of space launches in a year determines the number of sociology doctorates awarded.

Now, has Tyler created new knowledge? Has he answered a previously unanswered question? More precisely, has he answered this question correctly? Do we now know that A causes B, that launching satellites causes people to become sociologists?

I’m think you all are comfortable seeing through this trick. Tyler has found some a strange correlation, but correlations do not prove causation. For example, it takes at least 4 years to get a doctorate in sociology. So let’s re-plot the data, showing doctorates earned 4 years after launches:

They really don’t seem to have much to do with each other, do they?
II. Need:

You might think this is silly, that people wouldn’t be fooled by bad science like this. But they are, all the time! For example, a remarkably high fraction of the health news reported on the Internet, from the Huffington Post to the New York Times, is reporting correlation, not causation, but then the writers suggest that maybe we should rely on these correlations to make important individual or social policy decisions.

For example, I went to the Huffington Post the other day and read one of the first stories that came up on its health page. The title was “How Gratitude and Kindness Go Together for Brain-Changing Happiness”. The author, Kellie Edwards, is a registered psychologist — sounds credible. She tells us that expressing gratitude is “something that will ... transform your life.” She tells us “I know that my brain is changing for the better.”

Wow. How does she know this? First she refers to peer-reviewed research by Barbara Frederickson, who “discovered that people who flourish have a 3:1 ratio of positive to negative emotions.” Note that this is a correlation: among people who flourish, we see a higher positivity ratio. From Frederickson’s correlation Edwards asserts causality: “increasing the frequency of these [positive] moments matters.” But she offers zero evidence that positivity causes people to flourish.

One of the great reasons to spend time at a university, even if you have no intention of becoming a professional scholar, is to improve your ability to answer previously unanswered questions. One part of this is becoming aggressive about spotting attempts to substitute correlation for causality. This is an example of what we call “critical thinking”. If you want to have positive impact on the world, your actions need to cause the right effects!

________________________

Today’s take-away — the tl;dr version — is that anytime someone presents you with a conclusion, your immediate response should be to ask at least two questions. Don’t just accept what you are told. Challenge it: is this really true?

III. Satisfaction of the need:

You students are lucky: we have the world’s most talented information school faculty. Let’s learn something from them. I’m going to illustrate how our faculty go about discovering causality — and how much work it takes — by discussing three examples of their recent research.

Yang, Newman, Forlizzi

The first study is by Rayoung Yang, Mark Newman and Jodi Forlizzi, “Making Sustainability Sustainable”. Yang is a UMSI PhD student; Newman is a professor here. Forlizzi is at Carnegie-Mellon.

Yang et al. ask the following previously unanswered question: does a smart home device — the Nest Learning Thermostat — cause people to change their heating and cooling behavior over a period of at least a year? This is an important question, helping us in the quest to design Internet of Things devices to reduce our harm to the environment, and to save money.

Their first finding: people with Nest did engage in more active heating and cooling control after they bought it. Yang et al. learned this by having the subjects keep activity diaries, and by interviewing them.

Have they answered their research question? Can they publish an article now, get famous, and stimulate a lot of profitable investment in smart home devices? Are we ready to advise the government to raise our taxes to give rebates to people who buy a Nest?

NO. Let’s try my advice: when faced with an assertion, ask at least two questions.

1) The first challenge question might be: User behavior was different after they bought the Nest...but how do we know the Nest caused the change in behavior? Maybe there was a lot of public concern about global warming that month, and people generally got more
careful about heating and cooling, whether they had a Nest or a conventional thermostat. Then if we start spending tax money on subsidies for Nest buyers based on the mere correlation we might have no effect, except to make taxpayers poorer, and to divert resources from something that might actually work, like a vaccine for HIV.

2) A second challenge question: User behavior was different after they bought the Nest, but does the behavior change last for more than a few weeks? Maybe we’re seeing a new toy effect, but no lasting impact on sustainability.

So, this time consuming, expensive study, which required finding willing subjects, paying them to participate, getting them to spend hours writing diary records and being interviewed...this effort was not enough to answer the question.

But Yang et al. knew they hadn’t answered their question yet, so they performed two additional studies.

First, to determine whether it was getting a Nest, or something else that caused the behavior change (like lots of global warming stories in the media), they found another set of people willing to be studied who had conventional thermostats. What Yang et al. found is that there was a difference: people without a Nest controlled their heating and cooling less actively.

So that helps answer the question: there is now at least some evidence that the smart home device — the Nest — made a difference.

What about the next challenge: does the Nest cause a change in sustainability behavior that lasts at least a year? It costs about $250 — if it’s just a toy that we play with for a few weeks before we go back to our old habits, it’s probably not worth it to buy one!

To deal with this part of the question, Yang et al. did yet another study: they recruited most of the original Nest users to participate in another in-depth interview about their behavior, this time a year later. What did they find? No, behavior changes did not persist. So there we are: despite some supportive evidence, the final answer is NO: the smart home device may have caused a short-term behavior change, but it did not last.
Do you see how hard it is to create new knowledge — true facts and true theories? Anyone can spin a story to explain a correlation, but it takes patience, creativity, and a drive to keep asking questions if you want to demonstrate that A causes B — or doesn’t.

**Reinecke - Gajos**

Next example: UMSl professor Katharina Reinecke has been working on answering previously unanswered questions about cultural differences in visual preferences. People increasingly acquire and interact with information visually, on screens. Suppose you are trying to help people across multiple cultures through providing health information about, say, diabetes. We know you’ll have more impact if you offer versions in the local languages. Should you also offer multiple versions with different visual designs based on local preferences?

Reinecke, in a CHI 2014 article with Krzysztof Gajos, asked whether preferences for specific visual features are different around the world. For example, should the amount of color used vary by country?

You can’t just ask people their preferences. For one thing, if Indians say they like more color, you’re left wondering, more color than what? More color than Swedes?

So Reinecke and Gajos constructed a controlled, randomized experiment. They collected web pages from around the world. They then presented a set of 30 pages to nearly 40,000 subjects from 179 countries and asked them to rate them. Thus, subjects in different countries were comparing a given set of pages against each other: we know that “more” means “more like site B”.

40,000 subjects: that’s a lot of work. From it, Reinecke and Gajos discovered many differences in visual preferences around the world. But let’s look at another example: older folks prefer more complex web sites than do younger folks. Here is some of their evidence:

When comparing pairs of sites, under-20s prefer the simpler sites on the left, while the over-51s prefer the more complex web sites on the right.
Do we conclude that as people age, they prefer information to be presented in denser, more complex formats? Remember today’s lesson: whenever anyone presents a conclusion, ask at least two questions.

So let’s challenge this conclusion with a question: how do we know that it isn’t just that people are more comfortable with styles with which they are more familiar? Older folks are used to Web 1.0 web sites from the 1990s and early 2000s: text-heavy, graphics-light. Maybe they are just more used to getting information from complex web sites, so it is experience, not age that is causing their stated preference. By the way, this is not a criticism of Reinecke and Gajos: they make this point about interpreting the age result themselves.

The difference between two different stories — which is the true cause — matters. Suppose you lead a billion dollar web design agency and think age is the cause, and so create two presentations for every client web page: if the real cause is familiarity, not age, then within a few years this all this extra effort would be wasted, because as people of all ages get more used to simpler Web 2.0 designs the age correlation will disappear. Your business will fail because you incorrectly assumed a correlation was the true cause.

Chen / Mei

The third and final example for today: UMSI professors Yan Chen and Qiaozhu Mei have been studying how to increase contributions to social causes. In fact, Yan and I ran an experiment on how to increase donations to support the online Internet Public Library about a decade ago, but I’m going to talk about the more recent research Yan and Qiaozhu are doing.

They want to answer this previously unanswered question: does the opportunity to form teams in a crowd-funding setting increase charitable contributions? To answer this, they study Kiva.org, which is “a non-profit organization with a mission to connect people through lending to alleviate poverty.” On Kiva people put up loans to help the poor start or improve businesses, and lenders can form and join teams. Kiva is charity: there is no interest paid on the loan, and on some loans the principal is never repaid.
In a 2012 study of over 5000 lenders, Chen, Chen, Liu and Mei compare lending by those who join teams and those who do not. After controlling for different charitable motivations, they measured the difference in lending behavior between team and solo lenders. They found that team members made nearly one more loan per month, on average.\(^2\)

Great! They answered the question: does the opportunity to form teams increase lending?  
Yes! Right?

Wrong. That is not the question they answered. They found that belonging to a team was *correlated* with more loans, but they don't know whether team formation *causes* more loans. For example, those who joined teams were, of course, different people than those who were not. Maybe something unrelated — call it Factor X — causes them to join teams and to make more loans. If so, take away teams and these lenders will still have Factor X and will still make more loans — the ability to join teams may have been irrelevant. Answering the original unanswered question requires more work.

Kiva was so excited by the initial results that they allowed Chen, Chen, Liu and Mei to run a randomized field experiment on 550 lending teams. To determine whether being on a team *causes* more lending, they tested two features of team behavior that are only possible if you belong to a team: encouraging your team members through goal-setting, and helping your team members through coordination.

Their experiment was clever: they created artificial lenders with made-up names, had them join teams, and make real loans (research can be expensive as well as time-consuming). The artificial lenders then engaged in one of four behaviors: a control case or one of three treatments. Either they made the loan and did nothing else (the control), or they made the loan and announced to other team members a goal, or they provided coordinating information, or both. They sent these messages through the private team message forum (which Kiva

creates); because the message forums exist only for teams, non-team lenders can’t be influenced by this kind of team activity.

They found that when members of previously inactive teams were encouraged by both goal setting and coordination, they increased their loan activity by about one-third loan per month.³

Now we have a convincing answer to this previously unanswered question: being on a team is not only correlated with higher contributions, but it causes higher contributions (on average) when team members engage in goal setting and coordination.

But, remember I said every time someone asserts a conclusion, you should ask at least two questions? One question we might ask here is whether team formation helps only in social lending settings, like Kiva, or would it also help pure charitable donation programs, like we run every winter on Crowdrise to support our Alternative Spring Break program? We could guess whether these results generalize...but we don’t know. For now this is an unanswered question.

Another question might be: choosing to be on a team increases lending, but if we nudge reluctant members to join a team, will that increase their lending? Perhaps this effect only works for those predisposed to join a team.

This new knowledge is having an impact in the world: this summer Kiva invited two UMSI PhD students, Wei Ai and Yang Liu, out to San Francisco, where they worked with Kiva engineers to develop a system to recommend an appropriate team to new members. Kiva is currently trying out this new feature, to increase lending to fight poverty around the world.

And so the circle of research goes on.

V. Action:

What have we learned today? A fundamental reason to study or work here is to improve your ability to answer previously unanswered questions — to create knowledge. And when you

aren’t the one creating the knowledge, you should be questioning those who claim to provide it.

Every time someone asserts a conclusion, come back with at least two questions.

One critical question might be: “does the evidence truly answer the question?” For example, every day there is a new study reported about surprising behavior on Twitter, Facebook, Wikipedia. Ask the question: Is this surprising behavior caused by something about Facebook, or is something else the cause? This matters because if we don’t know the cause for behavior, then we can’t use system design or social policy to change behavior: to create impact.

A second critical question might be: “Does this evidence, collected in one setting, tell us something useful about other settings?” For example, do older people prefer more complex information displays across the generations, or is this a current phenomenon that won’t be true when millennials become the middle-aged? People are quick to apply the results of studies to designing in a different setting ... which often doesn’t work.

Answering previously unanswered questions, solving previously unsolved problems, is hard. We have the world’s best information school faculty here, and they have each devoted their adult life to learning how to create new knowledge. Their research projects take months or years before they succeed. But because they are careful, because they aren’t satisfied with casual guesses and informal evidence, we are making genuine new discoveries here, and helping move the world forward.

You students need to work hard if you want to have successful careers. You don’t need to be here to do that hard work. But you chose to be here, and if you work hard here you can benefit from one of the things that top universities do best: learning how to ask questions, and judging carefully when those questions are actually answered, correctly and usefully.

Today’s lesson: Every time someone asserts a conclusion, formulate at least two questions.

Including here, now, today: don’t just accept my assertion. Ask questions. Challenge what you have heard. And don’t leap to casual answers.