Mistaking Statistics for Rigor in American Business School Research: Should China Follow, or Lead in a New Direction?

Forum on Case-based and Qualitative Research in Business Administration in China

9 November 2013

Allen S. Lee
Professor of Information Systems
Virginia Commonwealth University
I would like to thank Dr. Ji-Ye Mao 毛基业 for giving me the opportunity to address the business administration scholars attending this forum. For over 25 years, I have been a proponent of qualitative research, and I have been a proponent of doing qualitative research rigorously so that it qualifies as science.

Most or all of us attending this forum already believe in the merits of case-based and qualitative research. We don’t need to be convinced. However, most of our colleagues in schools of business administration are different from us. They do research that is quantitative -- specifically, statistical. Some of my statistical colleagues genuinely want to learn how qualitative research can, in its own way, qualify as objective, valid, and scientific. I refer to this group as the “friendly statistical researchers.” But unfortunately, there are still some statistical colleagues who are not sympathetic to us and continue to automatically judge statistical research as scientific, and qualitative research as unscientific. I refer to this group as the “unfriendly statistical researchers.” In this keynote address, I will share with you the line of reasoning that I have developed to communicate with both the friendly and unfriendly statistical researchers.

The line of reasoning that I will share also leads to a bigger lesson. The lesson is: (Lee, Briggs, and Dennis, 2014):

- The statistical analysis of a theory can be rigorous, but this means that [the] rigor is in the statistical analysis, not necessarily the theory or an explanation which uses the theory.

But this lesson is hardly well known. Researchers in American schools of business administration have long made the mistake of conflating statistics with rigor. This is a mistake that the next generation of business administration researchers should not, and need not repeat, especially here in China, where business school researchers should consider declaring their independence from any American-imposed or other western-imposed research norms. Business administration researchers in China may, can, and should choose to lead business administration research in new directions.

In the United States, we already know that our business administration research is so statistically rigorous that it often has no practical relevance. Executives in companies and our own MBA students cannot understand the research that we do because it is so statistically sophisticated. In the United States, our business administration research is so rigorous that the high level of rigor has had the effect of severely limiting, or even eliminating, the research’s applicability or relevance. This is a major flaw of American business administration research. Why should any business administration researchers continue to follow the
As a doctoral student at M.I.T. in the late 1970’s, I initiated my research program of establishing the scientific basis of qualitative research. I searched the literatures of the philosophy of science and history of science in order to learn the criteria for identifying, and doing, research that deserves to be called scientific. I distilled these criteria and then, in my research publications, I demonstrated how qualitative research can be conducted so as to satisfy the criteria of science.

But then, about five years ago, I began to wonder: Does statistical research live up to the criteria of science? I had spent many years establishing how qualitative research can be done scientifically, where I assumed that statistical research was already being done scientifically. But that was only an assumption. So, I started doing research on this. I began to take long, hard looks at actual instances of statistical research published in our top business administration journals. I searched for lessons and insights from the literature of the philosophy of statistics. I eventually concluded, on the one hand, that business administration research has performed its statistical analyses correctly, at least from the perspective of rote mathematics. But I also concluded, on the other hand, that business administration research has often performed its statistical analyses without satisfying the criteria of science. Sure, more often than not, business administration research is quantitative and uses numbers, but just because research is quantitative and uses numbers does not make it scientific. What, then, makes research truly scientific? I will share my lessons with you in Part I of my keynote, “What statistical analysis is good for in science,” Part II, “What statistical analysis is not good for in science,” and Part III, “What statistical analysis is, and is not, good for in business administration research.” And these lessons have major ramifications for qualitative research.

Please know that, unlike many qualitative researchers, I am not opposed to statistics. In fact, in my recent research, I have dedicated my efforts to making contributions to statistical methods, where I have been strengthening statistical research by making explicit what its scientific foundation is.
Then, in Parts IV and V, I will discuss a criticism that has often been leveled against case-based and qualitative research. It is the criticism that the study of a single case is not generalizable. This criticism, which is incorrect, is based on the assumption that a large sample size is always needed. However, I will explain exactly what a large sample size is, and is not, needed for – and this, in turn, will allow me to explain the way in which a single case is generalizable.

Finally, in Part VI, I will turn your attention to action research and design research. I will present only highlights, not details, of these two approaches because I will be using them just to transition to the Nobel laureate Herbert Simon’s concept of “the sciences of the artificial,” which he contrasts to the “the sciences of the natural.” The final point of my keynote address will be that we need to pursue business administration research as a “science of the artificial” (instead of as a “science of the natural,” which is what we in the United States have been pursuing for decades). And qualitative research has a special role to play in establishing business administration research as a “science of the artificial.”
To explain what statistics is good for, I am choosing an example from natural science – physics. My reason for this is that the social sciences have modeled many of their own social-science research methods on the research methods used in the natural sciences. Therefore it behooves us, and even requires us, to take a look at how statistical methods are used in natural science. Statistical researchers in business administration have been copying the statistical methods used in the natural sciences. But have they been copying the statistical methods correctly?

I know that most or all of you are expecting me to talk about qualitative methods, certainly not statistical methods in physics. However, as qualitative researchers often say, especially those of us who do interpretive research, context is everything. When we do an interpretive case study, for example, it is not enough to interview research subjects. We also have to collect data on their context – their social context and their historical context. Well, the same applies to us. Much of the research that we qualitative researchers do depends on our own social and historical context in the community of scholars. And for us qualitative researchers in business administration, our social and historical context has been dominated by quantitative, statistical researchers, who have adopted the natural sciences as the norm for what science is.

For that reason, let’s now step into the framework of our statistical colleagues – which is a framework they have adopted from the natural sciences – where I will show you, using the framework of statistics itself, what statistics is, and is not, good for.

Let’s take the example of an object in vertical free fall in a liquid. We would expect that a denser liquid, like oil instead of water, would cause the object to slow down more in its velocity and that the object would exert a smaller and smaller force. Now, how would we explain that?

To keep this example simple and uncontroversial, I am using an exercise I found for undergraduate students on a web site for the University of Wyoming’s College of Engineering and Applied Science.

Please take a look at the equation in red. Let’s say that the dependent variable is the object’s force, F, and that the independent variable is the object’s velocity, V. And $\beta_0$, $\beta_1$, and $\beta_2$ are constants. The equation in red shows how the dependent variable F is related to the independent variable V.
The table on the left and the graph on the right are from the University of Wyoming web site. In an experiment, 18 observations (or data points) are taken, where each data point consists of a measurement of the object’s velocity and force at the same moment in the object’s free fall in the liquid. The curved line is a graph of the equation in red font, \( F = \beta_0 + \beta_1 V + \beta_2 V^2 \). The graph also allows us to show the dispersion of the 18 data points around it.
For the equation you see in red, how would it be fitted to the object in this experiment? The equation in green shows how – the beta’s take specific numerical values as constants. And statistically speaking, the fit is great. It is over 99%. We never see an r-square like that in the social sciences.

In a different experiment using a different object that is falling in a different liquid, these constants would be different. But still, we would have the same theory, represented by the equation in red, that is simply fitted to whatever the instance is, as would be represented by another equation in green.
Now, I offer this as the “big picture” of what it is that statistical analysis is supposed to do in natural science and social science. It is for fitting a theory to the real world. This is what statistical analysis is good for telling us.

Statistics is good for simply calibrating a general theory to fit a specific instance.

But note that the statistical analysis does not create, develop, or build the theory. The statistical analysis requires that a theory already exist before the statistical analysis can even begin.

And also note that the statistical analysis does not, in itself, care about whether the theory is true. The statistical analysis cares about whether the numbers in the green equation, such as -0.1278 and -0.0271, are statistically significant, but the statistical analysis already presumes that the theory, in the red equation, is correct, or has already been tested or will be tested. But this is not a job for the statistical analysis itself.
The next question, therefore, is this: Is the theory correct? I am referring to the theory both in the form of the red equation, which represents it in general, and in the form of the green equation, which represents it in the instance of this one experiment. We saw that the graph of the equation did not fit the 18 observations perfectly. There is error. Because of the presence of error, may we still consider our theory to be correct?

I have created a new table on the right, using an Excel spreadsheet. The new table shows what the theory predicts for dependent value – the force $F$ of the object – for every observed value of the independent variable, the velocity $V$. The new table allows us to compare the predicted values and the actual, observed values.
Sure, the fit of this equation to the data set is excellent. The $R^2$ value is an astounding 99.86%. But…
...what about all of these prediction errors? There are 18 data points. For each data point, the predicted value of the dependent variable F is different from the observed and measured value of the dependent variable F. All of these errors are telling us that our original theory could be wrong. A correct theory predicts correctly. Our theory, as instantiated in this experiment, is not predicting correctly.
Therefore, what statistical analysis is not good for is telling us is whether the theory is correct, or true, or valid. A statistical analysis, even when conducted properly and completely, does not tell us if the theory is correct and does not test the theory for us.

Those of us who have done statistical analyses know that they involve much hypothesis testing. The hypothesis testing is about details such as whether the numerical coefficients, such as -0.1278 and -0.0271 from our equation in green, are significantly different from 0 despite sampling error. This is important to know, but this does not involving testing the theory. In fact, statistical hypothesis testing even makes the assumption that the theory is true. But what we want to do is to test our theories to see if they are true – this is the point of science. We may not simply assume that they are true – this would be unscientific.
I would like now to shift gears in the discussion. I will apply the foregoing lessons about statistics to business administration research. The article I am selecting, "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," was published by Fred Davis, Richard Bagozzi, and Paul Warshaw in the top business administration research journal, *Management Science*, in 1989. They derived their theory, the technology acceptance model or “TAM,” from an earlier well known theory, Fishbein and Ajzen’s theory of reasoned action.

This instance of research is illustrative of much, though certainly not all, business administration research in the United States and Canada. First, much behavioral research in business administration originates in organizational behavior and organizational studies, which includes Fishbein and Ajzen’s theory of reasoned action. Behavioral research in organizational behavior and organizational studies has set the pace for much other behavioral research in business administration, such as behavioral accounting research, behavioral marketing research, and behavioral information systems research. And methodological innovations in research in organizational behavior and organizational studies, such as structural equation modeling, eventually diffuse to other disciplines in business schools, such as accounting, marketing, and information systems.

Second, another reason I am selecting the article by Davis is its high citation count – over 9,000 since the year 1989, according to Google Scholar. It has influenced, and it has been emulated by, many other articles published in business administration.

Davis’ theory, the “technology acceptance model” or TAM, offers an explanation of how an individual person does, or does not, accept a computer technology. Now, let me go over the arrows in the diagram.

TAM theorizes that an individual’s actual use of a computer technology follows from his or her behavioral intention (BI) to use the technology. Then it theorizes that the individual’s behavioral intention (BI) to use the computer technology follows from three things: her attitude (A) towards the computer technology, her perceived usefulness (U) of the computer technology, and her perceived ease of use (E) of the computer technology. It also theorizes that the individual’s perceived usefulness (U) of the computer technology follows from the individual’s her perceived ease of use (E) of the computer technology.
Davis’ behavioral theory consists of three propositions. I will read just the first one. “TAM [the technology acceptance model] postulates that computer usage is determined by BI [a person’s behavioral intention to use the technology], but differs [from the theory of reasoned action] in that BI is viewed as being jointly determined by the person’s attitude toward using the system (A) and perceived usefulness (U) [p. 985].” Davis then translates or operationalizes these words into a mathematical equation, which I show in blue font. The equation is, “BI = β_{1,0} + β_{1,1}A + β_{1,2}U.”

Davis does the same for two more theoretical propositions.

So far, what Davis is doing is standard for much business administration research.
The “boxes and arrows” diagram, which we see in many research articles in business administration, is translated into the three equations.

\[
\begin{align*}
BI &= b_{12} + b_{13}A + b_{14}U \\
A &= b_{23} + b_{24}U + b_{26}EOU \\
U &= b_{33} + b_{34}EOU
\end{align*}
\]
Here are the statistical results that Davis presents. Notice, in the bottom half of the diagram, how all the results indicate a successful statistical analysis. The numerical values that have been measured for the constants in the three equations are statistically significant, and even one of the the r-square values is impressively high for behavioral research, 51%.

So far, so good. In the United States and Canada, this is what a typical behavioral researcher in business administration hopes to achieve in a study.
So, are there any problems?
Yes, there is a major problem. It is that Davis and his co-authors, in his article, never tested their theory, TAM. And we can be sure there is prediction error. Their theory does not predict correctly. This is exactly analogous to the situation in physics.

Statistical researchers in business administration pride themselves on calling their research "science," but statistical research itself does not include performing the scientific task of actually testing the theory. The natural sciences, like physics, test their theories by using them to make predictions and then seeing if the predictions come true. Business administration research that claims to be scientific must do this too, but it has not been doing this.
Let me emphasize some points I've already made as well as make some new ones.

First, a statistical analysis is invaluable for calibrating a theory so that the theory can be tested subsequently, but the statistical analysis itself does not test the theory.

Second, a statistical analysis does not create, build, or develop a theory. The theory must already be created, built, or developed before the statistical analysis can take place.

Third, still, a statistical analysis (even a perfect one) is just good for telling us how to calibrate a theory so that it best fits the specific instance where it is applied.
Fourth, “The statistical analysis of a theory can be rigorous, but this means that [the] rigor is in the statistical analysis, not necessarily the theory or an explanation which uses the theory” (Lee, Briggs, and Dennis, 2014).

Fifth, to follow the example of business administration research in America would mean submitting to (a) perpetuating the predominance of statistical research and (b) foregoing the strategy of taking newer and more innovative paths.

Sixth, given these problems, should China follow American research, or lead in a new direction?
Now, let’s take a step back. How did business administration research in America become so statistical in the first place? The reason is historical.

Way back in the 1950’s, there was concern that American business schools were too vocational. There was concern over whether they deserved to be in universities because they were not scholarly enough.

In 1959, two studies were published by influential bodies: the Ford Foundation and the Carnegie Corporation.

The Ford Foundation report said:

“There has been too little pure research…”

“…business research needs to become more analytical, to develop a more solid theoretic underpinning, and to utilize a more sophisticated methodology”

“This in turn requires that the business schools turn to the underlying disciplines such as the behavioral sciences and mathematics and statistics…”

So clearly, this pointed business schools in the direction of pure research and statistical research.
But that's not all. In the same year, the Carnegie Corporation also published a report. It said:

"...business schools need to concentrate on developing a body of widely applicable generalizations which have been scientifically tested…"

"Both hypothesis forming and hypothesis testing are essential."

"...very rarely [in 1959] is emphasis placed on developing analytical findings which can be fitted into a general system of principles and tested in a scientific manner."

The result of the Ford Foundation report and the Carnegie Corporation report was the major effort to create business administration research that was scientific and statistical.

However, today, we may legitimately question whether the Ford Foundation’s and Carnegie Corporation’s conception of science and statistical research was correct. We know much more about different kinds of science today, thanks to studies in the philosophy of science and history of science. And we also know that statistical research and statistical hypothesis testing are just one part of science, where it is even just an optional part of it.

Today, business administration research in America is still laboring under the legacy of the Ford Foundation and Carnegie Corporation reports, including their questionable ideas about what scientific research should be. Should the historical accident of these reports in America, in 1959, establish the agenda for business administration research in China in 2013?

Surely, there must be a better way of doing rigorous, academic research in business administration. I will return to this point later, in Part VI.
We just took a step back to look at the historical context that led to the predominance of statistical analyses in business administration research.

Now, let’s return to the present: the predominance of statistical research supports a criticism that we qualitative researchers have all heard: “You can’t generalize from a sample size of one.” And as I will explain, it is a criticism which is incorrect.

The truth is that case-based and qualitative research is generalizable, and not generalizable, in the same way that statistical research is.

The first logical truth we need to deal with in considering sample size and generalizing is this: \textit{No observations consistent with a theory, no matter how numerous, ever offer any proof that the theory is true}. I’ll give three reasons for this.
First, let me introduce a concept from symbolic logic. It is “the fallacy of affirming the consequent.”
Consider the major premise, "if p is true, then q is true."

For example, if p, "it is raining outside" is true, then q, "the street is wet" is true.
Next, for the minor premise, we observe that q, "the street is wet" is true.

Now, when we see that "the street is wet," does this necessarily mean that "it is raining outside"? Absolutely not! The street could be wet for many other reasons. In general, just because q is true does not mean that p true.

That's why this form of reasoning is incorrect. It is the fallacy of affirming the consequent.
Now, how does this apply in scientific reasoning?

Let p be the theory T.

Let q be a prediction made by the theory T in an experiment.

We see that it would be a fallacy to reason that just because a prediction p is true, the theory would be true.
Now let's say that we use the theory to make a prediction $q_1$, and this prediction turns out to be true.

Then we use the theory to make a new prediction $q_2$, and this prediction turns out to be true, too.

Then we use the theory to make a new prediction $q_3$, and this prediction turns out also to be true.

And we do this for 100 or even a 1,000,000 different predictions, and all of them turn out to be true.

Would 100 or even a 1,000,000 true predictions prove the theory to be true? If you were to answer "yes," then you would be committing the fallacy of affirming the consequent 100 or a 1,000,000 times.
Therefore, even with a large "sample size" size of confirming observations, a theory is not generalizable.
The fallacy of affirming the consequent is only 1 of 3 reasons that a large sample size is not
generalizable. The second reason is Goodman's new riddle of induction.

Now, suppose you observe many, many emeralds before time $t$. You make the following
correct statements:

- Emerald #1 is green.
- Emerald #2 is green.
- And so on.

So you draw the conclusion, “all emeralds are green.” We will call this conclusion #1. For
now, we won’t say if this conclusion is right or wrong. We will save that question for later.
Next, instead of describing an emerald as green, let me describe as emerald as “grue.”

“Grue” is defined as describing an emerald that you encounter up through time $t$ and is green or that you encounter after time $t$ and is blue.

Now, all the very same emeralds that you saw before fit these statements:

- Emerald #1 is grue.
- Emerald #2 is grue.
- And so on.

In fact, all the emeralds that you saw before that were green, you can now correctly say that they are grue.

So you draw the conclusion, “all emeralds are grue.” We will call this conclusion #2. Could conclusion #2 be correct?
From CONCLUSION #1, it necessarily follows that “the next emerald to be observed will be green.”

From CONCLUSION #2, it necessarily follows that “the next emerald to be observed NEED NOT be green.”

So the result is that even if we are generalizing from the exact same things, we are ending up with two radically different conclusions. The two conclusions cannot both be correct at the same time.
So what this means is that, therefore even a large sample size (going up to 1,000,000 emeralds) is not generalizable.
Next, we have Hume’s problem of induction, which is about the “uniformity of nature assumption.”

According to the Oxford Dictionary of Philosophy, the “uniformity of nature” is “the principle that the future will resemble the past, in that when sufficiently similar situations recur, similar effects follow.”

We can also state it as the assumption that the more two things, F and G, are alike (“uniform”), the more F and G will behave alike in the future.

But is this assumption true?

Suppose that the two things, F and G, have behaved alike 100 times in the past, or 1000 times in the past, or 1,000,000 times in the past. Will F and G behave alike in the 1,000,001st time, in the future?

In other words, may we generalize past performance to the future?
So, let’s put “the uniformity of nature assumption” to the test.

Argument 1.1 has two statements. The first statement, with the straight underlining, is: In past experience, a large number of Fs have behaved like Gs. The second statement, with the curly underlining, is: Therefore, in the future, the next F will behave like a G.

Of course the difficulty is that the second statement does not necessarily follow from the first.

How may we solve this problem?
In Argument 1.2, we do this by introducing Theory 1: If the first statement, with the straight underlining, is true, then the second statement, with the curly underlining, is true.

Next, we observe that the first statement, with the straight underlining, is true.

Conclusion: the second statement, with the curly underlining, is true. In other words, based on the large number of Fs that have behaved like Gs in the past, we generalize into the future that the next F will behave like a G.

But now, we have a new problem. How do we know that Theory 1 is correct?

How may we solve this problem?
In Argument 2.2, we do this by introducing Theory 2: If the first statement, with the straight underlining, is true, then the second statement, with the curly underlining, is true.

Next, we observe that the first statement, with the straight underlining, is true.

Conclusion: the second statement, with the curly underlining, is true. In other words, based on the large number of tests that have confirmed Theory 1 in the past, we generalize into the future that the next test will confirm Theory 1.

But now, we have a new problem. How do we know that Theory 2 is correct?

How may we solve this problem?
In Argument 3.2, we do this by introducing Theory 3.

You see, this goes on and on. So ultimately, in trying to justify the uniformity of nature assumption, in Argument 1.1, we need to create Theory 1. And then, to justify Theory 1, we need to create Theory 2. And then, to justify Theory 2, we need to create Theory 3.

So David Hume, the philosopher who identified this problem, did not conclude that generalizing is necessarily invalid – only that it lacks a logical foundation. To justify it leads to an infinite regress in reasoning, not a logical foundation.
The previous slides on Hume’s problem of induction can all be condensed into the equivalent of just one paragraph of text. I found this in a chapter written by Roseberg in a book, *The Cambridge Companion to Hume*. I will not read it.
So the result is this: Therefore even a large sample size ("In past experience, a large number of Fs have behaved like Gs") is not generalizable.
David Hume's recognition of the problem of generalization is famous. He is even cited by Don Campbell, who is known among positivist statistical researchers for his classic book, *Experimental and Quasi-Experimental Designs for Research*. 

```
Even Don Campbell, the "father" of Experimental and Quasi-Experimental Designs for Research (1963, p. 17), has stated explicitly:

"Induction or generalization is never fully justified logically.

```
So, let’s review. The philosophy of science and symbolic logic are exceedingly clear on this point: induction or generalization has no objective, scientific foundation. This follows from any one of three independent arguments: the fallacy of affirming the consequent, Goodman’s new riddle of induction, and Hume’s problem of induction.

But suppose that you, as a business administration researcher, would like to advise a manager of a corporation that you’ve performed a study with lessons that can be applied to the manager’s corporation. On the one hand, there is no scientific or logical basis for generalizing your study of other settings to the new setting of this corporation. On the other hand, you may still proceed ethically, by making a “full disclosure” of the lack of scientific basis for generalizing, and letting the practitioner audience know exactly all of the judgment calls you are making.

Richard Baskerville and I identified that a full disclosure can usefully include four judgment calls.

The first judgment call that a researcher needs to share with a practitioner is the “uniformity of nature” judgment call. The first judgment call is simply to assume that the uniformity of natural assumption is true. To review, the uniformity of nature assumption allows us to generalize from a past situation to a future, similar situation, where sufficient similarity or uniformity between them means that similar effects or outcomes will occur in the future situation. The problem is, as Hume’s problem of induction told us, there is no scientific or logical basis for generalizing.

The first judgment call leads to the second judgment call, which is that the second setting we want to generalize to, from the first setting, is sufficiently similar to the first setting. This requires that we know what “sufficient similarity” is.

The second judgment call leads to the third judgment call, which is that we know what the relevant conditions and variables are in the first place.

And the fourth judgment call is that the theory we want to generalize is true. Remember that, in the logic of science, we may never show a theory to be true. We may only show a theory to be false. So the fourth judgment call is another leap in judgment, just like the first three judgment calls.

If the necessity of making these four judgment calls makes generalizing from a single case sound very tenuous – remember that these four judgment calls are also required for statistical studies too.
Now, here is a concrete example of how to generalize from a case. This is an example that Richard Baskerville and I developed. We published this in *MIS Quarterly*. The case is a previously published one – M. Lynne Markus’ “Power, Politics, and MIS Implementation.”

Now, consider this theoretical proposition from Markus’ case study. Markus instantiated this proposition in an organization that she calls “GTC”:

An increase in “resistance [to the new information system], as instantiated in the situation where GTC’s divisional accountants resist FIS” leads to a decrease in “shift in power from GTC’s divisional accounting to GTC’s corporate accounting.”
Here, we begin to generalize.

As the yellow arrow on the left indicates, we assume that the Financial Information System FIS in the original organizational setting, GTC, is sufficiently similar to the new information system AIS in the new organizational setting QED.

And as the yellow arrow on the right indicates, we assume that the original organizational setting, GTC, is sufficiently similar to the new organizational setting, QED.

In these two simple generalizations, we are making all four judgment calls.

This includes the first judgment call, which is that the uniformity of nature assumption is correct. We are assuming that two information systems FIS and AIS are sufficiently alike, and that the two organizational settings GTC and QED are sufficiently alike, so that the theory that explains resistance to MIS implementation at GTC can also explain resistance to MIS implementation at QED.

This includes the second judgment call, which is that the similarity between the two settings is sufficient for us to generalize from the first setting to the second.

This also includes the third judgment call, which is that we know what the relevant variables are — such as R or “Resistance” (by the intended users of a system to the system) and PSR or “Power Shift Realized in the Organization.”

And the fourth judgment call is that Markus’ theory, which survived empirical testing in the first organizational setting, GTC, is true, and that it will continue to be true in the second organizational setting, QED.
Next, suppose we want to generalize even more.

In the green arrow on the left, we are generalizing from “QED’s divisional accountants” to “the managers in the QED division initially with power.” Of course, some people might feel that it’s a “stretch,” to generalize from accountants to managers in general. But this is what generalization involves.

In the green arrow on the right, we are generalizing from “QED’s corporate accounting” to “the QED division initially without power.” Again, this generalization could be a “stretch.”

Again, we are making the four judgment calls.

I wish to emphasize that by our making each step in generalizing so explicit, the practitioner audience can see all of the judgment calls we are making, and any practitioners can decide for themselves whether the benefits of accepting the judgment calls outweigh the risks and costs of the judgment calls.
So far in this keynote, I have addressed how we case-based and qualitative researchers can, and must, respond to the majority of business administration researchers who are statistical – some of whom are genuinely sympathetic to qualitative research, and some of whom are not.

Now, I will end this keynote on a more constructive note: how can we case-based and qualitative researchers help steer the course of future business administration research?

Almost all of the current American research in business administration takes what the Nobel laureate Herbert Simon calls a “sciences of the natural” approach. It is the approach of the natural sciences, which basically seek to describe and explain what already exists in the world. This is the approach taken in physics, as I earlier covered in Part II. And this is also the approach copied in behavioral business administration research, as I earlier covered in Part III.

Where a “sciences of the natural” approach seeks to describe and explain what exists, a “sciencies of the artificial” approach seeks to change what exists. It creates new things – artifacts, artificial things – instead of just looking at already existing things.

Herbert Simon says that professional schools – such as schools of engineering, medicine, architecture, and business – should not take a “sciences of the natural” approach, but instead, they should take a “sciences of the artificial” approach.

But the sad story is that, in business administration, we have been taking a “sciences of the natural” approach almost exclusively.

Let me show you, using action research and design research as examples, what a “sciences of the artificial” approach can accomplish.
There are many forms of action research, but all of them involve:

- using a theory, if only a tentative one, to diagnose a problem in an organization or other specific setting in the real world,
- planning or designing an action, based on the theory,
- applying or taking the action in the real-world setting, for the purpose of solving the problem, achieving a goal, or otherwise improving a situation,
- evaluating both the action for both its efficacy as a solution to the real-world problem and its efficacy as a prediction made by the theory, and
- specifying learning, which means improving the theory based on lessons learned from the action that was taken.

And then, the new, improved theory allows a new, improved diagnosis. So the cycle begins anew.

Many of you are already familiar with action research. However, my new lesson is that...
...here, our research actually creates something new, instead of just describing and explaining something that already exists. Traditional business administration research in the United States has almost never done this. The very fact that an intervention, creating something new, is an oddity does not speak well about American business administration research at all.
What information systems researchers call “design research” also demonstrates Simon’s “sciences of the artificial.”

There are many forms of design research, but this diagram from Vaishnavi and Kuechler captures the essence of design research in its many forms. Yes, there is a strong similarity to action research.

The steps are:

- defining the problem,
- suggesting a solution based on a theory or other knowledge,
- developing or building the solution in the form of an artifact,
- evaluating how well the artifact worked,
- and then concluding with lessons that lead to better defining the problem, better suggesting a solution based on a better theory or other knowledge, better develop or building a new solution, and so forth.

So the cycle begins anew.

Many of you are already familiar with action research. However, my new lesson is that...
...here, our research actually creates something new, instead of just describing and explaining something that already exists. Traditional business administration research in the United States has almost never done this. The very fact that an intervention, creating something new, is an oddity does not speak well about American business administration research at all.
Why is it important for us to consider the sciences of the artificial, such as action research and design research?

Business administration is a profession. And we can see that the professions of medicine, engineering, architecture, and law have all taken a “science of the artificial” approach not only as professions in the real world, but also as academic disciplines in universities. But strangely, business administration disciplines in the university – such as accounting, finance, information systems, marketing, and personal/human resource management – have taken a “science of the natural approach,” which, furthermore, they have interpreted as heavily statistical. And yet, action research and design research are in themselves proof that business administration disciplines can also take a “science of the artificial” approach.
I started out this keynote address with the question, “What strategies should business administration researchers in China take in the global community of research?”

I have argued that the approach of American business schools is overly statistical, and they have been pursuing this approach not necessarily in incorrect ways, but surely in incomplete ways. So, for Chinese business administration researchers to follow the American lead would not make sense. Furthermore, to follow the American lead would be to enter an overcrowded “market.” There are already so many scholars who are doing statistical, “science of the natural”-style business administration research. And they are truly excellent scholars. Why enter a field overcrowded with people (and competition) like them, where they have already been enjoying a “head start” compared to you, where any additional research contribution would only be incremental at best?

Case-based and qualitative research is hardly an overcrowded “market.” The competition here is less. And not only that, case-based and qualitative research has the potential to produce much more relevant research – research that managers, executives, organizations, and governments would find useful. In America, there are political constraints inside the community of business administration researchers that block many of them from pursuing case-based and qualitative research. Why replicate American problems and constraints here in China?

For many years in the United States, I have lamented that the vast majority of our young scholars – PhD students and assistant professors – have been institutionally forced into pursuing statistical, “science of the natural” research. I am told that, in China, business administration research is much newer, and it is at a crossroads. I encourage business administration researchers to choose qualitative, case-based, “science of the artificial” research, not only because it is a more promising intellectual thing to do, but also because it can better benefit business administration in China, and the world.

Business administration researchers in China can lead the world in business administration research. I look forward to seeing this happen.