Mitchell Musings 4-11-11: Rationalizations, Explanations, Incentives

Last week, I happened to hear a radio newscaster attribute a downturn in the American stock market to an aftershock earthquake in Japan. Now I suppose there must be a model that indicates that bad news for Japan is automatically bad news for the U.S. economy or the U.S. stock market. But I can make a plausible model that would say that the bad news in Japan would be good for the U.S. economy or stock market (less competition globally; more demand for U.S. exports). Most plausible of all is that there was absolutely no connection between the two events and that whoever wrote the script for the newscast just needed an explanation for a financial development.

The stock market seems particularly prone to attract such non-explanation explanations since it fluctuates from day to day. When it goes up, you often hear that the explanation is that everyone was buying. When it goes down, the explanation given is that everyone was selling. That is an odd idea when you take a moment to reflect. Every stock that is bought is one that is sold. Repeat: every stock that is bought is thereby sold. Buyers need sellers. (Sometimes, however, when I point this out to folks, I get surprised looks.) The everyone-was-buying or everyone-was-selling story is really saying nothing more than that the market consensus was that prices should be higher or lower than they were before. But the story sounds plausible and seems to provide an explanation. Human beings want explanations; they want stories and rationalizations. We really don't like the idea of randomness and luck or just not knowing.

There is a tendency, because we want stories of causation, to over-attribute good or bad outcomes to leaders. In politics when the economy is performing well, the president, governor, or mayor is apparently given credit. Political leaders understand this tendency and often take credit for the jobs "they" created. Of course, in downturns, political challengers to incumbents blame the incumbents for the jobs that "they" lost. Various studies do suggest that election outcomes are influenced by the direction of the economy. I am not saying political leadership has nothing to do with economic trends. But I am saying there is a tendency to give more credit or blame for what happens to leaders than they deserve.

Similarly, successes of enterprises - or organizations within enterprises – tend to be over-attributed to top managers. As any reader of Dilbert will undoubtedly attest – and anyone who has ever taught or studied in a management school should know – the field is prone to fads and gurus who sell fads. And there is a tendency – because we want to attribute causation to something – for fads and gurus to arise. Let's consider a possible mechanism.

Suppose that there are two outcomes for a firm or division within a firm: success or failure. Suppose a guru/fad is in vogue and top management has a choice: follow the guru/fad or don't follow. Suppose following the guru/fad does not in fact contribute to success. The chart below illustrates the four possible combinations, designated as A, B, C, and D.

	Don’t Follow Guru/Fad	**Follow Guru/Fad**
Success	**A**	**B**
Failure	**C**	**D**

If you did nothing new as a top manager, i.e., you didn’t follow the guru/fad, and your organization experienced success (A), you have offered evaluators no explanation in your own behavior for the success. Evaluators are likely to attribute the success to something else, whatever handy explanation they can find. On the other hand, if your organization experienced success and you followed the latest guru/fad (B), you can be seen an up-to-date manager who adopted the wise course of action and therefore succeeded. You should certainly be rewarded with pay, promotion, or even a better position in some other organization. (B is better than A in the face of success.)

Failure (C and D) can never be good. But if you followed the latest guru/fad and yet the organization experienced failure (D), it is possible you might escape blame. Perhaps something else can be found to be the needed explanation for the failure since you were doing the right thing. If the organization experienced failure and you are perceived as having done nothing (not followed the guru/fad; position C), surely you will be blamed and penalties will likely ensue. (D is better than C in the face of failure just as B is better than A in the face of success.) In short, everything points to following the guru/fad as the low-risk strategy since you will end up in the preferred B or D.

There is much more to be said about this conclusion – which, after all, is based on a “story.” It does not explain why some guru/fads catch on and others don’t. And you might be a brave leader who did something unique that was plausibly related to success other than following the guru/fad approach. In that case, you might well be credited with being exceptionally perceptive and innovative as a manager if the organization experiences success. But the risk is that you might end up in position C, in which case the risk increases that you will be blamed for the failure. (You took an oddball approach rather than follow the guru/fad and look what happened.)

The key point here is that we tend to commit to stories and explanations and then interpret events as consistent to that commitment. Our institutions – as in the management case – can reinforce such beliefs and commitments. Challenges to prior commitment are likely to provoke annoyance at best.

The commitment problem comes up in other contexts. An example appeared in my hometown newspaper, the *Los Angeles Times*, last summer. There is general concern about the quality of education in the Los Angeles Unified School District. The *Times* discovered that the District had been compiling "value-added" data on teachers but was not using these data in personnel evaluations. There is a whole literature, I gather, in the field of education on the use of value-added scores and how valid they are. The idea seems to be to look at student <u>improvement</u> via standardized tests for students under a particular teacher. Since what is being measured is the <u>change</u> in the test score rather than the absolute value of the score, the value-added methodology is seen as adjusting for pre-existing student advantages or deficiencies.

The *Times* got access to the value-added scores and published them on its website <u>by teacher name</u>. As might be expected, teachers who had low scores were unhappy, although the *Times* gave teachers the option of commenting on their scores on its website.

I don't want to get into the validity of such scores, since I don't know the education literature. I can, however, see at least some questions about a) interpreting the scores (since, for example, students are not randomly assigned to teachers) and b) about the wisdom or ethics – given that uncertainty – about publishing the scores by teacher name. It would seem (to me) that if there were doubts about validity, publishing the scores by name was not appropriate. The *Times* could have run averages or distributions for particular schools, for example, without names.

The more the publication of the scores by name was challenged, however, the more the *Times* insisted that it had done a service to the community. Parents could see if their children had good teachers, for example. Yet if the *Times* had looked at some of the comments that teachers put on its website, those comments might have at least raised some doubts.

In some cases, it appeared there may have been errors in the data set, e.g., the teacher was not teaching the subject in which he/she appeared to be rated. One comment stood out to me as suggesting a problem with the methodology. The particular teacher who commented was rated as "average" in effectiveness overall, including specifically average in effectiveness in English instruction, but he/she still had complaints. I reproduce his/her comments below in italics. I have inserted *, **, etc., at points in the comment to serve as footnotes indicating problems in grammar.

“My value-data score change* after other teacher complaints that the data was** incorrect based upon pool for teacher* with 60 or less*** students**** it just goes to show you how other teacher's***** data can effect****** mine. Not only are scores based upon what my students are doing, but how other teacher's***** students are doing. Just goes to show you how there are lies, damn lies, and then statistics.”

- - - - - - - - -

*Possible typos; should be “changes” in the first case and “teachers” in the second. **Purists would say “were” but “was” is now acceptable. ***Should be “fewer.” ****Sentence should end with a period at this point and a new sentence should start with “It.” *****Should be “teachers’” in both instances. ******Should be “affect.”

- - - - - - - - -

There are various interpretations that might be drawn from this teacher’s comment. It could be that all teachers who are rated “average” in effectiveness in English are actually sadly ineffective and surprisingly inarticulate. But it could also be that the value-added approach is subject to wide error in its ratings. This teacher, perhaps, should have received a poor rating – particularly in English - but was incorrectly rated as average. Since the possibility of error cannot be ruled out, the *Times’* continued assertion that all was well – its commitment to that view - seems inappropriate (to me, but not to it).

What is the lesson from this musing? At some level, we know that we tend to seek and see explanations for random events. We know that political leaders cannot be responsible for everything that happens when they are in office. We know that management is particularly prone to faddism and gurus. And we know that once we commit to a particular interpretation or course of action, it is very difficult to change course and admit error, or even possible error. It’s fine to argue forcefully for a position we believe in the social sciences. I often do. But there should always be room for doubt.