Harvard Business Review

November 2009

Two leading researchers discuss the value of oddball data

A Conversation with Stephen Scherer by Roger Martin

Businesses often face big, messy problems—challenges that defy precise definition, change constantly, and have numerous causes. To find solutions, says Roger Martin, dean of the University of Toronto's Rotman School of Management and author of The Design of Business (Harvard Business Press), managers must move through a "knowledge funnel" comprising three stages: staring into the mystery; developing a heuristic, or rough rule of thumb; and creating an algorithm, or step-by-step formula for addressing the problem. To explore this process, Martin turned to someone who has wrestled with a confounding problem for two decades: world-renowned autism spectrum disorder and genomics researcher Stephen Scherer of Toronto's Hospital for Sick Children. This is an edited version of their conversation.

Martin: I'm very interested in how managers, scientists, and designers tackle so-called wicked problems—big, unwieldy puzzles that look utterly unmanageable. Take the origins of autism: You could spend years just figuring out what angle of attack to take—is it genetic, environmental, viral? Most problem solvers would start with a hypothesis, test it, and then look for data that confirm or disprove it. But you took a different approach. Can you explain your process?

Scherer: Autism is a vast problem; no single researcher or lab can take on its full breadth. I focused on just one piece of it: the data that everybody else was throwing away. I call it the garbage-can approach. My belief is that answers to really difficult problems can often be found in the data points that don't seem to fit existing frameworks. To me, those little variations are like signposts saying "Don't ignore me!" Evolution doesn't tolerate useless junk for long, so all data points, even the oddball ones, need to be considered seriously.

Martin: I think most people would sweep the anomalous stuff under the carpet—they'd say it's just noise, and we should focus on the results we see again and again, the confirming data. What did you see in the oddball data?

Scherer: I noticed a pattern in genetic variations: Children with autism had more deletions and duplications of genes than usual, and the variations tended to occur on specific genes. Scientific literature contended that 99.9% of the human genome—the genetic instruction book that tells our bodies how to develop—was exactly the same for all of us and that the variations amounted to just an altered letter here and there, so to speak. But our group found that the deletions and duplications were important—that they were more like textual differences than mere typos. We set out to explain them in autism.

Martin: So you attacked the mystery by asking what could be true, rather than what was true. And you narrowed the field of inquiry and came up with an evidence-based heuristic that whittled the mystery down to a manageable size.

Scherer: I focused closely on what children with autism do—repetitively overanalyzing numerical patterns, for example—and why they do it. And I inferred that the genetic deletions and duplications

I saw, known as copy number variations, predisposed some children to have developmental imbalances, which led to behaviors typical of autism. Since then we've been able to identify a few autism-susceptibility genes.

Martin: Most research is focused on creating reliable outcomes—outcomes that can be consistently recreated. That's a big part of the scientific method. But I think the major leaps forward in knowledge come from focusing, as you did, on achieving a valid outcome—one that actually answers the question to give us the information we really need. It seems to me that reliability and validity are inherently incompatible—that to achieve a valid outcome, one must incorporate some aspects of the subjectivity and judgment that are typically eradicated in the quest for a reliable outcome. I feel that the distinction between reliability and validity is at the heart of the innovation dilemma.

Scherer: I agree. Quantitative measures that produce reliability often strip away nuance and context and thus sometimes even prevent the discovery of a valid outcome. By emphasizing validity, even at the expense of reliability, a researcher can get to a heuristic that moves the work forward. Focusing on the anomalous data, I was able to see things that others couldn't. When the scientific establishment didn't believe it, we knew we were onto something big. In retrospect, it's so simple to see that these copy number variations were not at all biological outliers, just outliers of the scientific dogma of the time.

Martin: The orthodoxy is to use what we already know, to hone and refine it. I used to be a consultant, and whenever we started a case, the first question of maybe 85% of my colleagues would be: "Where can I find a template? Is there a PowerPoint summary of another case in the database that will tell me to do this, this, and this?" We want to have an algorithm to fall back on. Formulas are important—we need them to develop efficiencies. When you have a methodology for solving the problems—a script to follow—there are enormous time savings, but the danger is that you stop looking at mysteries entirely.

Scherer: We haven't quite progressed to the algorithm stage in autism research; we're still exploring aspects of the mystery and refining our heuristics in areas we understand better. But the time will come. Look at the Human Genome Project. Once we had figured out the mystery of the genome, we could create a map of it—an algorithm. And labs around the world could coordinate using that algorithm. So there are huge benefits. But it takes time to get there.