Corporate Event December 2005

INSIGHT

Event Measurement: Reliability vs. Validity

Companies are spending more and more money to reliably measure the impact of corporate events on their customers. But how valid are the results?

ROGER MARTIN

Roger Martin is dean of the Rotman School of Management at the University of Toronto and a professor of strategic management. Before joining Rotman, he was a director of Monitor Co., a global strategy-consulting firm in Cambridge, MA, with clients such as The Proctor & Gamble Co., Honeywell, and AT&T. In addition to several articles published in Harvard Business Review, he wrote the book "The Responsibility Virus: How Control Freaks, Shrinking Violets — and the Rest of Us — Can Harness the Power of True Partnership" (Basic Books, 2002).

You've tabulated post-event survey results, tracked event attendance, and counted media impressions after your event — all reliable data that can help you prove the effectiveness of your event. But that's not enough according to Roger Martin, dean of the Rotman School of Management at the University of Toronto. In fact, Martin argues that when companies focus exclusively on collecting reliable, consistent data, they often do so at the expense of validity. Martin's innovative concept was recognized by Harvard Business Review as one of 20 "Breakthrough Ideas of 2005" (Harvard Business Review, February 2005). We talked to Martin to find out how companies can use both reliability and validity to measure the success of their strategic corporate events.

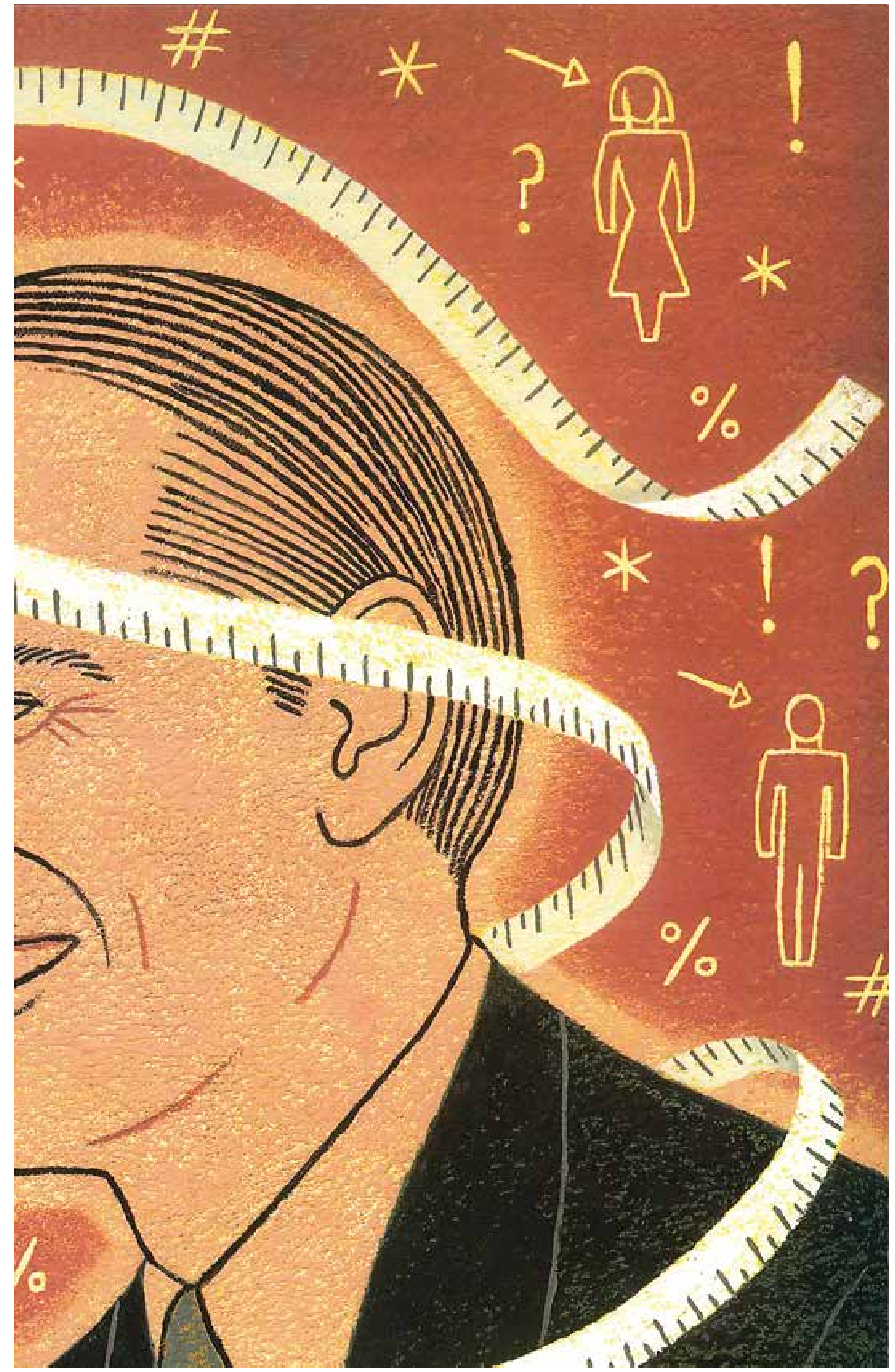
CE: WHAT IS RELIABILITY AND WHAT IS VALIDITY?

RM: A process is *reliable* if it produces the same outcome every time you run it. For example, if you take a blood sample and divide it into 100 subsamples and test each sample for hepatitis B, a perfectly reliable test would give you the same answer 100 times.

A *valid* process is one that produces an accurate outcome consistent with your desires. So if you took one vial of blood and tested it for hepatitis B, and the test was negative and you did not go on to develop the symptoms of hepatitis B, it would be a valid test. If the test were positive and you developed symptoms, it would be valid. But the test would be invalid if it were negative and you went on to exhibit symptoms.

A reliable test could report 100 times out of 100 that you don't have hepatitis B and be completely invalid.

ILLUSTRATION: JAMES STEINBERG



INSIGHT

CE: HOW DOES THIS APPLY TO MARKETING MEASUREMENT?

RM: Over the past 20 years, businesses have used management science and information technology to generate reliability — from customer-relationship management (CRM) systems to Six Sigma. These reliable measurements concentrate on a limited number of quantitative, objective, bias-free data, such as customer demographics, and data from past purchases. As a result, they are consistent and repeatable.

On the other hand, valid measurements also seek to identify whether a marketing program met the desired objectives, even if the system used to collect the data does not produce a consistent, predictable outcome. To measure validity, you measure qualitative data, such as the mood of the customer or his attitude toward a new product, which require subjective judgment to assess.

Reliable measurement tends to focus on inputs instead of outputs. For example, one thing you can measure reliably in marketing is dollars spent. I've known marketers who say, "We have a good marketing program because we're spending as many dollars as our competitors." We can measure dollars spent and be exact, but it's a long way from the desired output: selling products.

A reliable metric that's closer to the desired output is the number of impressions a marketing program generates. But again, what ultimate outcomes are you trying to achieve with marketing initiatives? Do you really want to generate impressions? No. You ultimately want to sell products.

If companies depended solely on reliable data vs. valid data, they would never spend any money on marketing and they wouldn't hold events. They would spend money where they could more easily make the connection to an actual purchase. These reliability-oriented companies would say, "Spending money on marketing is like piling up dollar bills on of the floor and pouring gasoline on them and lighting a match. You can't prove to me beyond a shadow of a doubt that marketing spending is generating sales."

CE: SO SHOULD A MARKETER ABANDON RELIABLE METRICS?

RM: No. The optimal situation is to have both validity and reliability in every system. It is not healthy for any organization to tip the balance too much to either reliability or validity.

Unfortunately, there is a natural tension between the two measurements. And in the corporate world today, reliability usually wins for two reasons: the comfort in the idea of "proof" and an aversion to bias.

Measure the information that's reliability-oriented, because it's all helpful. But then view the information as flawed, because it doesn't necessarily contribute to an in-depth understanding of the customer. A company should spend an equal amount of time and effort gaining a valid, deep understanding of its customers, finding out how they actually make their purchase decisions, and identifying the role of event marketing in their purchase decisions.

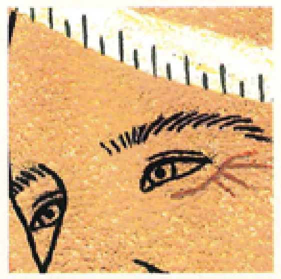
Get as much reliability as you can get, without going overboard and ignoring the downsides and the validity consequences. We're better off erring more on the side of validity.

CE: WHAT ARE SOME EXAMPLES OF RELIABLE MEASURES FOR CORPORATE EVENTS?

RM: One reliable measure is the cost of the event. You can also track things like the number of customers who exchanged business cards and held conversations with salespeople.

Post-event surveys also generate reliable data, even if you're asking about attendees' opinions. This is because traditional, multiple-choice surveys are designed to interpret these opinions as quantitative data. All the nuances are taken out of them.

Another reliable metric is attendance. You can accurately measure the number of attendees you influenced at an event. But that's just one part of the equation. The number doesn't tell you if you influenced them in a positive way. You could have great attendance at a dreadful event. That's why the validity-based measurements are so important.



If companies depended solely on reliable data vs. valid data, they would never spend money on marketing and they wouldn't hold events. They would spend money where they could more easily make the connection to an actual purchase.

CE: WHAT ARE SOME EXAMPLES OF VALID MEASURES?

RM: The number one valid measure is none other than, "How much did we sell?" The trouble is, of course, there are confounding variables. We may have had a fabulous event that lots of people attended, but a competitor dropped prices 20 percent in the week after the event and we didn't sell anything.

That's the challenge. In marketing, the link between cause and effect is not obvious. It's either separated by time or has a bunch of other variables that are in between the two.

A more practical way to measure validity is what we call mother-in-law research — spending quality time with customers. Go out and actually talk to your customers, get a deep understanding of how they think and act. Find out how they respond to marketing and what kind of marketing they respond to. Get to know them deeply and profoundly. You can't do that with a random sample of 10,000 consumers.

I'd rather have a CEO visit 10 consumers and spend a couple of hours with each of them to develop a sophisticated, interpretive understanding of those consumers than to spend time analyzing a customer survey.

Face-to-face interviews are not high on reliability, because the CEO has to use her judgment to say, "I think this is really what they meant." But now, when the CEO sees a statistic that says, "Eighty percent of consumers say blah blah blah," she can say, "When we got to that subject, customers didn't actually divide it up into those three categories like that. They actually thought about it in this way."

When I talk about customer interviews, some people say, "It's incredibly dangerous what you're doing here, Roger. You're going to have a CEO going out and visiting 10 people and making all the corporate decisions based on that. That's not a statistically significant or representative sample."

My response is that you have to gather this valid data as an important input to the way you analyze your reliable data. If not, you sacrifice a lot of depth and nuance. In a reliable study, you're asking 1,000 customers to respond to a 15-minute survey that has 20 questions. They can check A, B, C, or D. They can't give you a paragraph, because if you allowed for that, the survey would take an hour instead of 20 minutes.

CE: WHAT ABOUT FOCUS GROUPS — AREN'T THEY A MORE EFFICIENT WAY TO FIND OUT WHAT CUSTOMERS THINK?

RM: I am not a fan of focus groups. You tend to get lots of anchoring/following behavior. For example, if the first person in the group says, "I hate this aspect of the event," everyone else says, "Me too."

If I had the choice of having a focus group with 10 people for two hours or two one-hour interviews with individuals, I'd choose the interviews. I would get higher-quality insight from that. The only thing I'd use focus groups for are for social products, such as fashion or group vacations, because in that case, the social behaviors of leading and following are relevant and important.

CE: HOW CAN MARKETERS CONVINCE RELIABILITY-ORIENTED MANAGERS OF THE VALUE OF VALID MEASUREMENT?

RM: To convince management that validity is important, you have to make predictions and then prove that your predictions were correct.

To prove the value of valid measurement, marketers should predict what this research will tell them, and how it will help them interpret the reliable data they are collecting.

They can also show management the gaps in the current reliability-oriented system. Valid measurements let marketers evaluate factors that are otherwise difficult or impossible to measure with reliability-oriented research, such as customer intimacy. It allows them to know what their customers' motivations are, how the product makes them feel as a human being, and what emotional and psychological connotations the product has for them.

A more practical way to measure validity is what we call mother-in-law research — spending quality time with customers. Go out and actually talk to your customers, get a deep understanding of how they think and act.