**Harvard
Business
Review**

ASSESSING PERFORMANCE

Most HR Data Is Bad Data

by Marcus Buckingham

FEBRUARY 9, 2015

How good a rater do you think you are? If you were my manager and you watched my performance for an entire year, how accurate do you think your ratings of me would be on attributes such as my "promotability" or "potential?"

How about more specific attributes such as my customer focus or my learning agility? Do you think that you're one of those people who, with enough time spent observing me, could reliably rate these aspects of my performance on a 1-to-5 scale? And how about the people around you – your peers, direct reports, or your boss? Do you think that with enough training they could become reliable raters of you?

These are critically important questions, because in the grand majority of organizations we operate as though the answer to all of them is *yes, with enough training and time, people can become reliable raters of other people*. And on this answer we have constructed our entire edifice of HR systems and processes. When we ask your boss to rate you on "potential" and to put this rating into a nine-box performance-potential grid, we do it because we assume that your boss's rating is a valid measure of your "potential"— something we can then compare to his (and other managers') ratings of your peers' "potential" and decide which of you should be promoted.

Likewise, when, as part of your performance appraisal, we ask your boss to rate you on the organization's required competencies, we do it because of our belief that these ratings reliably reveal how well you are actually doing on these competencies. The competency gaps your boss identifies then become the basis for your Individual Development Plan for next year. The same applies to the widespread use of 360 degree surveys. We use these surveys because we believe that other people's ratings of you will reveal something real about you, something that can be reliably identified, and then improved.

Unfortunately, we are mistaken. The research record reveals that neither you nor any of your peers are reliable raters of anyone. And as a result, virtually all of our people data is fatally flawed.

Over the last fifteen years a significant body of research has demonstrated that each of us is a disturbingly unreliable rater of other people's performance. The effect that ruins our ability to rate others has a name: the Idiosyncratic Rater Effect, which tells us that my rating of you on a quality such as "potential" is driven not by who *you* are, but instead by *my* own idiosyncrasies—how I define "potential," how much of it I think *I* have, how tough a rater I usually am. This effect is resilient – no amount of training seems able to lessen it. And it is large – on average, 61% of my rating of you is a reflection of me.

In other words, when I rate you, on anything, my rating reveals to the world far more about me than it does about you. In the world of psychometrics this effect has been well documented. The first large study was published in 1998 in *Personnel Psychology;* there was a second study published in the *Journal of Applied Psychology* in 2000; and a third confirmatory analysis appeared in 2010, again in *Personnel Psychology*. In each of the separate studies, the approach was the same: first ask peers, direct reports, and bosses to rate managers on a number of different performance competencies; and then examine the ratings (more than half a million of them across the three studies) to see what explained why the managers received the ratings they did. They found that more than half of the variation in a manager's ratings could be explained by the unique rating patterns of the individual doing the rating– in the first study it was 71%, the second 58%, the third 55%.

No other factor in these studies – not the manager's overall performance, not the source of the rating – explained more than 20% of the variance. Bottom line: when we look at a rating we think it reveals something about the ratee, but it doesn't, not really. Instead it reveals a lot about the rater.

Despite the repeated documentation of the Idiosyncratic Rater Effect in academic journals, in the world of business we appear unaware of it. Certainly we have yet to grapple with what this effect does to our people practices. Look closely and you realize that it will cause us to dismantle and rebuild virtually all of them.

Fueled by our belief in people as reliable raters, we take their ratings — of performance, of potential, of competencies — and we use them to decide who gets trained on which skill, who gets promoted to which role, who gets paid which level of bonus, and even how our people strategy aligns to our business strategy. All of these decisions are based on the belief these ratings actually reflect the people being rated. After all, if we didn't believe that, if we thought for one minute that these ratings might be invalid, then we would have to question everything we do to and for our people. How we train, deploy, promote, pay, and reward our people, all of it would be suspect.

And yet, is this really a surprise? You're sitting in a year-end meeting discussing a person and you look at their overall performance rating, and their ratings on various competencies, and you think to yourself "Really? Is this person really a '5' on strategic thinking? Says who – and what did they mean by 'strategic thinking' anyway?" You look at the behavioral definitions of strategic thinking and you see that a "5" means that the person displayed strategic thinking "constantly" whereas a "4" is only "frequently" but still, you ask yourself, "How much weight should I really put on one manager's ability to parse the difference between 'constantly' and 'frequently'? Maybe this '5' isn't really a '5'. Maybe this rating isn't real."

And so perhaps you begin to suspect that your people data can't be trusted. If so, these last fifteen years have proven you right. Your suspicions are well founded. And this finding must give us all pause. It means that all of the data we use to decide who should get promoted is bad data; that all of the performance appraisal data we use to determine people's bonus pay is imprecise; and that the links we try to show between our people strategy and our business

strategy — expressed in various competency models — are spurious. It means that, when it comes to our people within our organizations, we are all functionally blind. And it's the most dangerous sort of blindness, because we are unaware of it. We think we can see.

There are solutions, I'm sure. But I think, before we can even consider those, we must first stop, take stock, and admit to ourselves that the systems we currently use to reveal our people only obscure them. This admission will challenge us. We will have to redesign almost our entire suite of talent management practices. Many of our comfortable rituals — the year-end performance review, the nine-box grid, the consensus meeting, our use of 360's — will be forever changed. For those of us who want HR to be known as a purveyor of good data — data on which you can actually run a business — these changes cannot come soon enough.

Marcus Buckingham is the founder of TMBC, a company that builds strengths-based tools and training for managers. He is the author of several WSJ and NYT bestsellers, including his latest book and accompanying strengths assessment, *StandOut: Find your Edge, Win at Work*.