Predictive Analytics for Human Capital: Think Outside the Lines
By John Boudreau
Appeared in CFO Magazine, April 2015

Lead: Big data and predictive analytics are revolutionizing workforce planning, retention, deployment and performance, but they must be combined with the insight of leaders, because the most direct path to good decision is often not a straight line.

The year 2015 is the 200th anniversary of George Boole's birth. That's significant to HR and talent leaders, who must make sense of the coming tsunami of big data about their human capital investments.

Boole studied idea of necessary and sufficient conditions. For example, a powerful engine is a necessary condition for a fast car, but it's not sufficient. A fast car also requires a suspension system for stability at high speeds. The fact that cars with more powerful engines go faster does not mean that increasing engine power will always make a car speedier. A smaller engine, combined with an optimal suspension system and efficient aerodynamics, can produce a faster car than one with a larger engine. It is a fact that faster cars generally have more powerful engines, but that's not a good design rule. Savvy car designers recognize the fallacy of the straight line logic. The answer lies not simply with increasing engine power, but in the right configuration of engine power, suspension systems, and aerodynamics.

Today, even the most sophisticated applications of big data, analytics and evidence about human capital is often fixated on the straight line, not configurations. For example, a scientific study of auto manufacturing plants found that "high commitment" HR practices -- such as training and incentive pay -- were associated with higher plant productivity and quality. Many leaders and even seasoned data analysts might understandably conclude that more high commitment HR practices will increase plant productivity. That decision risks making the same mistake as the misguided auto designer pursuing speed with ever-larger engines.

Many investments in human capital don't follow a straight line. More is not always better. Effects depend on other conditions. As tempting as the straight line may be, wise leaders will approach their human capital data and analytics thinking about configurations and "it depends," not simply about straight-lines and "more is better."

My colleague, professor Peer Fiss, at the University of Southern California's Marshall School of Business, makes this point in his 2007 article in *Academy of Management Review*. He describes advances in social sciences that recast straight lines into conditions, that more clearly reveal the payoff from human capital and other investments. Such analysis focuses on "sets" and "truth statements," such as "If A and B are both true then outcome Y happens, but if C is also true, then Y does not happen."

The basic idea is not new. It is based on George Boole's ideas, called Boolean algebra, a fundamental pillar of modern technology. Boole developed algebraic rules for set-based questions, where the answers are either true or false. In 1937, Claude Shannon recognized that true-false could describe electrical circuits switched either on or off, and that electrical applications of boolean algebra could resolve any logical numerical relationship. That was the birth of the 1's and zeros that underlie today's digital computation.

Can this powerful idea help leaders with big data, analytics and evidence-based human capital decisions?

Recall the research on how investments in human capital affect manufacturing plant performance. A 2004 study by Kogut, MacDuffie and Ragin examined data on automobile plant performance in the 1990's. They gathered data on these attributes:

- Proportion of production steps that are automated
- Recently updated manufacturing and assembly designs and processes

- Using minimal inventory buffer stocks
- Constructing work that uses emphasize multiskilled workers who can span operations and solve problems as a team
- Using "high-commitment" HR policies such as extensive training and incentive pay

They first analyzed correlations, which extrapolate straight-line relationships, and they found that all of the attributes were positively related to plant productivity (lower labor hours per vehicle). Often, big data analyses are presented only this way, so it is understandable that decision makers might decide to maximize all of these attributes. More is better. If teamwork and high-commitment HR practices were positively correlated with performance, then more of these will increase productivity. The irony is that while this is consistent with the findings, it misses vital information that is revealed only when you approach the data like George Boole would, asking about necessary and sufficient conditions.

The study authors did that, and they found that some conditions were necessary to high performance: All high-performing plants had recently updated manufacturing and assembly processes, plus a high proportion of automation. The other factors of minimal buffers, team-based work and high-commitment HR practices affected productivity only if the first two were present.

When the two necessary conditions were met, three different auto plant configurations emerged:

One configuration combined high levels of teamwork and low buffers, and achieved 18.9 hours per vehicle and 53.5 defects per vehicle
A second configuration had low levels of teamwork but high levels of high-commitment HR practices, and achieved 24.4 hours per vehicle and 43.2 defects per vehicle
A third configuration had high levels of both teamwork and high-commitment HR practices, and had 19.1 hours per vehicle and 44.1 defects per vehicle

Investments in human capital can be key to high performance, but the right investment depends on the conditions.

Today, "predictive analytics" in human capital too often means searching for the straight lines. Yes, higher pay and valued co-workers can reduce employee turnover, but they only work up to a point. Smart leaders will look for inflection points, and that requires thinking about sets and conditions.

How are those sets and conditions defined? It turns out to require a sophisticated combination of human insight and computational analysis. How did the scientists define which automobile plants had "high" versus "low" levels of buffers, automation and teamwork? As Professor Fiss noted, that requires situational knowledge. Computation can help, but thinking about configurations still requires human insights and guidance, well beyond simply mining data for straight-line correlations.

The simple idea that 1's and zero's could represent and solve complicated logical problems ushered in the modern era of computation. Perhaps the same idea offers equally profound implications for big data and the future of human capital decisions.

The tsunami of human capital data that can be a boon or a distraction. Making the most of it requires that human resource professionals, analysts and leaders think outside the lines, and use their experience and insights to see the fuzzy sets and inflection points that reveal the true story.

John Boudreau is professor and research director at the University of Southern California's Marshall School of Business and Center for Effective Organizations, and author of the forthcoming books, "Lead the Work" with Ravin Jesuthasan and David Creelman and *"Global Trends in Human Resource Management"* with Edward E. Lawler III.

###