In Press, *Sage Handbook of Human Resource Management* (2nd edition), Wilkinson, A., Bacon, Lepak, L., & Snell, S. (Eds.).

BIG DATA AND HUMAN RESOURCE MANAGEMENT

Mark Huselid
Distinguished Professor of Workforce Analytics
Director, Center for Workforce Analytics
D'Amore-McKim School of Business
Northeastern University
315B Hayden Hall
Boston, MA 02115
Tel. (609) 651-6484
Email mark@markhuselid.com
www.markhuselid.com
www.centerforworkforceanalytics.org

and

Dana Minbaeva
Professor in Strategic and Global HRM
Copenhagen Business School
Kilevej 14, Copenhagen Frederiksberg
2000 Denmark
Tel. (45) 38152527
Email dm.smg@cbs.dk
www.cbs.dk/hc-analytics

INTRODUCTION

It would be difficult to overstate the influence of *Big Data* across a wide range of business and societal outcomes. In particular, the level of interest in the business community is substantial: the amount of available data is growing exponentially, cloud-enabled computing power has increased rapidly and storage and connectivity costs have dropped precipitously, and there are more and more sophisticated machine-learning techniques that help to translate Big Data potential into value added knowledge (McKinsey, 2017). As a consequence, firms are spending billions of dollars on data and infrastructure, and hundreds of blogs and thousands of Linkedin posts have been written on this topic.

Among the numerous definitions of Big Data (see Gandomi and Haider, 2015, for detailed review), the definition offered by the Gartner IT Glossary is the most prevalent in the literature: "Big Data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making." In addition to 3V's included in the definition (*Volume*, *Velocity*, and *Variety*), other dimensions of Big Data have also been presented in the literature. For example, IBM coined the term *Veracity* as the fourth V, which represents the unreliability, impreciseness and uncertainty inherent in some sources of data[1]. SAS introduced *Variability* and *Complexity* as two additional dimensions of Big Data[2]. Variability refers to the variation in the data flow rates, that is not consistent and has periodic peaks and troughs. Complexity refers to the fact that Big Data are generated through a myriad of sources. This implies another critical challenge - the need to connect, lean and merge data received from different sources. Finally, Oracle introduced *Value*, or rather low value, as a defining attribute of Big Data[3]. The Big Data in its original form has low value relative to its volume. This implies that the Big Data per se is not a strategic resource; the value added comes from analyzing large volumes of such data.

[1] http://www.ibmbigdatahub.com/infographic/four-vs-big-data
[2] https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
[3] https://www.oracle.com/big-data/guide/what-is-big-data.html

With the right analytics, Big Data can deliver richer insights since it draws from multiple sources and transactions to uncover hidden patterns and relationships. Big Data analytics have been applied to real-time fraud detection, complex competitive analysis, call center optimization, consumer sentiment analysis, intelligent traffic management, and the management of smart power grids, to name only a few applications.

Among strategic human resource management (HRM) and strategic human capital scholars, there is also significant interest in the potential for research and new insights using these sources of data and analyses. Recent special issues of the *Journal of Organizational Effectiveness: People and Performance* (Minbaeva, 2017), *Human Resource Management* (Huselid, 2018), books in the popular press (Bock, 2015; Guenole, Ferrar, Feinzig, 2017) and workshops at professional organizations such as the *Academy of Management* (AOM) and the *Society for Industrial and Organizational Psychology* (SIOP) all point to the growing importance of Big Data and analytics, especially in the domain of HRM. Yet, Marler and Boudreau (2017) reviewed the literature and concluded that while the promise may be real, there is much work to be done before Big Data can fulfill its promise for the science and practice of HRM.

We believe that the advent of Big Data in HRM represents both a significant opportunity and a significant challenge for our field. For example, most organizations routinely spend between 50 and 70% of their revenue on the workforce and related expenses (wages, benefits, investments in training and development, etc.), yet that the quality of analytics processes and infrastructure in most organizations is poor (Huselid, 2018). This form of "information failure" can be very costly. Talent (especially top talent) is more mobile than ever, and disruptions and global labor arbitrage have left firms with no choice but to increase their understanding of the quality of their workforce. Markets are changing much more quickly than most firms can adapt, so workforce analytics one of many potential solutions to help them survive and perhaps prosper in the current economic environment.

To address the challenges and opportunities of Big Data for HRM, we believe that both academics and practitioners should address several key questions to move the field forward:

1. Is the Big Data trend a positive development for the field of HRM?
2. Will Big Data and analytics transform the practice HRM as we know it?
3. Where is the biggest value added of Big Data and analytics for HRM?
4. What are the key priorities for the development of workforce analytics?

In this chapter we address these issues and provide a brief overview Big Data in the context of HRM.

IS THE BIG DATA TREND A POSITIVE DEVELOPMENT FOR THE FIELD OF HRM?

We believe that the advent of Big Data provides a potential opportunity, but one that is also fraught with peril if managed incorrectly. In a nutshell, ironically, more data isn't necessarily always a good thing. While one could argue at length whether the Big Data construct is best described by subfactors such as *Volume*, *Velocity*, *Variety, Veracity, Variability*, *Complexity, and Value*, we would argue that for HRM the discourse around Big Data should be concerned about the concept and definition of *Smart Data*.

In their recent editorial in the *Academy of Management Journal*, the editors point out that for management research, "big" is no longer the defining parameter, but, rather, how "smart" it is—that is, the insights that the volume of data can reasonably provide. "For us", they add, "the defining parameter of Big Data is the fine-grained nature of the data itself, thereby shifting the focus away from the number of participants to the granular information about the individual" (George, Haas, and Pentland, 2014: 321).

So what is *Smart Data* for HRM? "I know that we have a lot of HR data, but I do not know what kind of data we have." This is the most common response from managers when asked about their existing HR data. What data do we have? Where do we store our data? How has the data been collected? What rules have been applied? How can two (or more) different datasets be merged into one? What are the advantages and disadvantages of each dataset? Although these are basic questions, most firms do not have the answers.

Such poor organization of firm data can be very costly. When formal, centralized coordination of data collection is lacking, we often see data duplication, wrong entries, etc. Moreover, such a situation makes it impossible to combine different datasets; creates unexplained breaks in time-series/longitudinal data; and leads to data inconsistencies due to the proliferation of various metrics, codings, or time frames. Accordingly, analyses based on such data are rarely comparable or combinable. Answers to complex business problems that rely on the analysis of different variables observed over several time periods and at different organizational levels (e.g., individuals, teams, departments, business units) are difficult to derive. Moreover, firms usually do not collect data documenting changes in the organization (e.g., business-unit reorganizations). However, as organizational change can modify the relationships under study, this failure to model such processes biases the analytics-based decision-making process.

Furthermore, most firms do not necessarily have a full ownership of their own data. That is, most firms do not have access to individual level data gathered by the means of survey by external vendors, often due to contractual arrangements. Accordingly, they cannot connect at the individual level their existing HR data to the collected survey data. A major contributing factor to this is unclear deals the firms make with their external vendors regarding whether the collected data could be returned to the firms in raw form (i.e., as original responses at the individual level). External vendors often attribute this to the need to ensure respondent anonymity. However, in their argumentation, the external providers often do not distinguish between confidentiality and anonymity. The terms "anonymity" and "confidentiality" are often used interchangeably, but they have very different meanings. When data is collected and held "anonymously", there is no identifying information that can link the survey responses to a respondent – not even the researcher can identify a specific participant. In contrast, when data is collected and held "confidentially", the researcher can identify the participants, but that information is kept in a secure environment.

The problem with anonymous survey data is that matching it with other available data can only take place at the group level. As such, any explanatory and causal models accounting for individual variance cannot be developed. Why is this problematic? By averaging the individual responses at the group level, we lose a great deal of explanatory power. This means that we are unlikely to be able conclude anything about the true individual-level antecedents and consequences of employee engagement. In research terms, this is called an ecological fallacy. It occurs when we make conclusions about individuals based only on analyses of group data. Even if we are working with a collective concept that is, by definition, supra-individual (such as Barrick, Thurgood, Smith, & Courtright's 2015 discussion of collective organizational engagement), the individual level data are needed to ensure discriminant validity between aggregated individual-level engagement and collective organizational engagement.

How can this issue be addressed? If a firm promises its employees confidentiality rather than anonymity, there are several ways to handle this issue. Patrick Coolen, HR Analytics Manager at ABN-AMRO Bank, explains: "We partnered with an external partner … in some cases, to protect the anonymity, we are not allowed to handle data at an individual level within our organization. This simply means our external partner can perform richer models and therefore can create better insights than we can internally" (Ignostix, 2016). Employees may trust that third parties will not inappropriately share information with their employer. The aspects of the third-party relationship that support trust in confidentiality include a reputation for independence, explicit rules for research ethics, academic integrity, and traditions.

Another solution is to encrypt the individual-level data. Encryption is the conversion of data into a form that cannot be easily understood by unauthorized people. In practice, one file is created in which individual identifiers are connected with a code. In all other files, the code is used instead of individual identifiers. One person in the company (for example, Data Protection Officer) may have access to this file or it could be held by an external party (e.g., the survey provider or an academic partner).

So, it is not about having more data, but doing more with the data you have. Further, moving from data to actionable information requires that we understand behavioral science theory and ask the *right questions* about how the workforce contributes to your success. In this regard, it is useful to remember that *Big Data Require Bigger Theory*! The typical statistical approach of data mining, searching for significant p-values and moving towards more and more sophisticated econometrics will probably result in a decent, perhaps slightly over-fitted statistical model (since with the immense volume of data, everything is significant), but it is very unlikely it will result in an impactful model for practitioners of in a research publication that would be acceptable by editors and reviewers in top journals. In the above referred editorial of the *Academy of Management Journal* the editors stress: “Given the unstructured nature of most Big Data, causality is not built into their design and the patterns observed are often open to a wide range of possible causal explanations” (George, Haas and Pentland, 2014: 323). The idea is to shift from reporting on what is happening to using rigorous analysis based on a solid conceptual model to help the firms understand and address current challenges and to plan for the future (Davenport, Harris, & Morison, 2010).

WILL BIG DATA AND ANALYTICS TRANSFORM HRM AS WE KNOW IT?

We believe that the advent of Big Data can have substantial and positive implications for the field of HRM, provided that leaders and analysts stay focused on data for decision-making and strategy execution through the workforce. Indeed, for us, a focus on the workforce (in contrast to the HR function) is central to the effective use of Big Data in organizations. This reflects the transition of both academics and professionals to shift their focus from the activities of the HR function (a relatively low value added activity) to measuring the output of the workforce (an activity with much greater potential). Huselid (2018) defines *workforce analytics* as follows:

> *Workforce Analytics refers to the processes involved with understanding, quantifying, managing, and improving the role of talent in the execution of strategy and the creation of value. It includes not only a focus on <u>metrics</u> (e.g., what do we need to measure about our workforce?), but also <u>analytics</u> (e.g., how do we manage and improve the metrics we deem to be critical for business success?)*

As we have said elsewhere, workforce analytics is both a very new and a very old discipline (Becker, Huselid, & Ulrich, 2001; Huselid, 2018). Managers have been making decisions about the workforce many years – whom to hire, how to appraise performance, whom to promote, etc. What is new about Big Data is that we not have the potential to improve the quality of these decisions substantially. In short, the major potential contribution of Big Data is not the data *per se*, but rather the insights and intelligence that the data can potentially generate.

From an historical perspective, the field of workforce analytics is rooted in the conventional disciplines of economics, statistics, social psychology, law, and of course, HRM. So, how can we understand the impact of Big Data on the field of HRM? Can we expect the impact to be one of evolution or of revolution? Like any other business function, HRM is exposed to various disruptive forces that push business function to transform themselves.

Consider a case of introduction of strategic workforce planning (SWP) and devolving the workforce planning decisions to line managers. SWP is technological tool that systematically forecasts risks, finds the right balance of quantity, quality, and location of critical talent, and pinpoints internal supply of and demand for critical skills and roles under multiple business scenarios. It can be developed in-house (e.g. Novo Nordisk's analytics team developed such tool for the whole organization using just Excel) or in-sourced from external providers (e.g. such solutions are offered by various external providers). When introduced properly, SWP presents a unique case of how HR technological advancements and easy access to actionable analytics push people-related decisions out of the hands of HR professionals and into the hands of line managers. Minbaeva (2017) noted that

> *"With the introduction of strategic workforce planning and actionable analytics, do line managers need HR business partners to discuss the changes in their workforces driven by market growth and talent supply? Would line managers prefer to obtain their figures by playing with scenario planning in the strategic workforce planning application? Given the expansion of digitalization and the rise of e-HR, what should be outsourced to robots or automated, and what should be kept for HR? How will the rise of analytics shape the employable HR profile over the next three to five years? "*

In summary, we believe that capitalizing on the opportunity afforded by Big Data will require changing the mindset of HRM. We argue that these tremendous advancements in information technology, disruption of the main business processes and growing stakeholder expectations for economic gains pose *significant challenges to HR*, but also offer *tremendous opportunities for reinventing HR* for organizational value creation. "Technology and analytics are needed to translate data, because deciding on human capital value is no different from deciding on capital investments in the business with an expected return on investment" (EY, 2016a: 2). To rise to this occasion and meet these higher expectations "many HR legacy mind-sets that may have been true in the past need to evolve to modern realities" (Ulrich, Schiemann, & Sartain, 2015: 2).

Although firms are improving their abilities to act on the results of their analytics, too few collect data focused on the consequences of their analytics-based decisions and actions. What actions have been taken and where? How they have been operationalized? What changes are evident in the variables? The formal analysis of follow-up data reveals the effectiveness of the decisions and actions, helps identify how actions can be modified or changed to better achieve the expected output, and highlights those actions that are actually harmful and should therefore be stopped.

In HRM, the situation is very different, and an atheoretical (or unmonitored) search for results with "statistical significance" can be ill-informed, or perhaps even illegal. For example, one of us was told (with great enthusiasm) by a workforce analytics specialist that their analyses showed that single, white males in their firm had the highest performance evaluation ratings, the highest salaries, and received the highest raises (both in percentage and absolute dollar values). As a consequence, this analyst suggested that the organization should consider devoting more resources to this group, because of their "obvious" higher performance and potential. It hadn't occurred to this person that correlation most certainly does not equal causation, and that there were a range of alternative explanations for these findings, beginning with the firm's own biases in the recruiting, selection, development, and promotion processes.

WHERE IS THE BIGGEST VALUE ADDED OF BIG DATA AND ANALYTICS FOR HRM?

The great irony of the advent of Big Data is that more data have the potential to distract rather than inform. The danger is that that we get distracted and overwhelmed by the availability of data, and pursue avenues of research that are either not focused on strategy execution or not supported by the previous research. Just as with any form of scientific inquiry, some questions are much more important than others, and not all questions warrant the investment of time and resources needed to generate a quality answer. What is important to understand is that data (of any size, large or small) are only valuable to the extent to which they can create new insights and knowledge for business.

Sander's (2018) review of the implications of Big Data for supply chain management makes some important points with specific relevance for HRM. Based on interview and survey data from executives in over 300 firms, Sanders concluded that the advent of Big Data has created three new domains of opportunity for leaders:

- *Opportunity for Inquiry with Big Data*. Sanders makes the point that the availability and quality of data allows both scholars and practitioners to explore questions and issues in a way that simply wasn't possible even a decade ago. For Sanders, the sheer type and variety of data sets available make it possible to explore a wide range of potential questions.
- *Changing the Nature of Inquiry with Big Data*. For Sanders, Big Data allow old questions to be asked and answered with much greater speed, but also to change the way in which questions are asked and answered. Extremely large datasets, low storage costs, and very high computational speed have allowed the development of machine learning algorithms that enabled the exploration of new questions in new ways.
- *Changing the Nature of Experimentation with Big Data*. Finally, Sanders notes that Big Data allow us to exploit the naturally occurring field experiments that occur in every organization. This is perhaps the most important difference with relevance to workforce analytics, because it will allow us to potentially assess causality in ways that wasn't possible before.

For HRM, the biggest value added of Big Data and analytics will be around the key unanswered question in HRM: does HRM pay off? Couple of years ago, the cover story of Harvard Business Review claimed "It's time to blow up HR and build something new".

As Capelli (2015) explains "HR managers focus too much on "administrivia" and lack vision and strategic insight" (p. 56). Another article in the same issue highlighted the problem that HR has a tendency to fall in love with the problem, not the solution, thereby focusing too little on *the actual value* of HR initiatives and *their contribution* to fulfilment of organizational goals (Boudreau & Rice, 2015). Big Data and analytics offer a possibility to demonstrate HR's actual value and contribution, thereby making HR a more credible partner for the business. As Green (2017: 137) argues, "successful people analytics teams focus on projects that actually matter for business." To be viewed as a valuable partner for the business, HR must speak a language that stakeholders - the language of value creation. As Ed Iames, Wawa Inc.'s Senior Director of HR, says: "We've found that the more data we [HR] produce and send to our business partners, the more questions we get and the more they want. They become very engaged with what we are doing, very engaged with the solutions."[4]

The advent of Big Data and analytics will also help HRM to move away from treating all employees equally towards starting to treat them equitable (Becker, Huselid and Beatty, 2009). For example, analytics in talent management can provide input to the core talent management decisions: (1) the identification of pivotal or strategic positions within the organization that have the potential to affect organizational performance; (2) the identification of a talent pool (both external and internal) to fill those positions; and (3) the monitoring of talent performance and active management of talent retention (Minbaeva and Vardi, 2018). Similarly, well set-up and executed analytics projects will help HR to create "a clear sense of the HR management practices (selection, development, performance management, and so on) that you [the organization] wish to *improve* vs. those you would like to *do differently*" (Becker, Huselid and Beatty, 2009: 129). This will ultimately lead the organizations towards building differentiated HR architecture and enable them to effectively execute their strategies.

So, where will the potential impact of workforce analytics be the greatest? What should we measure, and how? While the answers to these questions will almost certainly differ

---

4 http://www.nugress.com/resources/images/HR%20Analytics%20%20Gaining%20Insights%20for%20the%20Upturn%20[1].pdf.

by firm, the key point we would make is that business logic drives measurement, and for us this means that the metrics and analytics a firm develops should be focused on executing the firm's strategy. More specifically, we believe that workforce analytics will have the greatest impact in when they are focused on the firm's strategic work embedded in strategic jobs (Becker, Huselid, and Beatty, 2005; Huselid, 2018). These jobs may appear at any point in the firm's value chain, and they exhibit two key attributes. First, they are almost always located within one of the firm's most essential strategic capabilities (i.e., supply chain analyst in a logistics firm). And second, there is substantial variability in the performance of the individuals holding those roles – that is, the range in performance from the top to the bottom is very large. It is this unique combination of *importance* and *opportunity* that makes strategic jobs such a priority for both the development of analytics and improvement by managers.

WHAT ARE THE KEY PRIORITIES FOR THE DEVELOPMENT OF WORKFORCE ANALYTICS?

Develop analytical <u>competencies</u> at the level of the <u>individual</u>. Effective implementation of analytics programs requires a wide range of skills and abilities, some of which are likely to be resident in most well-developed HR functions, some may need to be "borrowed" from other functional areas (e.g., marketing, accounting, finance, supply chain, etc.), while still others will need to be developed internally or brought in from the outside.

The point we wish to make here is that world class analytics don't just occur on their own – they are created by competent and capable leaders who know and understand workforce analytics. In Becker, Huselid, and Ulrich (2001) we argued that effective workforce analytics design and implementation require HR leaders with the following skills (in addition to general HR manager competencies):

- Critical causal thinking
- Principles of good measurement
- Estimating causal relationships
- Communicating HR strategic performance results to senior line managers

Taking this idea further, Kryscynski, Reeves, Stice-Lusvardi, and Ulrich (2018) tested found in a sample of 1,117 HR professionals from 449 organizations, they that HR professionals with higher levels of analytical capabilities out-performed their peers.

Clearly, analytical competencies matter, and the field of analytics is growing rapidly (Davenport and Patil, 2012). While this is a positive development for HRM, it is also important to form an analytics team with great care. Managers training analytics may not have much experience in the science and practice of HR, while points to the importance of a breadth of skills on the analytics team, and also a focus on the organizational level of analysis when considering investments in analytics capabilities.

***Develop analytical <u>capabilities</u> at the level of the <u>organization</u>*.** In addition to changing the mindset it will also require new *organizational-level capabilities*, which are built upon a foundation of *individual competencies*. Minbaeva (2018) defined Human Capital Analytics (HCA) as an "*organizational capability that is rooted in three micro-level categories (individuals, processes, and structure) and comprises three dimensions (data quality, analytical competencies, and strategic ability to act)."* (p.701) She argued that at three different levels – individual, process and structure levels – developing of HCA as organizational capability requires different components, as well as interactions within and across components:

- *Individual:* (a) having committed individuals to ensure flawless data organization; (b) acquiring and developing analysts with needed KSAs; and (c) encouraging boundary-spanning behavior outside of the HCA team.
- *Process*: (a) building systems and establishing workflows to continuously support data quality, (b) linking the results of analytics projects with existing organizational processes, and (c) encouraging experimentation and enabling follow-up actions via HRBPs.
- *Structure:* (a) continuous investments in formal, centralized coordination of data collection and organization; (b) creating a culture of inquiry and a habit of making evidence-based decisions; and (c) equipping top management with tools for action, which should be linked to current and future strategy discussions.

Related to this is a discussion around where analytics should be placed within the organization: should it belong to HR, line managers, or business intelligence? Andersen (2017) weighs the pros and cons of moving analytics outside of HR. Van den Heuvel and Bondarouk (2017) argue that moving analytics to the HR department or to a general business-intelligence department is much more desirable. Regardless, where the analytics function is based it should fulfill boundary-spanning roles, bridging between HRBPs, line managers, and executive team. The interdependency between analysts and HRBPs is crucial since articulating a business problem in analytical terms is a joint effort between HRBP and analyst. However, direct links with business and line managers are also needed since the communication of the defined problem and interpreting the results would happen directly between the business leaders and the analyst. Finally, having a support from executive team is also crucial. Green (2017) warns: "Without CHRO and senior executive involvement your people analytics adventure is likely to be doomed from the start" (p. 172). Boudreau and Cascio (2017) also point out that "a fundamental requirement is that HCA address key strategic issues that affect the ability of senior leaders to achieve their operational and strategic objectives" (p. 122). Analyzing Shell's analytics journey, Minbaeva (2017) concludes: " … one of the decisive factors for the success of Shell's analytics journey is the close cooperation between Jorrit van der Togt, the Executive Vice President of HR Strategy and Learning, and Thomas Rasmussen, the Vice President of HR Data and Analytics, as well as the strong support from the senior business leaders" (p. 114).

Understand business problems and translate them into questions about the workforce. Perhaps the most important counsel we can provide is that it is workforce measures and analytics are intended to provide answers to questions, presumably about the quality and progress of the workforce in support of the firm's strategy. Asking the right questions about how the workforce contributes to firm success is among the most important things that an analytics team can do.

This can be more of a challenge than it might appear, especially in the context of a business (or leadership team) that is pressuring the HR function to "do something" about analytics and provide results quickly. We have worked with a number of analytics teams that have actually impeded their own long-term progress by moving too quickly to the data analysis phase. The rationale we most often hear is that there will be time "later" to go back and collect the "right" data, but for now it is important to "do something." The answers generated by this approach are either not compelling or simply incorrect, and the analytics team loses credibility and line managers lose interest in the concept.

We can provide two brief examples of how a focus on the "data we have not the data we need" can distract an analytics team's attention from the ultimate goal of helping leaders make better and evidenced-based decisions about the workforce. The first example has to do with benchmarking common HR processes, such as time to fill an open position or cost per hire. Measuring HR function activities such as these is very appealing to leaders because it seems straightforward and relevant; who could argue that we shouldn't try to fill open positions quickly and efficiently? Unfortunately, a decrease in time to fill an open position is frequently associated with lower candidate quality and ultimately, higher costs and lower organizational performance (Becker and Huselid, 2003). How could the firm address this problem? Instead of measuring time to fill, some firms we are working with measure time to competence, or time to first promotion. Others measure performance at the 1, 2, and 5-year work anniversaries as a measure of recruiting competence. These time-lagged measures are more complex than simple time to fill measures, but ultimately they are a much better fit for the recruiting construct.

The second example has to do with an over-reliance on ERPs and data warehouses as a data source for workforce analytics. Part of the Big Data implementation process in many firms is development and installation of system-wide data wharehouses that are intended to integrate not only the functional areas within HR (e.g., linking their performance management and compensation systems), but also linking all of these systems with data in the other functional areas, e.g., marketing, sales, supply chain, finance, etc. This sound to be an ideal situation for the workforce analyst, and it can be – but the devil is often in the details with this type of system. Given the scope, magnitude, and costs associated with these systems, there is enormous pressure to standardize the data feeds and elements for the workforce. Customization of the software to meet the needs of the workforce analysts, especially after is has been installed, is often staggeringly expensive. To avoid this type of problem, we believe it is important for the workforce analytics team to be involved in the design and implementation of the system at the outset.

The point of these examples is that it is crucial for workforce analysts to focus on collecting the *relevant* data, not on analyzing *available* data. First determine *what to measure – then* collect reliable and valid data. As Becker, Huselid, & Beatty (2009) have described, the process needs to start with the development a clear statement of the strategic capabilities (bundles of information, technology, and people) that are needed to execute firm strategy. As we have said above, the greatest potential opportunity to impact firm performance is likely to be located in (some, very specific) *strategic positions*.

Once this is done, it is important that someone on the team look to the literature to see what is known about a topic. How is the performance of our project managers measured? What do we know about the predictors of their performance? How difficult are these predictors to change or influence? In short, it is very important to read the research and build a theory or model that shows *what causes what* in your organization. In the long run this will save a tremendous amount of time and energy. Moreover, these analyses should be focused on the entire work system, as and not just individual HR policies or practices (Levinson, 2018).

Most of the focus in the domain of workforce analytics has to do with quantitative data, with performance appraisal data, salary data (base, bonus, other incentives) and employee movements (quits, promotions, etc.). These data are relatively easy to acquire. However, much of the interesting and important analytics data is qualitative in nature, and firms are generally much less skilled in dealing with these data (Gandomi and Hader, 2015).

Finally, we also use managers to develop and install audit functions for workforce analytics. Audit procedures are commonplace in many organizational functions, and we believe that they are especially important in the domain of workforce data because most because the data collections processes will be both newly created and widely distributed throughout the firm, thus increasing the possibility of errors in the process.

Match the rigor of the data quality and analyses with the importance of the question. Our experience is that the most quantitative part of the process (estimating statistical relationship among variables) is actually the easiest and least controversial part of the process. There is an extremely well-developed literature in psychometrics (how do we measure, for example, employee attitude such as satisfaction, job involvement, or engagement), and statistics and econometrics (how do we assess relationships among variables once we have measured them).

One key point to keep in mind is that workforce outcomes (e.g., performance, turnover, satisfaction) is not the results of just a single driving factor. Rather, work outcomes are determined by a variety of factors, to the way in which we think about and model those outcomes will need to be multivariate as well (Huselid, 2018). Managers should be wary of simple correlations in organizations. Only focusing on the relationship between, for example, engagement and performance, is likely to overstate the importance of engagement in the model. Instead, manager should focus on multivariate models such as regression or network analyses (Robinson, 2018), using predictors that have been identified in the prior (extensive) body of research in HRM.

In addition, another defining characteristic of measuring and managing the impact of the workforce on firm success is that the effects of the workforce are nested or multilevel in nature. For example, employees work together in teams, which develop a product or service (or supports the development of), which then influences the production, merchandising, and distribution processes, which then affect customer sentiment and purchase (and repurchase) behavior, which then turns into sales and cash flow, and ultimately, profitability and shareholder value. The rich, multi-level nature of this research can also be modeled with existing research techniques (Gibson, 2017). The reliance on a single-level view yields an "incomplete understanding of behaviors occurring at [any] level" (Hitt et al., 2007, p. 1385). We believe that firms that can understand and act to improve the direct and indirect ways that employees affect firm value can achieve a source of competitive advantage that will be very difficult to replicate.

Influence the right decisions through workforce analytics, and build an evidence-based decision culture. Workforce analytics without managerial influence represents a substantial missed opportunity. Thus, it is very important to develop an implementation plan to ensure that workforce data and analytics are used to help execute strategy and to improve the quality of our workforce. Managers need help in focusing and prioritizing their workforce decisions and investments, and to make better decisions about the firm's most expensive (and valuable) resource.

In this context, Big Data and the analytics team can help managers to focus their decisions buy collecting and presenting data on the extent of workforce success. Data visualization software and HR or workforce scorecards help managers understand complex and often nuanced data. At the level of the HR function, one approach is the HR Scorecard (Becker, Huselid, & Ulrich, 2001), while metrics for the broader workforce can be presented in a workforce scorecard (Becker, Huselid, & Beatty, 2009). Whatever the approach, HR leaders and decisionmakers need to understand 1) what is the specific process through which the workforce affects our success, and 2) how are we doing on those elements and where can we improve.

Work to Integrate the Academic-Practitioner Gap in Workforce Analytics. The final point that we would make is that we believe it very important for the scholarly and practitioner communities to work closely together as the field evolves. In our work with the *Human Capital Analytics Group* at the Copenhagen Business School (Minbaeva) and the *Center for Workforce Analytics* at Northeastern University (Huselid), we have observed numerous cases where the applied analytics teams made substantial mistakes because they were not aware of the prior research on a topic or the appropriate analytical tools. Similarly, we have also worked with analytics teams who were exceptionally advanced and were performing analyses much more sophisticated than has typically appeared in the literature – so much so that they were hesitant to publicize their work because they felt it was a potential source of competitive advantage. Clearly, both the academic and practitioner communities have much to learn from each other (Simon, 2018).

We also believe that the Big Data trend represents a significant potential opportunity for HRM scholars to conduct new and innovative research that simply wasn't possible even a short time ago. Workforce analytics exists within the context of broader business analytics. HR function analytics are likely to be a subset of workforce analytics, but they don't have to be. For example, prior research on the impact of HR management systems on firm performance can help firms position their work in the context of the broader business and its strategy (Combs, Liu, Hall, and Ketchen, 2006; Huselid, 1995).

For scholars, we believe that it is very important to reach out to practitioners who are doing this work in organizations. Like the problems experienced by evidenced-based management scholars, scholars can firms to understand what we know about the HR practice – talent – customer outcomes – firm-level outcome relationships, and then translating these findings into a form and structure easily accessible to practitioners developing and implementing workforce analytics. As such, we believe the field of workforce analytics will face many of the same challenges and obstacles as does evidence-based management, especially in the process of translating existing research into testable internal research design (Rynes and Giluk, 2007). Excellent examples of this process can be found in case studies at *Google* (Bock, 2015), *Jack in the Box* (Schiemann, Seibert, and Blankenship, 2018) and at Zara (Simon, 2018).

CONCLUSION

We began this chapter with a focus on four broad questions:

1. Is the Big Data trend a positive development for the field of HRM?
2. Will Big Data and analytics transform HRM as we know it?
3. Where is the biggest value added of Big Data and analytics for HRM?
4. What are the key priorities for the development of workforce analytics?

Our conclusions were that Big Data in the domain of HRM has the potential to contribute substantially to effective workforce management and ultimately firm success, but much of this potential is as yet unrealized. Our analyses show that that the shift toward workforce analytics (or the broader construct of evidenced-based management) represents a real and enduring transition. But is this transition real or a fad (Rasmussen and Ulrich, 2015)? Only time, of course, can inform the final outcome. But we believe that workforce analytics will represent a significant shift in HR management, 1) meets a significant managerial need, and 2) is at its core based on fundamental social science research methods that are well understood and well proven. Perhaps the most important contribution is to help managers to develop a causal understanding of the role of their workforces in their firm's success, and then to act on this information. However, in most firms there is much work still to be done.

The challenge for both scholars and practitioners is to manage the signal to noise ratio carefully, and not get distracted by data and questions that are not relevant to the firm's overall success. The HR team cannot pull the analytics challenge alone, and the most effective organizations build a specific organizational capability in analytics through the creation of interdisciplinary teams. Broad, integrated business problems will require an equally broad and competent analytics team to address them

Analytics can drive that significant makeover that HR needs (Cappelli, 2015). HR has a tendency to fall in love with the problem, not the solution, thereby focusing too little on *the actual value* of HR initiatives and *their contribution* to fulfilment of organizational goals (Boudreau & Rice, 2015). "A critical analysis of many HR functions today would reveal between 60 per cent and 80 per cent of activity and associated cost remains focused on what are primarily transactional or compliance-based activities, suggesting the function may not be that different to what it was 30-plus years ago" (EY, 2016b: 1). We believe that carefully designed workforce analytics can go a considerable distance in closing this gap.

REFERENCES

Barrick, M.R., Thurgood, G.R., Smith, T.A., & Courtright, S.H. 2015. Collective organizational engagement: Linking motivational antecedents, strategic implementation, and firm performance. *Academy of Management Journal*, 58(1), 111–135.

Becker, B.E., & Huselid, M.A. (2003). Measuring HR's Performance? Benchmarking is Not the Answer! *HR Magazine*. December, 57-61.

Becker, B.E., Huselid, M.A., and Ulrich, D. (2001). *The HR Scorecard: Linking People, Strategy, and Performance*. Cambridge, MA: Harvard Business Press.

Becker, B., Huselid, M., & Beatty, R. (2009). *The Differentiated Workforce: Transforming Talent Into Strategic Impact*. Harvard Business Press, Boston, MA.

Becker, B. E., & Huselid, M. A. (2006). Strategic human resources management: Where do we go from here? *Journal of Management*, 32(6), 898–925.

Bock, L. (2015). *Work Rules!: Insights from Inside Google That Will Transform How You Live and Lead*. New York: Twelve Books.

Boudreau, J., & Rice, S. 2005. Bright, shiny objects and the future of HR. *Harvard Business Review*, July–August, 72–78.

Boudreau, J., & Cascio, W. (2017). Human capital analytics: why are we not there? *Journal of Organizational Effectiveness: People and Performance*, 4(2), 119–126.

Cappelli, P. (2015). Why we love to hate HR….And what HR can do about it. *Harvard Business Review*, July-August,

Combs, J., Liu, Y., Hall, A., & Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59: 501-528.

Davenport, T.H., and Patil, D. J. (2012). Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, October, 70–76.

Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at Work: Smarter Decisions, Better Results*. Boston: Harvard Business School Press.

Deloitte (2017). *Rewriting the Rules for the Digital Age: 2017 Deloitte Human Capital Trends.* New York: Deloitte University Press.

Economist Intelligence Unit. (2016). *Use of Workforce Analytics for Competitive Advantage*. Washington, DC: SHRM Foundation.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big Data concepts, methods, and analytics. *International Journal of Information Analytics, (35),* 137-144.

Garvin, D. A. (2013). How Google sold Its engineers on management. *Harvard Business Review*, 21, 75–82.

George, G., Haas, M.R., & Pentland, A. 2014. Big Data and management. *Academy of Management Journal, 57(2*), 321-326.

Gibson, C.B. (2017). Elaboration, generalization, triangulation, and interpretation: On enhancing the value of mixed method research. *Organization Research Methods*, 20(2), 193-223.

Guenole, N., Ferrar, J., and Feinzig, S. (2017). *The Power of People: Learn How Successful Organizations Use Workforce Analytics to Improve Performance*. New York: Pearson.

Huselid, M.A. (2018) The science and practice of workforce analytics: Introduction to the HRM special issue. *Human Resource Management, 57*, 679-684.

Huselid, M.A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, *38*, 635-672.

Huselid, M. (2015). Workforce analytics for strategy execution. In Ulrich, D., Schiemann, W.A., and Sartain, L. (Eds.) *The Rise of HR: Wisdom from 72 HR Thought Leaders*. Human Resource Certification Institute.

Huselid, M.A., & Becker, B.E. (2005). Improving HR's Analytical Literacy: Lessons from *Moneyball.* In *The Future of HR: 50 Thought Leaders Call for Change,* page

278-284. Dave Ulrich, Mike Losey, and Sue Meisinger (Eds). New York: John Wiley and Sons.

Huselid, M.A., Becker, B.E., & Beatty, D. (2005). *The Workforce Scorecard: Managing Human Capital to Execute Strategy*. Cambridge, MA: Harvard Business Press.

Huselid, M. A., Beatty, R. W., & Becker, B. E. (2005). A Players or A Positions? The strategic logic of workforce management. *Harvard Business Review*, 83(12), 110–117.

Ignostix. (2016). http://www.inostix.com/blog/en/the-hr-analytics-journey-at-abn-amro-interview-with-patrick-coolen/

Kryscynski, D., Reeves, C., Stice-Lusvardi, R,, Ulrich, M., and Russell, G. (2018). Analytical abilities and the performance of HR professionals. *Human Resource Management, 57*, 715-738.

Levenson, A. (2018). Using workforce analytics to improve strategy execution. *Human Resource Management.*

Levenson, A. (2017). *Strategic Analytics: Advancing Strategy Execution and Organizational Effectiveness*. New York: Barrett-Koehler Publishers.

Marler, J.H. & Boudreau. J.W. (2017) An evidence-based review of HR Analytics, *The International Journal of Human Resource Management*, 28:1, 3-26.

McKinsey Global Institute (2016). *The Age of Analytics: Competing in a Data Driven World*. Brussels: McKinsey & Company.

McKinsey and Company. (2017). https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-to-win-in-the-age-of-analytics?cid=other-eml-ttn-mip-mck-oth-1801

Minbaeva, D.B. (2018). Building credible human capital analytics for organizational competitive advantage. *Human Resource Management, 57*, 701-713.

Minbaeva, D.B. (2017). Introduction to the special issue. *Journal of Organizational Effectiveness: People and Performance*, 4(2), 110-118.

Rasmussen, T., & Ulrich, D. (2015). Learning from practice: How HR analytics avoids being a management fad. *Organizational Dynamics*, 44(3), 236–242.

Robinson, M.A. (2018). Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management, 57*, 739-750.

Rynes, S.L., and Gulik, T.L. (2007). The very separate worlds of academic and practitioner publications in human resource Management: Implications for evidenced-based management. *Academy of Management Journal, 50*, 987-1008.

Sanders, N. (2016). How to use Big Data to drive your supply chain. *California Management Review*, 58(3), 26-48.

Schiemann, W.A., Seibert, J.H., and Blankenship, M.H. (2018). Putting human capital analytics to work: Predicting and driving business success. *Human Resource Management, 57*, 795-807.

Simon. (2018). Workforce Analytics: A case study of scholar-practitioner collaboration. *Human Resource Management, 57*, 781-793.

Ulrich, D., Schiemann, W. A., & Sartain, L. (2015). *The rise of HR: Wisdom from 73 thought leaders*. Alexandria, VA: HR Certification Institute.

van der Laken, P., Bakk, Z., Giagkoulas, V., van Leeuwen, L., and Bongenaar, E. (2018). Expanding the methodological toolbox of HRM researchers: The added value of latent bathtub models and optimal matching analysis. *Human Resource Management, 57*, 751-760.

van der Togt, J., & Rasmussen, T. H. (2017). Toward evidence-based HR. *Journal of Organizational Effectiveness: People and Performance*, 4(2), 127–132.