

6-2018

# Big Data Techniques and Talent Management: Recommendations for Organizations and a Research Agenda for I-O Psychologists

Michael C. Champion

Michael A. Champion

Emily D. Champion

*Old Dominion University*, [ecampion@odu.edu](mailto:ecampion@odu.edu)

Follow this and additional works at: [https://digitalcommons.odu.edu/management\\_fac\\_pubs](https://digitalcommons.odu.edu/management_fac_pubs)



Part of the [Business Administration, Management, and Operations Commons](#), [Industrial and Organizational Psychology Commons](#), and the [Organizational Behavior and Theory Commons](#)

---

## Repository Citation

Campion, Michael C.; Champion, Michael A.; and Champion, Emily D., "Big Data Techniques and Talent Management: Recommendations for Organizations and a Research Agenda for I-O Psychologists" (2018). *Management Faculty Publications*. 31. [https://digitalcommons.odu.edu/management\\_fac\\_pubs/31](https://digitalcommons.odu.edu/management_fac_pubs/31)

## Original Publication Citation

Campion, M. C., Champion, M. A., & Champion, E. D. (2018). Big data techniques and talent management: Recommendations for organizations and a research agenda for I-O psychologists. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 11(2), 250-257. doi:10.1017/iop.2018.14

- Lee, S., Han, S., Cheong, M., Kim, S. L., & Yun, S. (2017). How do I get my way? A meta-analytic review of research on influence tactics. *The Leadership Quarterly*, 28, 210–228.
- Lefkowitz, J. (2008). To prosper, organizational psychology should ... expand the values of organizational psychology to match the quality of its ethics. *Journal of Organizational Behavior*, 29, 439–453.
- Lefkowitz, J. (2010). Industrial-organizational psychology recurring identity crises: It's a values issue! *Industrial and Organizational Psychology*, 3, 293–299.
- Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2, 34–46.
- Rotolo, C. T., Church, A. H., Adler, S., Smither, J. W., Colquitt, A. L., Shull, A. C., ... & Foster, G. (2018). Putting an end to bad talent management: A call to action for the field of I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 11(2), 176–219.
- Salas, E., Kozlowski, S. J., & Chen, G. (2017). One hundred years of the *Journal of Applied Psychology*: Background, evolution, and scientific trends. *Journal of Applied Psychology*, 102, 237–253.
- Sikula, S. A., Olmosk, K., Kim, C. W., & Cupps, S. (2001). A “new” theory of management. *Ethics & Behavior*, 11, 3–21.
- Weick, K. E. (1984). Small wins: Redefining the scale of social problems. *American Psychologist*, 39, 40–49.
- Yost, P. R., McLellan, J., Ecker, D., Chang, G. C., Hereford, J., Roenicke, C., ... Winberg, Y. (2011). HR interventions that go viral. *Journal of Business and Psychology*, 26, 233–239.

## Big Data Techniques and Talent Management: Recommendations for Organizations and a Research Agenda for I-O Psychologists

Michael C. Campion  
Campion Services  
University of Texas Rio Grande Valley

Michael A. Campion  
Purdue University

Emily D. Campion  
Old Dominion University

Big data and its applicability to talent management (TM) as defined by Rotolo et al. (2018) has already been recognized by many *outside* the field of I-O psychology. The market is beginning to include offerings from vendors for products that use some combination of big data techniques to process

Michael C. Campion, Campion Services and University of Texas Rio Grande Valley; Michael A. Campion, Purdue University; Emily D. Campion, Old Dominion University.

Correspondence concerning this article should be addressed to Michael C. Campion, Campion Services, 3336 Dubois St., West Lafayette, IN 47906. E-mail: [michael@campion-services.com](mailto:michael@campion-services.com)

vast amounts of data or previously unanalyzable data, which they claim will improve components of TM for organizations. Unfortunately, as noted in the focal article, this “frontier” issue makes it difficult for organizations to separate the wheat from the chaff. Further, with few exceptions, I-O psychology is just beginning to inform organizations about whether and how big data can be used for the purposes of TM.

For the purposes of this article, we define *big data techniques* as those that use advanced computer programs that apply a wide range of statistical and other analytic frameworks and procedures, including text mining, to analyze large datasets to discover relationships, create models, and predict outcomes to help make decisions in TM. Currently, most examples exist in selection contexts, but the potential applicability of these techniques could range far broader to many other human resource practices such as job analysis, performance and succession management, turnover prediction, engagement surveys, and so on.

In addition to our own experience researching the use of big data techniques in staffing contexts (e.g., recruitment, scoring of essays, interviewing; e.g., Campion & Campion, 2018; Campion, Campion, Campion, & Reider, 2016) and ongoing consulting projects where we develop and administer such systems, we are often asked to advise organizations who are considering big data products from vendors. In that role, we serve as outside evaluators to help determine whether the companies should adopt the products and how to improve the quality (and prove the value) of those products.

Based on this, we respond to the recommendation of the focal article to end bad talent management by helping to preemptively prevent the adoption of big data products in the selection domain that are not reasonably grounded in science. First, we present a list of recommendations for organizations to consider when faced with an opportunity to adopt new technology from outside vendors that claim to enable an organization to leverage big data for the purposes of TM. Second, we provide a research agenda for I-O psychologists that attempts to ensure that future research aligns with important questions organizations have or will have moving forward regarding whether to, or how best to, leverage big data for TM purposes. This agenda is not intended to be exhaustive but rather to focus on topics of more immediate importance in terms of their potential to inform organizations. These recommendations and agenda are provided in Tables 1 and 2, respectively.

**Table 1. Recommendations to Improve the Application of Big Data to Talent Management****Basic understanding of the product**

1. Avoid “black-box” proposals. If consulting firms cannot explain their processes to you, then either they do not understand the processes themselves or they are hoping to figure them out as they apply them to your organization. An I-O psychologist has the training to understand these concepts and procedures; they are not only interpretable by those in the emerging field of “data scientists.”
2. Determine whether you are buying a product or a research study. Determine the extent to which the vendor has figured out the product completely, or whether you are buying basic research to create a product. There is nothing wrong with the latter, but just be clear.
3. Insist on knowing what the process measures. Components of normal content and construct validation approaches probably still apply, even if their application might not be as straightforward as validating a traditional assessment. There may be an analogy between complex big data models applied to application data and research on empirically scored biodata. The lack of construct validity in biodata probably impeded its adoption, despite its superior prediction compared to rationally scored biodata. As a different example, in the domain of interviewing, some vendors are exploring the possibilities of developing products that model everything from facial features and eye contact to speech patterns and voice pitch.

**Validation issues**

4. Insist on proven success regarding what products and/or process are proposed, not vague self-report success stories. Watch out for exaggerated promises, of course. Insist on proof in your organization. Do not buy a product supposedly “validated” in other organizations alone. Recall that a vendor’s claims of validity are insufficient by themselves according to the *Uniform Guidelines on Employee Selection Procedures* (1978).
5. Normal validation requirements still apply, and consider associated legal challenges. Make sure these are covered in the proposal; the Equal Employment Opportunity Commission (EEOC) and Office of Federal Contract Compliance Programs (OFCCP) will not forget these requirements just because you are using a nontraditional selection procedure.
6. Insist on strong methodological designs in the research to prove the value. For example, make sure the hypothesis is actually falsifiable, meaning that the lack of value can also be proven by the design. Many of the supposed validation studies are actually pilot tests that only prove that the software works and neither predict meaningful criteria nor show incremental validity beyond existing procedures. As another example, if the vendor uses personality predictors in an attempt to reduce adverse impact, which is extremely common, make sure they are validated in predictive designs (rather than concurrent designs) because validities often do not hold up from the latter to the former due to candidate faking in operational settings.

**Table 1. Continued**

- 
- 
7. Especially consider cross-validation with big data techniques. This is especially critical to big data techniques because many of them seek statistical solutions and models that optimize prediction, and thus they fit to the error in the data and necessitate cross-validation. Note that this applies not only to the weights but also to the selection of the subset of variables included in the model. In the authors' experiences, vendors often recognize the former but seem to be unaware or not bother with the latter, yet that may be the biggest cause for overestimation of validities.
  8. Ensure the consulting firm has adequate I-O psychology talent. I-O talent is necessary to address the validation requirements and other topics where our expertise is necessary, such as the measurement of job performance, embedding the process within the broader selection system in the organization, determining cutting scores and other issues with using the data to make decisions, conducting adverse impact analyses, writing validation reports, and so on. The skill sets of many vendors in the big data space appear to be 50% entrepreneurship, 45% IT, and 5% content (e.g., of personnel selection).
  9. Ensure that the customer organization has adequate independent I-O psychology advisement. If the company does not have an internal I-O psychologist who is familiar with these techniques (or does not have time to learn), consider using an outside reviewer who knows the topic as an independent advisor who works for you and is unassociated with the vendor.

**Legal issues**

10. Watch for inadvertent race or gender scoring by the software. Big data approaches will utilize whatever data it is given to enhance prediction. This also includes markers that reflect race or gender, such as schools, associations, and other information on applications, for example. As such, if race or gender related information improves prediction, it will be included in the model. Another problem is if the computer model is constrained to not show differences by diversity subgroups, it could inadvertently create within-group norming, which was made illegal by the 1991 Civil Rights Act.
11. Beware of value claims that rest too heavily on the lack of adverse impact. Many vendors tout this as one of their key product advantages. However, the lack of any adverse impact may mean the product simply does not measure anything and thus will not offer you any advantages. The tradeoff between validity and adverse impact is unlikely to be easily remedied by big data techniques, given the current state of the science. However, there are some statistical techniques being currently researched that might help marginally (e.g., Song, Wee, & Newman, 2017), and some recent research suggests that perhaps big data techniques might also help avoid the trade-offs by allowing high validity and low impact procedures to be used less expensively (e.g., Campion et al., 2016).

**Hidden costs**

12. Consider applicant reactions to the big data selection procedure. Candidates react negatively to selection procedures that they do not understand. Not knowing what the selection procedure measures violates their perceptions of job relatedness and opportunity to perform. Negative candidate reactions may create hidden costs for your organization in terms of decreased applications or company reputation. For example, the initial implementation of computer scored essays to evaluate reading in educational contexts met with some public resistance (e.g., referring to such scoring as "robo-raters").
- 
-

**Table 1. Continued**

- 
- 
13. Make sure your IT system can handle the product if it is to be integrated. This might reveal incompatibilities or create hidden costs.
  14. Watch for hidden profits. You are only paying for the research and programming. Once developed, it should be very cheap to use.
  15. Determine the “lifecycle” of the proposed product. This includes information regarding substantive changes that will be necessary not only for general maintenance and upkeep purposes but also how often the algorithm will need to be updated and at what cost.
- 
- 

**Table 2. Future Research Directions****Attracting and selecting talent**

- 
- 
1. Can big data techniques be used to mine potential applicant information that is available online to improve targeted recruitment strategies of an organization? Research on targeted recruitment in general is still in its infancy despite organizations engaging in these efforts on a regular basis. With the advent of big data techniques for various TM purposes, it is only a matter of time before organizations begin to consider their use for candidate targeting purposes. With the sheer volume of potential applicants using social media today (e.g., LinkedIn, Facebook, Twitter), a rich corpus of data, both qualitative and quantitative, is available.
  2. Can big data techniques that text mine prior applicant data be used to mitigate the adverse impact dilemma? Initial research has shown that it may be possible to approach the adverse impact-validity dilemma from a less traditional angle—by reducing the costs associated with using more valid but traditionally more costly assessments (Campion et al., 2016). However, big data techniques may provide many alternative ways to approach this problem. As one example, over time, organizations accumulate a large volume of information on applicants from interviews, application information, narrative descriptions of skills, and so on. Big data techniques may enable organizations to take advantage of this information by identifying underlying combinations of applicant characteristics that are generally ignored but offer incremental validity above and beyond current assessments it uses. For example, this would be particularly important if such generally “invisible” characteristics differ across subgroups. To use an example for application data, it may be the case that Blacks and/or Hispanics differ from Whites in terms of the general trends of their human capital investment strategies (e.g., types of jobs held, regions where jobs are sought, greater frequency and duration of holding multiple jobs) and the experiences and specific types of knowledge and skills they accumulate as a result (e.g., foreign languages, social networks they have developed, challenging work experiences that create leadership development opportunities, ability to engage in multiple subcultures). These might, in turn, be more strongly associated with some criteria of interest to the organization such as employees’ abilities to transition among work roles in different locales or its current and future strategic needs.
  3. Are there problems associated with using big data techniques as a surrogate for human interviews? Already, many vendors are offering video platforms for online interviewing. Many still require a human interviewer (e.g., Skype), but some do not and claim to offer
- 
-

**Table 2. Continued**

- 
- validated assessments that incorporate verbal responses, intonation, and nonverbal cues. Such assessments beg many further questions related to the one identified above including: When used as a surrogate for human interviewers, do big data techniques reduce, model, or enhance subconscious biases of human interviewers? If they do, is there a way to create computer algorithms that do not model these biases? To what extent are speech patterns, eye contact, intonation, and other nonverbal cues predictive of job performance (or any other criteria of interest that is relevant to TM)?
4. Can big data techniques used for personnel selection increase in validity over time? Just as Google's search algorithm is updated over time to influence the results listed, it may be possible for computer algorithms providing the basis for operation of big data techniques used for personnel selection to be made to update over time automatically through machine learning. As an example, it may be possible to create an algorithm that continuously updates itself to include data collected on candidates' job performance over time and comparing it to their test scores continuously to maximize validity (e.g., by adjusting weights of selection procedures, by revising cutting scores).
5. How do applicants react to organizations that use big data techniques for assessment? It seems that completely automated recruitment and selection processes are not too far out of reach, where applicants do not speak to an organizational representative during the process but are, instead, hired based on a score earned via those assessments. According to social presence theory, the medium through which interactions take place contributes to the perception of whether or not the receiving end is manned by a "real person" via nonverbal cues such as posture and attire, and so on. Without a sense of a real person, an individual may doubt the authenticity of the interaction and its outcomes. This has critical implications for automated recruitment and selection assessments such that the candidates may perceive low procedural or distributive justice and be less likely to feel as though they understood what was being asked of them due to the lack of social cues.
6. How might candidates use an organization's big data techniques to their advantage? Are there fair and unfair advantages produced by this? Prior research on staffing has characterized the relationship that exists between organizations and applicants as a competitive one involving at least partially misaligned goals, which breeds adaptation behaviors on both sides and an overarching trend of escalation. The use of big data techniques for the purposes of selection, and, more broadly, TM, is a prime example of an adaptation on the part of organizations. This, of course, highlights the concern of applicants counter-adapting by finding ways to "game the system."

#### **Developing and deploying talent**

7. Can big data techniques be used to improve team efficiency and/or productivity? Not only do teams in organizations interact with other teams throughout the organization, but individuals, of course, interact with others within their own team. Big data techniques may be useful toward improving efficiency and productivity at both the team and individual levels because they are generally capable of incorporating both quantitative data (e.g., frequency with which different individuals and teams interact) and qualitative data (e.g., the information communicated among these parties). Gathering this information over time may permit a computer algorithm to be developed that specifies the "best" way to group individuals into teams and teams that work together based on some set of criteria of interest such as by enhancing the external range of teams' social networks or specifying the degree of coordinative and collaborative behaviors necessary within a cross-functional team.
-

**Table 2. Continued**


---

8. Can big data techniques be used to model employee decision making? Prior research in marketing suggests that Bayesian frameworks can be used to understand nuances regarding consumer behavior. Using these frameworks, research has explained counterintuitive (i.e., generally unpredictable) behaviors when it comes to preference formulation and decision making regarding products and brands, such as why consumers often choose products and brands that they do not prefer or that are of lower quality. According to this research, individuals' expectations should be conceptualized as distributions rather than as point estimates, where large distributions are viewed as being associated with greater risk (due to greater uncertainty) and thus lower utility. These posterior probabilities are continuously updated with each decision outcome. Employees make many decisions that could potentially be modeled using this technique as well. For example, when it comes to succession management, research often ignores that employees actually have to accept positions, and there is research suggesting that promotions and developmental assignments are often not accepted. When deciding whether to accept job assignments, developmental or otherwise, they are likely to consider job assignment characteristics when formulating preferences and decisions in the same way they would any other major investment; it may be possible to model these over time to identify the likelihood of employees accepting assignments, thus enhancing the efficiency of the succession management system.

#### **Retaining talent**

9. Can big data techniques be used to reduce burnout and emotional exhaustion among employees? Prior research suggests employee burnout and emotional exhaustion predict voluntary turnover. Big data techniques may be useful toward retaining talent by mitigating these factors. For example, a plethora of information is often available to organizations regarding individuals' daily routines. These data can come from many sources such as online calendars they use, programs that track social media and email use on personal computers, and productivity curves derived from real-time productivity data if the job permits. From this information, algorithms can be developed that specify "micro breaks" for individuals at key times they might need them and when meetings should be scheduled, for example, the goal of which would be to create a personalized and optimized "work rhythm" for individuals.
10. Can big data techniques be used to identify "hidden" predictors of turnover? As noted above, vendors are beginning to consider alternative or "hidden" predictors of job performance in interviewing for personnel selection. It will not be long before they also begin to search for previously unnoticed predictors of turnover in organizations—if they have not already. Thus, it is important for future research to focus on this important topic as it relates to the use of big data techniques in organizations. Although research has identified a wide range of predictors of turnover, a major advantage of big data techniques is that the software they employ makes developing complex models of data much easier for those who are not "data scientists." These models can be "taught" to incorporate a greater amount, and a wider range, of types of data and variables. As such, they render the search for these hidden predictors easier and more available to a wider range of researchers of all skill levels when it comes to data analysis. Further, they make survival analysis available as an analytical strategy for understanding the relationships among a vast number of variables and expected durations of employment for individuals over time.
-



## References

- Campion, M. C., & Campion, E. D. (2018, April). Using text mining to identify and quantify strategically aligned applicant brands. Paper presented at the 33rd Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.
- Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology, 101*, 958–975.
- Rotolo, C. T., Church, A. H., Adler, S., Smither, J. W., Colquitt, A. L., Shull, A. C., ... Foster, G. (2018). Putting an end to bad talent management: A call to action for the field of I-O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 11*(2), 176–219.
- Song, Q., Wee, S., & Newman, D. A. (2017). Diversity shrinkage: Cross-validating pareto-optimal weights to enhance diversity via hiring practices. *Journal of Applied Psychology, 102*, 1636–1657.
- Uniform guidelines on employee selection procedures. (1978). *Federal Register, 43*, 38290–38315.

# I-O Psychologists Can Help Make Sure Your HiPos Aren't NoPos

Lisa M. Finkelstein  
Northern Illinois University

David P. Costanza  
The George Washington University

Gerald F. Goodwin  
US Army Research Institute for the Behavioral and Social Sciences

We agree with Rotolo et al.'s (2018) assertion that talent management is a space where the academic–practice gap in industrial and organizational (I-O) psychology is quite cavernous and where the vulnerabilities to anti-I-O (AIO) are high. As researchers who began a journey a few years ago to explore the high potential (HiPo) identification process from the science perspective (largely inspired by Silzer & Church, 2009), we echo the frustration that the current focal authors express with the science-side lag in this area. For us, what started as a question from a senior officer in the Army turned into the development of a theoretical model and the start of multilocation research lab designed to further the understanding and success of the HiPo identification process. Our objective is to share a bit of our journey that got us to this point and some lessons for others inspired by this focal article to become anti-anti-I-O (AAIO) warriors.

Lisa M. Finkelstein, Northern Illinois University; David P. Costanza, The George Washington University; Gerald F. Goodwin, US Army Research Institute for the Behavioral and Social Sciences.

Correspondence concerning this article should be addressed to Lisa M. Finkelstein, Northern Illinois University, DeKalb, Illinois 60115–2828. E-mail: [lisaf@niu.edu](mailto:lisaf@niu.edu)