The International Journal of Human Resource Management

ISSN: 0958-5192 (Print) 1466-4399 (Online) Journal homepage: www.tandfonline.com/journals/rijh20

HRM's financial value from obtaining more star performers

Harry Joo, Herman Aguinis, Joowon Lee, Hannah Kremer & Isabel Villamor

To cite this article: Harry Joo, Herman Aguinis, Joowon Lee, Hannah Kremer & Isabel Villamor (2022) HRM's financial value from obtaining more star performers, The International Journal of Human Resource Management, 33:21, 4179-4214, DOI: 10.1080/09585192.2021.1948890

To link to this article: https://doi.org/10.1080/09585192.2021.1948890

THE INTERNATIONAL JOURNAL OF HUMAN RESOURCE MANAGEMENT
2022, VOL. 33, NO. 21, 4179–4214
https://doi.org/10.1080/09585192.2021.1948890

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

HRM's financial value from obtaining more star performers

Harry Joo[a,#], Herman Aguinis[b,#], Joowon Lee[b], Hannah Kremer[b] and Isabel Villamor[b]

[a]University of Dayton, Dayton, OH, USA; [b]The George Washington University, Washington, USA

ABSTRACT

We assessed the financial value of human resource management (HRM) as a function of obtaining more star performers. Specifically, we implemented utility analysis procedures on 206 samples of individual performance (i.e. output) encompassing 824,924 workers. We found that HRM adds greater financial value by obtaining more stars. Our results also offer several specific contributions to HRM theory. First, regarding how HRM produces greater value by obtaining more stars, our evidence points to a nonlinear model of HRM's value, where HRM generates significant yet diminishing returns by increasingly obtaining the most productive ones. Second, regarding when, our results show that diminishing returns from HRM are stronger when output differences among top stars are relatively small. Third, regarding why, our study explains that small output differences among top stars may create various costs which diminish the returns from obtaining the most productive stars. Our explanation of HRM's nonlinear pattern contributes to the star literature by helping integrate a variety of specific explanations for stars' curvilinear influence discussed in past research. Regarding HRM practices, we highlight the need to use utility analysis procedures that more fully consider the existence of stars.

KEYWORDS

Star performers; individual performance; normality assumption; utility analysis

Star performers are workers who produce disproportionately larger amounts of cumulative output compared to their peers (Asgari et al., 2021; Kehoe et al., 2018; Morris et al., 2021; O'Boyle & Aguinis, 2012; Taylor & Bendickson, 2021). A star's output can be a dozen or more standard deviations above the average (e.g. Veksler, 2010). Stars include workers who manage to accumulate disproportionate amounts of output even if there are constraints that prevent many other workers from producing much output, creating a large gap between stars' and non-stars' output (Aguinis et al., 2016). In terms of prevalence, stars exist in a

CONTACT Herman Aguinis haguinis@gwu.edu The George Washington University, Washington, USA
Supplemental data for this article is available online at https://doi.org/10.1080/09585192.2021.1948890.
[#]The first and second authors contributed equally.

wide variety of occupations such as salespeople, engineers, managers and laborers (Aguinis et al., 2018; Crawford et al., 2015; Turetsky, 2017). For example, one study found that 82.53% of 229 samples of worker output had significantly heavy right-tails containing stars, as depicted in Figure 1 (Joo et al., 2017). Not only are stars prevalent, but they also seem to be more productive than ever before because some workers are able to generate much larger amounts of output from technological advances than others (Cascio & Montealegre, 2016), further widening the output gap between stars and non-stars. In particular, the Internet now allows workers to access valuable information with greater ease, communicate with clients and coworkers on a more frequent basis, and work longer hours with superior efficiency. As a result, some workers (e.g. those with greater motivation, knowledge, skills and abilities) are more likely to produce vastly superior output compared to others (Van Iddekinge et al., 2021).

However, to date, human resource management (HRM) research conducted outside the star literature has not fully accounted for the prevalence of increasingly productive stars, providing a potentially inaccurate picture of HRM's financial contributions. This is because most HRM research assumes that individual output is normally-distributed, effectively denying the presence of stars who create heavy right-tails in individual output distributions (see Figure 1). Reflecting the central role of the normality assumption in HRM, consider Pearson's correlation coefficient. Despite the widespread use of Pearson's correlation, significant non-normality (i.e. created by star performers) must not exist to ensure accurate correlation values (Bishara & Hittner, 2012; Kowalski, 1972). If there are major departures from normality, alternative approaches can be used to derive results that are 'robust' (i.e. insensitive) to non-normality. Yet, so-called robust approaches work by artificially reducing or eliminating the influence of outliers (i.e. stars) and, therefore, still rely on the normality-based assumption that stars are not much different from non-stars or that stars do not exist. Given the normality assumption that partially or completely ignores the presence of stars, there remains the likely misleading premise that stars largely do not exist and thus HRM's financial value may not change by obtaining more stars.

The utility analysis literature in HRM has also adopted the normality assumption, largely denying the presence of stars and implying that HRM's value does not depend on stars. Utility analysis assesses the degree to which HRM can contribute to firms financially (Ock & Oswald, 2018; Oprea et al., 2019; Seijts et al., 2020). Illustrating its continued relevance, HRM researchers refer to utility analysis as an important way to translate the value of HRM practices into financial terms (e.g. Albrecht & Marty, 2020; Chapter 14 in Cascio & Aguinis, 2019). Nonetheless,

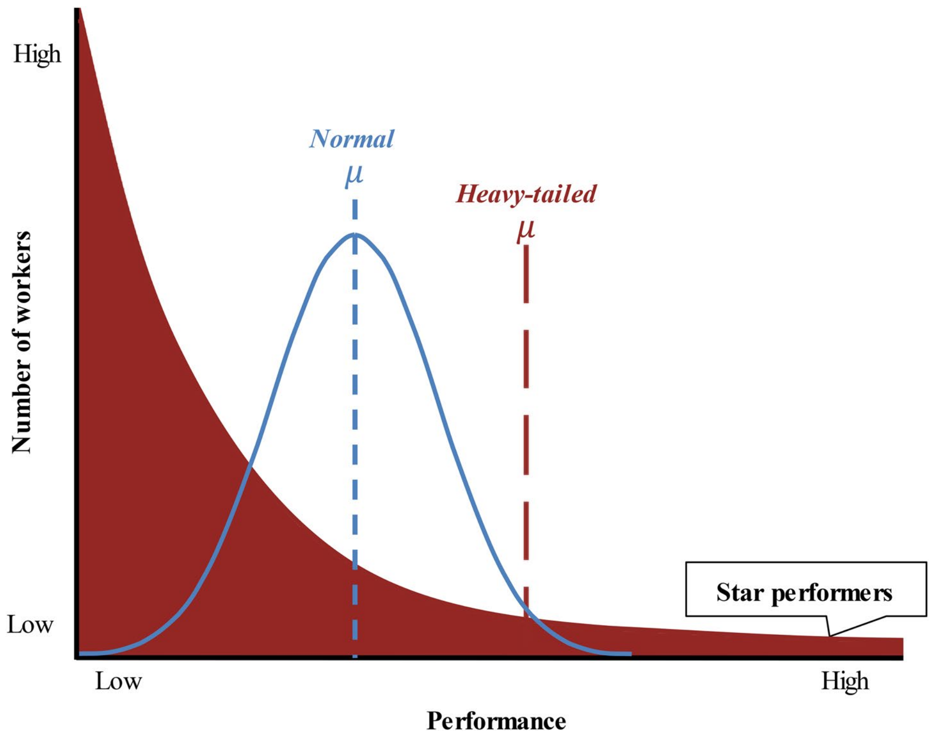

Figure 1. Normal distribution overlaying a heavily right-tailed distribution. *Note.* μ = mean value for each distribution. The normal distribution assumes that most scores cluster around the mean and fan out into short and symmetrical tails. The heavy-tailed distribution includes a larger proportion of extreme scores (i.e. star performers) and the majority of scores fall below the mean.

consistent with the broader HRM literature, utility analysis studies often explicitly assume that individual output distributions follow normality (Cascio & Ramos, 1986; Sackett & Yang, 2000; Schmidt et al., 1979) or, alternatively, state that estimates of HRM's financial value are essentially equivalent regardless of departures from normality (Anderson & Muchinsky, 1991; Burke & Frederick, 1984). In short, the HRM field, and the utility analysis literature specifically, currently offers the likely misleading premise that the financial value of HRM does not vary as a function of obtaining stars.

The present study

Our goal is to examine the financial value that HRM can offer by obtaining more stars. Given the need to operationalize HRM more specifically for empirical purposes, we evaluate the HRM practice of employee selection – that is, how HRM's financial value may vary from acquiring more stars. So, although our operationalization choice refers to employee selection, our results and implications broadly apply to other HRM domains including training (Carretero-Gómez & Cabrera, 2012), compensation (Sturman et al., 2003) and performance appraisal and feedback (Landy et al., 1982). A key feature of the utility analysis procedures we use is that they incorporate the existence of stars by varying degrees. So, if a utility analysis procedure that more fully incorporates the presence of stars produces higher financial valuations than another procedure which does less of that, the difference indicates financial gain (i.e. additional financial value generated from HRM obtaining more stars). Thus, we ask:

> *Research Question*: How does the financial value of HRM vary as a function of obtaining more star performers?

To clarify, we examine value created, not captured, from HRM. As noted in prior research, some stars may be better equipped than others for capturing value created (Kehoe et al., 2018). So, our study focuses on maximum possible value creation from obtaining stars, before the value is subsequently captured. We also clarify that our study consists of financial valuations of HRM resulting from more stars, instead of subjective or psychological evaluations.

Another aspect of our study is that we use abductive (i.e. discovery-oriented) reasoning (i.e. studying specific observations to find a plausible explanation without having to posit hypotheses in advance) (Behfar & Okhuysen, 2018; Saetre & Van de Ven, 2021), in contrast to the commonly-used deductive reasoning (i.e. using theory to generate specific predictions or hypotheses) (Cortina et al., 2017). We adopt abductive reasoning because, regarding our research question of how HRM's value may vary by obtaining more stars, there are largely two competing theoretical views derived from the star literature (Asgari et al., 2021): HRM produces less versus greater financial value as a function of obtaining more stars. In turn, 'when there is a need to explain inconsistencies [as is the case in our study], abductive reasoning is especially useful' (Behfar & Okhuysen, 2018, p. 328).

Based on 206 samples of individual performance (i.e. operationalized as individual output) encompassing 824,924 workers, our results indicated that HRM adds greater financial value to firms by obtaining more stars. Specifically, one of the utility analysis procedures we used – the observed distribution procedure – fully acknowledges the existence of stars by incorporating actual distributions of individual output to estimate HRM's financial value. In contrast, other extant utility analysis procedures we used assume away the presence of stars by some degree by, for example, ignoring the top 15% performers (per the global procedure), top 3% (modified global procedure) or top 1% (further modified procedure). We found that the observed distribution procedure in general yields higher financial valuations, up to nine times greater than the other utility analysis procedures which assume away stars. So, our study quantifies the extent to which firms benefit financially from obtaining more stars. In other words, by considering actual distributions of individual performance, our overall empirical contribution is to demonstrate that HRM resulting in more stars produces greater financial value than previously believed.

In addition to our overall empirical contribution, our results provide several specific theoretical contributions to HRM by helping clarify how, when and why HRM produces greater financial value as a function of obtaining more stars. First, given our conclusion that HRM obtaining stars adds greater value to firms, our evidence further indicates that increasing focus on obtaining the most productive stars often results in significant yet diminishing returns. Hence, we offer a nonlinear rather than a linear understanding of HRM's financial value, addressing not only whether but also *how* HRM generates greater value as a function of obtaining more stars. Second, we found that the diminishing pattern is more apparent when output differences among top stars are relatively small. As a result, we contribute to a better understanding of *when* obtaining stars results in more strongly diminishing returns. Third, in light of our results that diminishing returns are stronger when there are small output differences among top stars, we elaborate on past research explaining why such small differences may lead to costs which diminish the returns from obtaining the most productive stars. We thus provide an explanation for *why* obtaining more stars may generate significant yet diminishing returns. Next, we elaborate on studies in the star literature to offer two competing theoretical views addressing how HRM's value might vary as a function of obtaining more stars.

Theoretical background

HRM generates less financial value by obtaining more stars

According to one view, HRM generates less value by obtaining more stars because stars may have a largely harmful impact. Stars can contribute to negative stock-price movements (Groysberg & Lee, 2009), inhibit non-stars' learning (Li et al., 2020), reduce group effectiveness (Overbeck et al., 2005) and produce other costly outcomes (Chen & Garg, 2018; Kehoe et al., 2018; Prato & Ferraro, 2018). As an illustration, in biotech firms, some star scientists (e.g. who possess narrow expertise and are less collaborative) tend to be interested more in protecting their knowledge rather than expanding it. Such star scientists are potentially harmful given that they can inhibit non-star scientists from initiating and leading innovation (Kehoe & Tzabbar, 2015). Even in contexts where stars are mostly beneficial, relying on stars may make a firm more vulnerable to disruptions caused by their departure. The negative relationship between star departure and firm performance, for example, tends to be stronger among firms that rely on their existing stars so much to the extent of investing less in HRM (e.g. selection, training) (Kwon & Rupp, 2013). Star departure may also decrease coworkers' output quality (Oettl, 2012) and disrupt a firm's exploitation routines (Tzabbar & Kehoe, 2014), especially if the stars' colleagues heavily relied on the stars' helping and collaborative behaviors.

The perspective that HRM generates less value from obtaining more stars is further bolstered by studies indicating a possible inverted-U relationship between stars and outcomes (e.g. Call et al., 2020; Groysberg et al., 2011; Swaab et al., 2014). This research suggests that greater proportion of stars contributes positively to outcomes such as group performance or non-stars' individual performance, though the positive trend is followed by diminishing and then negative returns. One explanation for an inverted-U pattern is that too many stars can decrease the quality of coordination among members in a group (Groysberg et al., 2011). For example, to the degree that stars have large egos that lead to interpersonal clashes and less collaboration, greater proportions of stars in a group may aggravate those interpersonal dynamics. Deterioration in intragroup coordination, in turn, will be more harmful in reciprocal interdependence contexts (Swaab et al., 2014), where workers (e.g. basketball players) not only complete individual tasks (e.g. dribbling) but also create intermediate output (e.g. passes) that becomes the input that other members in the group need to produce final output (e.g. points scored) (Harder, 1992; Thompson, 1967). Thus, there will often be an inverted-U relationship between the proportion of stars and outcomes, suggesting that HRM may yield initially greater but subsequently less and then negative financial value from obtaining more stars.

HRM creates greater financial value by obtaining more stars

A competing perspective suggests that HRM resulting in more stars adds greater value to firms, to the degree that stars have mainly beneficial effects. Stars can contribute to firm revenue (Han & Ravid, 2020), improve the odds of firm survival (Bedeian & Armenakis, 1998), facilitate new product development (Zucker & Darby, 1996), motivate peers to perform better (Ammann et al., 2016) and achieve other desirable criteria (Bendapudi & Leone, 2001, Liu, 2014) – which are more likely if stars are supported by colleagues and firms (Amankwah-Amoah et al., 2017; Groysberg & Lee, 2008). Thus, stars may add value directly *via* exceptional output or indirectly by providing their firms with access to external resources and exerting significant influence on colleagues (Grigoriou & Rothaermel, 2014; Kehoe et al., 2018; O'Boyle & Aguinis, 2012). The greater the number and variety of ways in which a star contributes to value, the more sustainable is the star's value creation for the firm (Kehoe et al., 2018). The strategic HRM literature suggests that stars' prolific output is often due to the valuable, rare, inimitable and non-substitutable (VRIN) nature of their individual characteristics including knowledge, skills and abilities (Aguinis & O'Boyle, 2014). The VRIN nature of stars, then, can play a vital role in creating and maintaining the competitive advantage of firms in local and international contexts (Minbaeva & Collings, 2013).

Moreover, past research showing negative returns from greater proportions of stars can be explained by a methodological artifact (Gula et al., 2021). To detect the presence of negative returns accurately in correlational studies, it is necessary to use certain regression techniques (e.g. interrupted-regression) or compare the fit of inverted-U shapes with that of other nonlinear shapes (e.g. r-shape described by the logarithmic growth function) (Hansen, 2000; Marsh & Cormier, 2001, Simonsohn, 2018; Vaci et al., 2019). Yet, a common way to test for nonlinear relationships is to estimate a quadratic function by entering a predictor and a squared term of the predictor. The combination of a significant positive linear coefficient for the predictor, and a significant negative coefficient for the squared term, is usually interpreted as evidence of an inverted-U relationship. However, as demonstrated by Gula et al. (2021), such interpretation is often inappropriate because the squared term *a priori* assumes that there are both diminishing and negative returns, forcing the fitted regression line to bend downwards. As a result, there may be (very) little or no data points beyond the estimated inflection point (after which negative returns presumably occur). To the extent that this artifact explains negative returns found in the star literature, HRM may produce greater (though still possibly diminishing) financial returns by obtaining more stars.

In light of these two competing views, our goal is to better understand HRM's financial value as a function of obtaining more stars. We also sought to clarify possible diminishing or negative returns from HRM resulting in more star performers, as well as when and why, thereby integrating existing theoretical perspectives.

Method

Samples

We collected 206 samples of individual output encompassing 824,924 workers, where each case per sample refers to a distinct worker's output accumulated within a certain period of time. Given that past research defined individual performance as behaviors (i.e. actions taken to produce results), output (i.e. results produced), or a mix of both (Joo et al., 2017), our use of individual output is consistent with how performance has been defined in the utility analysis literature for the past 40 years (Chapter 14 in Cascio & Aguinis, 2019; Schmidt et al., 1979; Seijts et al., 2020). Moreover, our approach is theoretically appropriate for estimating HRM's value as a function of obtaining more stars because it considers workers who manage to accumulate disproportionate amounts of output, even in the presence of ceiling constraints that prevent many other workers from producing greater output (Aguinis et al., 2016). So, constraints on individual output likely serve as a force that creates a larger gap between stars' and non-stars' output, instead of acting to prevent stars from emerging. To highlight the representativeness of our samples, we list and describe the variety of typical occupations from which we collected our data in Table 1 (e.g. manufacturing, entertainment, high-tech, sports, banking, research, sales, customer service, agriculture and medicine). The online supplement provides more details about each of our 206 samples (Appendix A) as well as the source of every sample (Appendix B).

Per sample, we broadly define stars as workers who produce disproportionately larger amounts of cumulative output compared to their peers. As illustrated in Figure 1, this broad definition means that stars exist in heavy right-tails in skewed distributions of individual output. A star does not have to be the #1 performer in a group, firm or industry, and we did not limit ourselves to a single cutoff regarding how much output a worker must produce to be a star. Our broad definition of stars as those producing disproportionate contributions is consistent with studies that have similarly adopted a broad stance rather than specifying a particular cutoff. Almost identical to our definition, O'Boyle and Kroska (2017) noted that 'a star exhibits exceptionally high quality and/or exceptionally large quantity of output relative to his or her peers' (p. 43). Similarly, according to Cappelli and Keller (2017), 'a defining characteristic of a star is that he or she contributes a disproportional amount of output relative to his or her peers' (p. 26). In short, our broad conceptualization subsumes more specific definitions of stars based on varying cutoffs (e.g. top 1% or 3% of performers), allowing us to derive results that are comparable to previous studies that defined and measured stars using different cutoffs.

Overview of six procedures for estimating the financial value of HRM

To estimate HRM's financial value, we used six utility analysis procedures in each of the 206 samples. Among our six procedures, we first implemented the following four extant procedures based on prior research:

Table 1. Description of samples representing typical occupations used in the present study.

Sample #	Occupation	*N*	Worker output measure and comments
164	Bank tellers	75	Sales in month 1
165	Bank tellers	75	Sales in month 2
166	Retail sales associates	244	Sales over 1-month period
167	Call center employees	219	Hourly revenue over 3-month period
168	Call center employees	219	Hourly calls over 3-month period
169	Call center employees	86	Total revenue over 3-month period
170	Fundraising callers	57	Number of calls over 2-week period
171	Fundraising callers	101	Revenue over 2-week period
172	Fundraising callers	80	Calls per hour over 2-week period
173	Call center employees	71	Number of sales over 7-week period
174	Call center employees	71	Revenue over 7-week period
175	Paper sorters	18	Pounds sorted per hour over 2-year period
176	Pelt pullers	13	Number of pelts pulled
177	Toll-ticket sorters	13	Number of tickets sorted
178	Typists	43	Words typed per minute, adjusted for errors
179	Card punch operators	62	Average number of cards punched per hour
180	Lamp shade sewers	18	Number sewn
181	Lamp shade sewers	19	Number sewn
182	Card punch operators	113	Average number of cards punched per hour
183	Card punch operators	121	
184	Electrical fixture assemblers	40	Number assembled
185	Lawyers	1173	Number of new clients obtained in a year
186	Lawyers	417	Revenue generated in a year
187	Lawyers	1074	Number of new clients obtained in a year
188	Lawyers	693	Revenue generated in a year
189	Lawyers	1945	Number of distinct legal matters processed in past three months
190	Lawyers	717	Number of new legal matters obtained in a year
191	Lawyers	860	Estimated total value of one's portfolio of clients in a year
192	Managers	209	Number of people supervised
193	Managers	84	
194	Agricultural workers	142	Kilograms of fruit picked per hour
195	Agricultural workers	142	
196	Transcribers	15	Keystrokes per hour
197	Produce packers	17	Number of boxes scanned per hour
198	Recruiters	268	Ratio of actual productivity to expected productivity
199	Financial advisors	183	Number of appointments per day
200	Agricultural workers	3960	Number of trees planted per day
201	Outpatient care doctors	167	Number of patients examined per day in a year
202	Inpatient care doctors	130	Total number of patients admitted in a year
203	Inpatient care doctors	130	Total number of days that patients stayed in a year
204	Agricultural workers	377	Meters of rice planted per 10 min
205	Laundry workers	273	Number of garments processed per day
206	Programmers	20	Lines of code per hour, excluding blank or comment lines

Note. Given that the table here shows only 43 of our 206 samples (i.e. samples #164–206), Appendix A in the Online Supplement includes a description of each of the 206 samples used in the present study.

(1) 40% of mean salary, (2) 70% of mean salary, (3) global and (4) modified global (Burke & Frederick, 1986; Cabrera & Raju, 2001; Carretero-Gómez & Cabrera, 2012; Hazer & Highhouse, 1997). As our fifth procedure, the further modified procedure more fully considers the presence of stars compared to the other four procedures just mentioned above. Finally, for our sixth procedure, the observed distribution procedure most fully incorporates the presence of stars by examining actual distributions of individual output (while not assuming away any stars – which the other five procedures do by at least a small amount).

We implemented all six procedures quantitatively using output-based performance data rather than qualitatively using subjective assessments provided by judges, as 'the overwhelming reliance on subjective estimates' has limited the impact of past research (Becker & Huselid, 1992, p. 233). In other words, we conducted financial (not psychological) valuations of HRM resulting in more stars. Furthermore, each of the six procedures is based on the Brogden–Cronbach–Gleser (BCG) model (Brogden, 1949; Cronbach & Gleser, 1965; Sturman, 2000). The BCG model is the most widely used and influential among prior studies assessing the financial value of HRM (Cabrera & Raju, 2001). The BCG model, in the context of acquiring new workers, is described by Equation (1):

$$\Delta\mu(\$) = N_s SD_y r \frac{\phi}{p} - NC, \tag{1}$$

where $\Delta\mu$ ($) = financial value of using a valid predictor of worker performance to systematically acquire new workers; N_s = number of applicants acquired; SD_y = standard deviation of worker performance in dollars; r = validity coefficient for the focal predictor of worker performance; p = proportion of applicants acquired (i.e. acquisition ratio) while assuming top-down acquisition; ϕ = ordinate (i.e. height) of a normal curve associated with acquisition ratio p; N = number of applicants and C = average cost for assessing each applicant.

Detailed description of the six procedures for estimating HRM's financial value, and their differences

Each of the six procedures not only calculates HRM's financial value per sample, but also incorporates the presence of stars in varying degrees by differing on how the SD_y parameter in Equation (1) is calculated. First, 40% of mean salary most heavily ignores stars because it takes a small fraction of salary, as opposed to using a wide range of percentiles on a performance distribution. Second, 70% of mean salary ignores stars to a lesser degree by taking a larger fraction of salary (i.e. 70% instead of 40%). Third, the global procedure also ignores stars, but by a lesser degree compared to the 40% and 70% of mean salary procedures. This is because the global procedure, rather than simply taking a fraction of salary, uses a broad range of percentiles on a performance distribution including up to the 85th percentile in terms of monetary value of worker performance (while ignoring the top 15% performers). Fourth, the modified global procedure ignores the presence of stars less heavily than the global procedure by using a wider range of percentiles, all the way up to the 97th percentile (ignoring the top 3%). Fifth, the further modified procedure considers an even wider range of percentiles by including up to the 99th percentile (ignoring just the top 1%), thus incorporating stars more heavily compared to the modified global procedure. Sixth, the observed distribution procedure most fully considers stars (i.e. does not ignore any stars that exist) because the procedure calculates the standard deviation across all cases in the performance distribution. Table 2 provides the calculations involved in the implementation of each of the six procedures.

Full implementation of the six procedures requires not just calculation of the SD_y parameter, but also information on the parameters on the right-hand side of Equation (1). For each right-hand side parameter, we used a range of values to ensure generalizable results. Table 3 shows the values used for each right-hand side parameter and explains their rationale. Based on these values, we created nine distinct combinations of parameter values that we used to implement each of the six procedures.[1]

Calculation of financial gain from obtaining more stars

After using each of the six procedures to estimate HRM's financial value in each sample, we also assessed HRM's financial value as a function of obtaining more stars. We derived financial gain per sample by comparing the following set of estimated financial values: (1) financial value of HRM estimated from the observed distribution procedure (which incorporates the existence of stars fully) versus (2) financial value of HRM estimated from each of the five other procedures (which assume away the presence of stars by at least a small amount). As a result, we calculated five financial gains per sample. These five financial gains allowed us to quantify the extent to which the procedure that fully incorporates stars (i.e. observed distribution procedure) results in greater financial valuations compared to the other five procedures that ignore stars by varying degrees. This way, financial gain indicates the extent to which HRM, by obtaining more stars, offers greater financial value, and how much.

Each financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). So, a positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A

Table 2. Description of six procedures used for estimating the financial value of HRM.

1. 40% of mean salary

We obtained mean salary (as of 2016) for the focal occupation by first consulting bls.gov (the Bureau of Labor Statistics' website). If the mean salary for the occupation was not available from bls.gov, we used the median salary (as of 2016). If neither the mean nor the median salary was available from bls.gov, we obtained the mean or median salary (as of 2016) for the focal occupation from other websites (e.g. forbes.com). Next, we used the equation below:

0.40 * mean salary = SD_y (i.e. standard deviation of worker performance in dollars) based on 40% of mean salary

2. 70% of mean salary

We obtained mean salary (as of 2016) for the focal occupation by first consulting bls.gov. If needed, we instead obtained the median salary (as of 2016) from bls.gov or the mean/median salary (as of 2016) from other websites (e.g. forbes.com). Next, we used the equation below:

0.70 * mean salary = SD_y (i.e. standard deviation of worker performance in dollars) based on 70% of mean salary

3. Global procedure

We first calculated the difference between the 85th percentile and 50th; difference between the 50th and 15th; and difference between the 15th and minimum (where each percentile refers to a monetary value of worker performance). We then took the average of the three percentile differences. The resulting value was standard deviation in terms of worker output, or SD_o, as shown in the equation below:

{(85th percentile − 50th) + (50th − 15th) + (15th − minimum)}/3 = SD_o

Next, to convert the SD_o value into monetary terms, we borrowed from prior research suggesting that the ratio of SD_y (i.e. standard deviation of worker performance in dollars) based on the global procedure to SD_y based on the 40% of mean salary procedure is about 2.75, or 35,192 divided by 12,789 (Burke & Frederick, 1986). So, we multiplied 2.75 with SD_y calculated using the 40% of mean salary procedure. The resulting value was SD_y based on the global procedure, as shown in the equation below:

2.75 × (0.40 × mean salary) = SD_y based on the global procedure

4. Modified global procedure

We first calculated the difference between the 97th percentile and 85th; difference between the 85th and 50th; difference between the 50th and 15th; and difference between the 15th and minimum (where each percentile refers to a monetary value of worker performance). We then took the average of the four percentile differences. The resulting value was standard deviation in terms of worker output, or SD_o, as shown below:

{(97th percentile − 85th) + (85th − 50th) + (50th − 15th) + (15th − minimum)}/4 = SD_o

Next, to convert the SD_o value into monetary terms, we calculated the ratio of SD_o based on the modified global procedure to SD_o based on the global procedure. Then, we multiplied the resulting ratio with SD_y (i.e. standard deviation of worker performance in dollars) based on the global procedure. The resulting value was SD_y based on the modified global procedure, as shown in the equation below:

(SD_o based on modified global / SD_o based on global) × SD_y based on global = SD_y based on the modified global procedure

5. Further modified procedure

We added the difference between the 99th percentile and 97th, to the four percentile differences considered by the modified global procedure. We then took the average of the five percentile differences. The resulting value was standard deviation in terms of worker output, or SD_o, as shown in the equation below:

{(99th percentile − 97th) + (97th − 85th) + (85th − 50th) + (50th − 15th) + (15th − minimum)}/5 = SD_o

Next, to convert the SD_o value into monetary terms, we calculated the ratio of SD_o based on the further modified procedure to SD_o based on the modified global procedure. Then, we multiplied the resulting ratio with SD_y (i.e. standard deviation of worker performance in dollars) based on the modified global procedure. The resulting value was SD_y based on the further modified procedure, as shown in the equation below:

(SD_o based on further modified/SD_o based on modified global) × SD_y based on modified global = SD_y based on the further modified procedure

6. Observed distribution procedure

We first calculated the standard deviation across all cases of worker output (i.e. SD_o based on the observed distribution procedure). Next, to convert the SD_o value into monetary terms, we calculated the ratio of SD_o based on the observed distribution procedure to SD_o based on the global procedure. We then multiplied the resulting ratio with the SD_y (i.e. standard deviation of worker performance in dollars) based on the global procedure. The resulting value was SD_y based on the observed distribution procedure, as shown in the equation below:

(SD_o based on observed distribution procedure/SD_o based on global) × SD_y based on global = SD_y based on observed distribution procedure

Note. We incorporated the presence of the lowest performers (i.e. difference between the 15th percentile and minimum) across the global, modified global and further modified procedures (i.e. procedures #3–5). By considering the same range of lowest performers across these procedures, we were able to accurately isolate HRM's estimated value as a function of obtaining more stars, operationalized as workers who are above certain high percentiles (e.g. 99th percentile, or the top 1%). The presence of the lowest performers is already incorporated in the observed distribution procedure because it calculates the standard deviation among *all* cases of worker output.

Table 3. Values used for each parameter on right-hand side of Equation 1 and their rationale (except for SD_y).

Parameter	Description of values used and rationale for each value
r = validity coefficient for the focal predictor	r = 0.30, 0.40 and 0.50. The three values match the range of meta-analytic validity coefficients for some of the most widely used predictors (e.g. conscientiousness, employment interview and general mental ability) predicting criteria such as supervisory ratings and objective measures of performance as reported in Hunter (1986), Hurtz and Donovan (2000), Marchese and Muchinsky (1993), McDaniel et al. (1994), and Schmidt and Hunter (1998).
C = average cost for assessing each applicant	C = \$75.26, \$752.63 and \$1505.25. The three values were obtained and adapted from Hoffman and Thornton (1997), where the authors used \$50, \$500, or \$1000 for C. The \$50 value was used for the cost of administering an aptitude test per applicant. The \$500 and \$1000 values were used as low and high estimates for the cost of using an assessment center per applicant. To update the three values from 1997 to 2016 dollars, we averaged the results from three websites: dollartimes.com/inflation/inflation.php; usinflationcalculator.com; and calculator.net/inflation-calculator.html. The updated values were \$75.26, \$752.63 and \$1505.25.
p = proportion of applicants acquired (i.e. acquisition ratio) while assuming top-down acquisition	p = 0.50, 0.32 and 0.15. As shown in Appendix A of Boudreau (1991), the three ratios closely match some of the most commonly adopted acquisition ratios in prior studies. The acquisition ratio of 0.50 has been used for food and beverage sales managers (Cascio & Silbey, 1979), U.S. government computer programmers (Schmidt et al., 1979) and power plant operators (Dunnette et al., 1982). The acquisition ratio of about 0.32 has been used for nursing assistants (Schmidt & Hoffmann, 1973), convenience store managers (Weekley et al., 1985), telephone company office managers (Cascio & Ramos, 1986) and clerical employees (Cronshaw et al., 1987). The acquisition ratio of 0.15 has been used for welders, mechanics, technicians and machine operators (Wroten, 1984).
ϕ = ordinate (i.e. height) of a normal curve associated with acquisition ratio p	ϕ = 0.3989, 0.3572 and 0.2323. The three values correspond to our use of 0.50, 0.32 and 0.15 for acquisition ratio p, respectively. To identify the ordinate associated with a specific acquisition ratio, we used the Naylor–Shine table (Naylor & Shine, 1965), which we accessed from Appendix B in Cascio and Boudreau (2011).
N = number of applicants	N = 100, 20 and 7. N = 100 was based on Cascio and Silbey (1979), representing large-scale applicant pools that characterize many acquisition situations. But because many other acquisition situations involve small applicant pools, such as fewer than 20 applicants or even fewer than 10 applicants (Scullen & Meyer, 2014), we also used the values of 20 and 7 for N.
N_s = number of applicants acquired	N_s = 50, 32, 15, 3 and 1. These values follow from the pair of acquisition ratio p and applicant pool size N used. For example, if the acquisition ratio p is 0.15 and the applicant pool size N is 100, then N_s must be 15.

Note. SD_y = standard deviation of worker performance in dollars. How we calculated values for SD_y and their rationale is described in the text.

negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, obtaining more stars diminished value.

Next, per comparison between the observed distribution procedure and each other procedure, we averaged the financial gains calculated across our samples. For example, a multiple of 2 means that the observed distribution procedure led to a financial value that is twice the size of the financial value from another procedure which ignores stars by a certain amount. In contrast, a multiple of −2 means that the observed distribution procedure led to a financial value that is half the size of the financial value estimated from another procedure. The average of the two multiples is zero, indicating that obtaining more stars overall neither added to nor diminished value. Table 4 includes additional details on how we calculated financial gain from obtaining more stars.

Results

Overall empirical conclusion

Table 5 shows estimates of financial gains from obtaining more stars – averaged across our 206 samples except three samples constituting

Table 4. Procedures for calculating financial gain from HRM obtaining more stars, expressed in terms of multiples.

Overview: We calculated financial gain in terms of multiples by comparing two financial values against each other: (1) the financial value of HRM derived from the observed distribution procedure that fully considers stars versus (2) the financial value of HRM derived from each of the five procedures ignoring stars by varying degrees (i.e. 40% of mean salary, 70% of mean salary, global, modified global, or further modified). Because the observed distribution procedure is compared to each of the five other procedures, a total of five comparisons were conducted. Per comparison, we used one of the formulas below depending on the result derived from using the procedure ignoring stars by a certain amount, as explained below.

Condition #1: If the observed distribution procedure produced a larger positive financial value than the positive financial value from the procedure ignoring stars, we used the formula shown on the right-hand side to calculate financial gain in terms of multiples:	*Formula for condition #1:* = Financial value derived from the observed distribution procedure/Financial value derived from the procedure ignoring stars
Condition #2: If the observed distribution procedure produced a smaller positive financial value than the positive financial value from the procedure ignoring stars, we used the formula shown on the right-hand side to calculate financial gain in terms of multiples:	Formula for condition #2: = (Financial value derived from using the procedure ignoring stars/Financial value derived from the observed distribution procedure) × −1
Condition #3: If the observed distribution procedure produced a smaller negative financial value than the negative financial value from the procedure ignoring stars, we used the formula shown on the right-hand side to calculate financial gain in terms of multiples:	*Formula for condition #3:* = (Financial value derived from using the procedure ignoring stars/Financial value derived from the observed distribution procedure)
Condition #4: If the observed distribution procedure produced a negative financial value, whereas the procedure ignoring stars produced a positive financial value, we used the formula shown on the right-hand side to calculate financial gain in terms of multiples:	*Formula for condition #4:* = [(Financial value derived from the observed distribution procedure × −1) + Financial value derived from using the procedure ignoring stars]/Financial value derived from the observed distribution procedure
Condition #5: If the observed distribution procedure produced a positive financial value, whereas the procedure ignoring stars produced a negative financial value, we used the formula shown on the right-hand side to calculate financial gain in terms of multiples:	*Formula for condition #5:* = [(Financial value derived from the procedure ignoring stars × −1) + Financial value derived from the observed distribution procedure]/(Financial value derived from the procedure ignoring stars × −1)

The above calculations allowed us to derive a multiple per sample per comparison. Then, for each comparison, we averaged the multiples across numerous samples (e.g. all 206 of our samples), thus producing financial gain in terms of multiples.

Table 5. Financial gains from HRM obtaining more stars averaged across all samples.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
All samples ($k = 206$)	29.24	16.09	9.77	3.70	1.96
All samples except outliers ($k = 203$)	9.08	4.94	2.76	1.25	0.54

Note. k = number of samples. Each numeric score quantifies the financial gain from obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The three excluded samples from the 'All samples except outliers' row are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; the three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all 206 samples combined = 824,924.

extremely influential outliers (i.e. samples #114 and 190–191, as listed in Appendix A in the Online Supplement). Figure 2 includes box plots associated with estimates of financial gains in Table 5. Table 5 and Figure 2 show that estimates of financial gain were larger when the financial valuation from the observed distribution procedure (fully considering the presence of stars) was compared to financial valuations from extant procedures that more heavily assume away stars. That is, stars are most heavily ignored by 40% of mean salary, followed 70% of mean salary, which more heavily ignores stars than the global procedure, which does so more than the modified global procedure, followed by the further modified procedure. Given such, estimated financial gain was 9.08, 4.94, 2.76, 1.25 or 0.54 – when comparing the observed distribution procedure to the 40% of mean salary, 70% of mean salary, global, modified global or further modified procedure, respectively (after excluding the outlier samples). Thus, the answer to our Research Question is that HRM adds greater financial value by obtaining more stars.[2]

Results based on different groups of samples

To assess the generalizability of our overall empirical conclusion, we conducted more specific analyses based on different groups of samples. Next, we report estimates of financial gain across a wide variety of occupations, performance measures, team versus individual sports, actual output versus forced rankings and one-star-only versus multiple-stars sports.

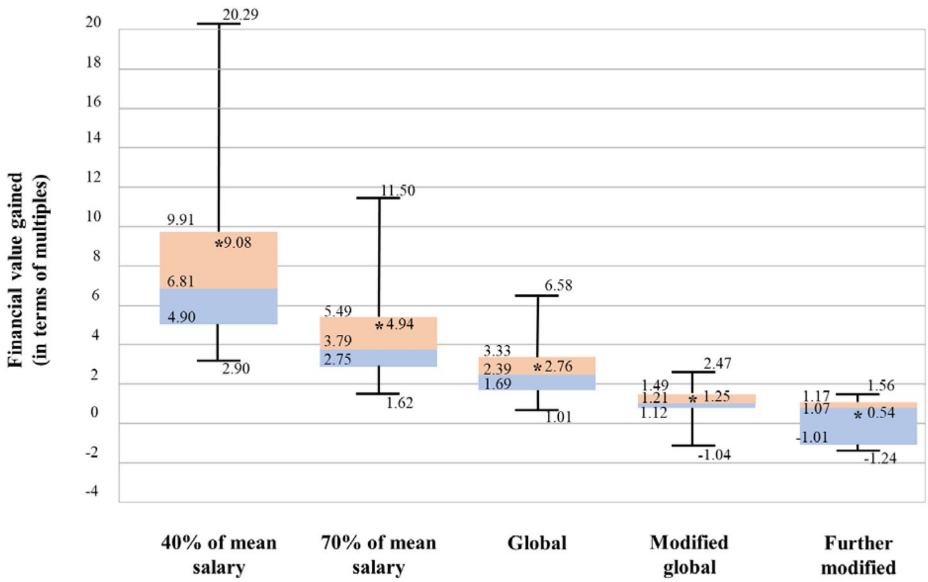

Figure 2. Box plots illustrating financial gains from HRM obtaining more stars. *Note*. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. We denote mean financial gains by using an asterisk (*); these mean values match the second row of results in Table 5. The median value is shown by the line dividing each box into two parts (with different colors). The ends of the box are the 25th and 75th quartiles. The top whiskers indicate the 95th percentile value and the bottom whiskers indicate the 5th percentile value. We calculated each box plot based on 203 samples (i.e. all our samples, except outlier samples #114 and 190–191 listed in Appendix A in the Online Supplement). These three outlier samples were not incorporated in the box plots due to their extreme influence.

Different occupations and performance measures

Tables 6 and 7 report estimates of financial gain across different occupations and performance measures, respectively. Across occupations and performance measures, HRM generally created greater financial value by obtaining more stars. So, our overall empirical conclusion remained unchanged.

Team versus individual sports

Appendix D in the Online Supplement shows estimates of financial gain based on samples indicating team (e.g. baseball, basketball) versus individual sports (e.g. bowling, tennis). We again found that HRM adds greater value as a function of obtaining more stars, regardless of team versus individual sports.

Table 6. Financial gains from HRM obtaining more stars across occupations.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Researchers ($k = 60$)	12.46	7.34	4.12	1.87	1.11
Entertainers ($k = 55$)	15.32	8.05	4.98	1.52	0.17
Entertainers except outliers ($k = 54$)	7.83	4.21	2.62	1.29	0.05
Athletes/coaches ($k = 48$)	7.75	3.46	2.03	1.03	0.45
Salespeople ($k = 11$)	4.75	2.43	1.30	0.94	0.59
Laborers ($k = 17$)	5.87	2.47	0.47	0.24	0.48
Lawyers ($k = 7$)	548.41	305.49	192.26	71.88	41.71
Lawyers except outliers ($k = 5$)	15.65	8.72	5.49	1.92	0.80
Managers ($k = 2$)	18.15	9.96	6.23	−1.10	−0.10
Medical doctors ($k = 3$)	5.35	2.93	1.83	−0.27	0.42
Recruiters ($k = 1$)	3.07	1.63	1.01	−1.10	−1.01
Financial advisors ($k = 1$)	10.18	24.77	4.72	4.69	−1.95
Programmers ($k = 1$)	4.59	2.44	1.51	−1.06	1.19

Note. k = number of samples. Each numeric score quantifies the financial gain from obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in some of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924.

Table 7. Financial gains from HRM obtaining more stars across performance measures.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Number of publications ($k = 46$)	12.40	7.50	4.08	1.68	0.94
Other research output ($k = 13$)	13.12	7.06	4.40	2.76	1.68
Nominations or appearances ($k = 41$)	8.44	4.54	2.83	1.28	−0.09
Number of movies ($k = 6$)	5.09	2.73	1.69	1.17	0.77
Number of wins ($k = 11$)	6.01	3.56	1.98	0.55	0.24
Number of medals ($k = 6$)	23.57	5.57	2.70	1.21	−0.34
Other sports output ($k = 30$)	5.38	3.06	1.94	1.17	0.67
Sales/revenue ($k = 17$)	151.76	84.38	53.07	21.47	10.30
Sales/revenue except outliers ($k = 16$)	6.55	3.49	2.16	1.35	0.65
Income from grants/awards ($k = 2$)	214.47	110.21	67.76	8.54	3.97
Income from grants/awards except outliers ($k = 1$)	9.35	4.82	2.97	2.97	1.55
Various productivity indices ($k = 3$)	4.19	2.32	1.46	−0.37	0.41
Other output ($k = 31$)	48.88	27.37	16.20	5.17	4.26
Other output except outliers ($k = 30$)	7.66	4.41	1.72	0.34	0.29

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in some of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924.

Table 8. Financial gains from HRM obtaining more stars across actual output versus forced rankings.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Actual output ($k = 159$)	34.82	19.46	11.82	4.42	2.57
Actual output, except outliers ($k = 156$)	8.69	5.02	2.75	1.25	0.73
Forced rankings ($k = 47$)	10.37	4.67	2.81	1.27	−0.12

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in one of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924.

Actual output versus forced rankings

Table 8 provides estimates of financial gain based on samples indicating actual output (e.g. meters of rice planted, lines of code per hour, revenue generated) versus forced rankings (e.g. expert rankings for Edgar Allan Poe Awards, Olympic medals). Regardless of actual output versus forced rankings, results supported our overall empirical conclusion that HRM adds greater value by obtaining more stars.

One-star-only versus multiple-stars sports

Some sports represented in our samples are one-star-only sports in that multiple players usually cannot be stars simultaneously. Appendix E in the Online Supplement shows estimates of financial gain based on samples indicating one-star-only (e.g. tennis, alpine skiers) versus multiple-stars sports (e.g. baseball, hockey). Regardless of one-star-only or multiple-stars sports, results again supported the overall empirical conclusion that HRM can add greater value by obtaining more stars.

Results indicating significant yet diminishing returns from HRM obtaining more stars

Beyond our overall empirical conclusion that HRM can add greater financial value as a function of obtaining more stars, Tables 5–8 in the manuscript and Appendices D and E in the Online Supplement also indicate that HRM often produces significant yet diminishing returns by increasingly focusing on obtaining the most productive stars.

Specifically, to estimate the financial gain from obtaining more stars, we compared the financial valuation from the observed distribution procedure (which does not assume away any stars) to the financial valuation from each of the other procedures (assuming away stars by varying degrees). Because the 40% of mean salary procedure assumes away stars by the greatest amount, financial gain estimated from comparing the observed distribution procedure to 40% of mean salary was generally the largest (e.g. 9.08 in Table 5, without outlier samples). In turn, given the 70% of mean salary procedure that assumes away stars by a lesser degree, estimated financial gain from comparing the observed distribution procedure to 70% of mean salary was generally the second largest (e.g. 4.94 in Table 5, without outlier samples). The global and modified global procedures assume away stars by increasingly smaller amounts (i.e. ignoring the top 15% or 3% of all performers in a sample, respectively). Accordingly, estimated financial gain from comparing the observed distribution procedure to the global and modified global procedures were usually the third and fourth largest (e.g. 2.76 and 1.25 in Table 5, excluding outlier samples). Finally, the further modified procedure assumes away stars by an even smaller degree by ignoring just the top 1%, so that financial gain estimated from comparing the observed distribution procedure to the further modified procedure was typically the smallest (e.g. 0.54 in Table 5, without outlier samples).

In particular, financial gains estimated from comparing the observed distribution procedure to the further modified procedure were usually close to or between −1 and 1, indicating that HRM overall neither added to nor diminished value by obtaining the top 1%. Even the financial gains estimated from comparing the observed distribution procedure to the modified global procedure were often close to values between −1 and 1 or just slightly higher than 1. This means that HRM generally added, at best, a relatively small amount of financial value by more broadly obtaining the top 3% (e.g. estimated financial gain of 1.25 in Table 5 indicates that obtaining stars including the top 3% was on average about 25% more financially valuable than obtaining stars except the top 3%). These findings show that HRM often produces significant yet diminishing returns by increasingly focusing on obtaining the most productive stars.

Results indicating when diminishing returns are stronger or weaker

At the same time, our results indicate that the diminishing pattern is not always very apparent. As shown in Tables 6 and 7, sometimes financial gains estimated from comparing the observed distribution procedure to the further modified procedure were significantly or at least incrementally higher than 1 (e.g. among researchers or programmers in Table 6, as well as among other research output and income from grants/awards in Table 7). In such instances, the diminishing pattern was weaker. To better understand when obtaining stars results in stronger or weaker diminishing returns (i.e. when increasingly focusing on obtaining the most productive stars overall adds little/no additional value or adds possibly slightly greater value), we conducted several follow-up analyses shown below.

Longer versus shorter time frames

Table 9 reports financial gains by longer versus shorter time frames for allowing more versus fewer stars to emerge, respectively. As noted by Joo et al. (2017), a relatively long time period is needed for some workers in a group to achieve star-level output. Financial gains were larger in samples indicating longer time frames, and smaller in samples indicating shorter time frames. In particular, financial gains estimated from comparing the observed distribution procedure to the modified global and further modified procedures were larger among samples with longer time frames, yet smaller among short time frame samples (i.e.

Table 9. Financial gains from HRM obtaining more stars across time frames.

		Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
		40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Longer time frames	Lifetime (i.e. career) [$k=41$]	9.33	4.98	2.32	1.08	0.46
	One decade ($k=42$)	12.22	6.69	4.19	1.66	0.90
	Multiple years ($k=13$)	12.22	6.50	3.89	2.61	1.90
	Summary of longer time frames: Multiple years to lifetime ($k=96$)	*10.99*	*5.93*	*3.35*	*1.54*	*0.85*
Shorter time frames	One to three months ($k=7$)	3.93	2.01	1.21	1.14	1.31
	Two weeks ($k=2$)	5.51	2.86	1.76	1.22	−0.03
	Per day ($k=4$)	6.78	8.92	1.94	2.05	−0.09
	Per hour ($k=12$)	6.58	2.32	0.69	−0.38	0.03
	Shorter than per hour ($k=2$)	4.23	4.06	0.06	1.24	1.89
	Summary of shorter time frames: Shorter than per hour to three months ($k=27$)	*5.67*	*3.39*	*1.04*	*0.61*	*0.48*

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. Total N (i.e. number of workers) for all samples represented in this table combined = 810,060.

1.54 and 0.85 versus 0.61 and 0.48, respectively). Thus, the diminishing pattern is weaker when the time frame is longer.

Higher versus lower performance ceilings

Table 10 shows financial gains by higher versus lower performance ceilings for allowing more versus fewer stars to emerge, respectively. As noted by Aguinis et al. (2016), more stars tend to emerge in contexts with higher ceilings. Financial gains were larger among higher ceiling samples with or without outlier samples, and smaller among lower ceiling samples. In particular, financial gains estimated from comparing the observed distribution procedure to the modified global and further modified procedures were larger among samples characterized by higher ceilings (without outlier samples), but smaller among lower ceiling samples (i.e. 1.42 and 0.56 versus 0.16 and 0.38, respectively). So, the diminishing pattern is weaker when the ceiling is higher.

Higher versus lower control over individual output

Table 11 reports financial gains based on higher versus lower control by workers over their own individual output. We broadly defined higher-control occupations as 'typical' occupations (e.g. sales, laborers, managers), and lower-control occupations as 'special' occupations (e.g. sports, entertainment). This is because in special occupations, environmental factors and luck (e.g. actions of competing players, getting noticed by a movie director in a chance event) likely play a stronger role while individual actions play a weaker role (e.g. Macnamara et al., 2016).

Table 10. Financial gains from HRM obtaining more stars across performance ceilings.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Higher performance ceiling ($k = 179$)	32.83	18.13	11.11	4.23	2.19
Higher performance ceiling except outliers ($k = 176$)	9.64	5.31	3.06	1.42	0.56
Lower performance ceiling ($k = 27$)	5.44	2.52	0.83	0.16	0.38

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in one of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924.

Table 11. Financial gains from HRM obtaining more stars across typical versus special occupations (i.e. higher versus lower control over individual output).

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Typical occupations (i.e. higher control)					
With outlier samples ($k = 43$)	94.44	52.67	32.40	12.03	7.11
Without outlier samples ($k = 41$)	7.33	4.14	1.83	0.58	0.44
Special occupations (i.e. lower control)					
With outlier samples ($k = 163$)	12.04	6.44	3.79	1.50	0.60
Without outlier samples ($k = 162$)	9.52	5.15	3.00	1.43	0.56

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in some of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924. The typical occupations included in this table are bank tellers, retail sales associates, call center employees, fundraising callers, paper sorters, pelt pullers, toll-ticket sorters, typists, operators, sewers, electrical fixture assemblers, lawyers, managers, agricultural workers, managers, transcribers, produce packers, recruiters, financial advisors, doctors, laundry workers and programmers (see Table 1 for more details).

Although a possible disadvantage of using individual output data is that workers often do not have much control over their output, our samples vary in terms of workers' degree of control over their individual output. Financial gains were smaller among higher control samples (without outlier samples), and larger among lower control samples (without outlier samples). In particular, financial gains estimated from comparing the observed distribution procedure to the modified global and further modified procedures were smaller among samples characterized by higher control (without outlier samples), but larger among lower control samples (i.e. 0.58 and 0.44 versus 1.43 and 0.56, respectively). These findings show that the diminishing pattern is stronger when workers have higher control over their output.

Different distribution shapes

Table 12 reports financial gains based on two groups of distribution shapes: (1) exponential-tail (characterized by right-tails that rapidly decay, or fall off, at the very end) versus (2) lognormal or pure power law (whose right-tails do not decay as rapidly as the exponential-tail) (Aguinis et al., 2018; Joo et al., 2017). In other words, output differences among stars at the top will tend to be smaller among exponential-tail samples, and larger among lognormal or pure power law samples. We thus expected a stronger diminishing pattern among exponential-tail

Table 12. Financial gains from HRM obtaining more stars across distribution shapes.

	Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Exponential-tail ($k = 136$)	26.38	14.63	8.94	3.75	1.66
Exponential-tail except outliers ($k = 135$)	8.24	4.53	2.58	1.23	0.45
Lognormal or pure power law ($k = 30$)	12.82	6.91	4.26	1.74	0.93
Normal or Weibull ($k = 17$)	5.17	2.21	0.33	0.07	0.34

Note. k = number of samples. Each numeric score quantifies the financial gain from HRM obtaining more stars. Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded sample in one of the rows is sample #191, as listed in Appendix A in the Online Supplement; this sample is an outlier due to its extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 700,603.

samples. Consistent with our expectation, financial gains were smaller among exponential-tail samples (without outlier samples), and larger among lognormal or pure power law samples. In particular, financial gains estimated from comparing the observed distribution procedure to the modified global and further modified procedures were smaller among samples characterized by exponential-tails (without outlier samples) rather than lognormal or pure power law (i.e. 1.23 and 0.45 versus 1.74 and 0.93, respectively).[3] Moreover, Table 13 shows that a large minority of the researchers samples best-fit the lognormal or pure power law, whereas the exponential-tail was the more clearly best-fitting shape among the entertainers except outliers, athletes/coaches and salespeople samples. Financial gains estimated from comparing the observed distribution procedure to the modified global and further modified procedures were larger among the researchers samples, but smaller among the other three sample groups (1.87 and 1.11 versus 1.29, 1.03, or 0.94 and 0.05, 0.45, or 0.59). In short, the diminishing pattern is stronger if the underlying distribution best fits the exponential-tail, and weaker if the best-fitting shape is the lognormal or pure power law.

Discussion

The observed distribution procedure (which does not assume away any stars) generally produced higher financial valuations than other extant utility analysis procedures (which assume away stars). Stated differently, we demonstrated empirically that previous methods of assessing financial value underestimated the value brought by obtaining more stars, indicating how much greater value HRM can add by

Table 13. Proportion of samples that best-fit competing distribution shapes, and financial gains from HRM obtaining more stars across different occupations.

	Proportion of samples that best-fit each distribution shape, per occupation			Financial gain calculated by comparing the observed distribution procedure to each of the five procedures below that ignore stars by varying degrees				
	Exponential-tail	Lognormal or pure power law	Other	40% of mean salary	70% of mean salary	Global	Modified global	Further modified
Occupations considered								
Researchers ($k = 60$)	51.67%	31.67%	16.67%	12.46	7.34	4.12	1.87	1.11
Entertainers except outliers ($k = 54$)	88.89%	3.70%	7.41%	7.83	4.21	2.62	1.29	0.05
Athletes/coaches ($k = 48$)	89.58%	6.25%	4.17%	7.75	3.46	2.03	1.03	0.45
Salespeople ($k = 11$)	54.55%	0.00%	45.45%	4.75	2.43	1.30	0.94	0.59
Not considered								
Entertainers ($k = 55$)	87.27%	3.64%	9.09%	15.32	8.05	4.98	1.52	0.17
Laborers ($k = 17$)	17.65%	11.76%	70.59%	5.87	2.47	0.47	0.24	0.48
Lawyers ($k = 7$)	28.57%	28.57%	42.86%	548.41	305.49	192.26	71.88	41.71
Lawyers except outliers ($k = 5$)	20.00%	40.00%	40.00%	15.65	8.72	5.49	1.92	0.80
Managers ($k = 2$)	0.00%	100.00%	0.00%	18.15	9.96	6.23	−1.10	−0.10
Medical doctors ($k = 3$)	100.00%	0.00%	0.00%	5.35	2.93	1.83	−0.27	0.42
Recruiters ($k = 1$)	0.00%	0.00%	100%	3.07	1.63	1.01	−1.10	−1.01
Financial advisors ($k = 1$)	0.00%	0.00%	100%	10.18	24.77	4.72	4.69	−1.95
Programmers ($k = 1$)	0.00%	0.00%	100%	4.59	2.44	1.51	−1.06	1.19

Note. For the purpose of comparing occupations that heavily follow the exponential-tail shape versus those that heavily follow the lognormal or pure power law, occupations under 'Not Considered' were excluded because they included outlier samples (in the case of Entertainers, $k = 55$); the majority of samples best-fit distribution shapes other than the exponential-tail, lognormal, or pure power law (in the case of Laborers, $k = 17$); or the number of samples in the occupation was likely too small to conduct a meaningful comparison (i.e. $k < 10$). Financial gain is expressed in terms of multiples (i.e. how many times the financial value estimated from the observed distribution procedure is greater or smaller than the financial value estimated from another procedure assuming away stars). A positive multiple greater than 1 means that the financial value estimated using the observed distribution procedure (fully considering stars) is greater than the financial value estimated using another procedure (ignoring stars by a certain degree); that is, HRM obtaining more stars added to financial value. A negative multiple lower than −1 means that the financial value estimated from the observed distribution procedure is less than that estimated from another procedure; in other words, HRM obtaining more stars diminished value. The excluded samples in some of the rows are samples #114 and 190–191, as listed in Appendix A in the Online Supplement; these three samples are outliers due to their extreme influence. Total N (i.e. number of workers) for all samples represented in this table combined = 824,924.

helping firms obtain more star performers. We also found that HRM often generates significant yet diminishing returns by increasingly focusing on obtaining the most productive stars. Financial gains estimated from comparing the observed distribution procedure to the further modified procedure, in particular, indicated that HRM overall neither added to nor diminished value by obtaining the top 1% of all performers in a sample. Our results offer several theoretical contributions to HRM, as described below.

Theoretical implications for HRM

We contribute to HRM theory by helping clarify whether, how, when and why HRM produces greater financial value as a function of obtaining more stars. First, regarding not just whether but also how HRM generates greater value by obtaining more stars, our results offer a nonlinear rather than a linear understanding of HRM's value. Given our overall empirical finding that obtaining more stars can add greater value to firms, our evidence further indicated that increasingly focusing on obtaining the most productive stars often generates diminishing returns. On one hand, we found that obtaining more stars can contribute as much as nine times greater financial value, reflecting the maximum extent to which obtaining stars may benefit firms (see the row 'all samples except outliers' in Table 5). These findings were replicated in a variety of occupations and performance measures, as summarized in Tables 6 and 7. At the same time, our results showed that obtaining more stars can often produce diminishing returns, as indicated by the rapidly decreasing gains that can be seen going from the left-hand columns to the right in Tables 6 and 7. Stated differently, Tables 6 and 7 showed that by obtaining more and more productive stars (e.g. the top 1% rather than more broadly the top 3%), HRM creates additional financial value that often becomes increasingly smaller, close to zero. We thus contribute to a nonlinear model of the value of HRM, which generates significant yet diminishing returns by increasingly focusing on obtaining the most productive stars.

Second, we contribute to a better understanding of *when* obtaining stars results in stronger or weaker diminishing returns. Results showed that the diminishing pattern is stronger when output differences among top stars are relatively small. For example, Table 9 organizes results based on longer versus shorter time frames, where the latter creates smaller output differences among top stars by imposing a lower ceiling constraint on how much output top stars can accumulate. Among samples with shorter time frames (and thus smaller output differences among top stars), the diminishing pattern was stronger. Similarly, Table 10 reports results based on higher versus lower performance ceilings, where the latter creates smaller output differences among top stars. The diminishing pattern was more heightened among samples with lower ceilings (and thus smaller output differences among top stars).

Results in Tables 11–13 offered further support for small output differences among top stars as a key context strengthening or diminishing returns from obtaining more stars. Table 11 shows results based on typical (characterized by higher control over output) versus special occupations (lower control). Workers who have a higher degree of control over their own output are generally more likely to attain star-level output, so that output differences among top stars are smaller – a context which facilitates stronger diminishing returns. As shown in Table 11, the diminishing pattern was indeed stronger among 'typical' occupations such as sales, laborers and managers where workers have higher control over output. Meanwhile, diminishing returns were weaker among 'special' occupations such as sports and entertainment, characterized by lower control. Table 12 indicates that the diminishing pattern was more apparent among samples that best-fit the exponential-tail shape, where top stars have increasingly smaller output differences. In contrast, the diminishing pattern in Table 12 was weaker among samples that best-fit other non-normal shapes (i.e. pure power law, lognormal), where top stars have increasingly larger (not smaller) output differences. Similarly, Table 13 shows that the diminishing pattern was stronger in occupations that heavily fit shapes with increasingly smaller output differences at the top, but weaker among occupations that heavily fit distribution shapes with increasingly larger output differences among top stars. Thus, as our second contribution to HRM theory, we clarify that diminishing returns from HRM are stronger when output differences among top stars are small.

Third, our study also provides a theoretical explanation for *why* obtaining stars can produce significant yet diminishing returns. Going beyond our finding that diminishing returns are more likely when there are small output differences among top stars, we further theorize that those small differences better enable various costs to manifest – to the effect of diminishing the returns from obtaining the most productive stars who are just slightly more productive than other top stars. The star literature offers several examples of such costs (i.e. created by small output differences among top stars). For instance, relatively small output differences among top stars (i.e. 'excessive similarity') create ambiguity in the status hierarchy, which those stars seek to clarify by competing with each other more intensely and perhaps in a disruptive manner (Groysberg et al., 2011, p. 725). As another example, when output differences among stars are small, the same psychological tendency would be considered biases that lead to overvaluing top stars (Massey & Thaler, 2013). As yet another example, if there are numerous and similarly productive top stars serving as many role models, top stars can (inadvertently) signal inconsistent sets of behavioral cues that inhibit non-stars' learning, especially since 'exemplary performance can be demonstrated in many different ways' (Call et al., 2020, p. 6). In short, we contribute to a better understanding of not only how and when, but also why obtaining stars often produce significant yet diminishing returns.

Contributions to the star literature

Our explanation of the nonlinear pattern also contributes back to the star literature from which we borrowed theoretical rationale. To explain why obtaining stars may often produce significant but diminishing returns, we had stated that relatively small output differences among top stars may create various costs which diminish the returns from obtaining the most productive stars. This explanation then helps integrate specific explanations for stars' curvilinear influence proposed in past research, thereby improving theoretical parsimony.

Specifically, to the effect of reducing theoretical parsimony, studies in the star literature have offered numerous and disparate explanations for stars' often curvilinear impact, such as disruptions caused by competition among stars to establish a clear status hierarchy (Groysberg et al., 2011); biases that lead managers to over-value top stars (Massey & Thaler, 2013); and stars' behaviors that reduce the learning of non-stars (un)intentionally (Call et al., 2020). Given such, small output differences among top stars may facilitate or allow these mechanisms to occur in the first place. If stars have large differences in output, a clear status hierarchy would exist and prevent disruptive competition between stars who seek to establish a clear hierarchy (Groysberg et al., 2011); if top stars have large enough differences in output, managers' tendency to disproportionately value the contributions of top stars may be justified (Massey & Thaler, 2013); and large output differences among top stars likely mean that there are just one or few of those stars who serve as clear role models for non-stars, reducing inconsistent cues from stars about how to work (Call et al., 2020). Thus, our emphasis on small output differences among top stars serves as a broad theoretical component that is common across and, therefore, links together the many specific explanations for stars' curvilinear impact.

Implications for HRM practice

Our results highlight the need to use utility analysis procedures that more fully consider the presence of stars. The financial value estimated from the observed distribution procedure (which fully considers stars) was on average up to nine times greater than the value estimated from other extant procedures (which ignore stars by varying degrees). In other words, extant procedures often significantly underestimated the value brought by obtaining more stars. It thus follows that firms should ideally use the observed distribution procedure, which estimated HRM's value most accurately by incorporating reality as is rather than ignoring stars who exist. By considering stars more fully, financial valuations of

HRM based on the observed distribution procedure are not only more accurate, but also more comparable with valuations of other business areas (e.g. marketing) that recognize the reality that often few products and services contribute disproportionately to a firm's bottom line (Mariotti, 2008).

Yet, full data on performance distributions needed to implement the observed distribution procedure are often unavailable. So, it may be necessary to use other procedures that do not require full data. Among the five procedures ignoring stars by at least some degree, we suggest using the further modified procedure because it led to the most consistently accurate results (i.e. estimated financial values closest to those derived from the observed distribution procedure). The superior accuracy of the further modified procedure is evidenced by the fact that financial gains in Tables 5–13 (and in Appendices D and E in the Online Supplement) were smallest (i.e. closest to values between 1 and −1), when the observed distribution procedure was compared to the further modified procedure. The further modified procedure enjoys superior accuracy by assuming away stars by the least amount (i.e. by ignoring just the top 1%), compared to the other four procedures that assume away stars by a larger degree (e.g. the top 3% ignored). This way, our results help address demands from top management to financially justify HRM practices.

Limitations and future research directions

First, when implementing utility analysis procedures, we did not incorporate all possible costs of obtaining stars because we examined value created from HRM, not captured. So, used in isolation, each of the utility analysis procedures in our study likely leads to valuations that are too high (Whyte & Latham, 1997). To examine net value after at least some of the initial value created has been captured and otherwise deducted, practical implementation of utility analysis should add cost-related parameters in Equation (1) by including additional costs such as opportunity costs, taxes and compensation (Boudreau, 1983; Cabrera & Raju, 2001; Sturman, 2000). Indeed, regarding compensation, some stars may be better equipped than others for capturing value created (Kehoe et al., 2018).

Second, our results may not generalize to contexts where performance is defined and operationalized as behaviors. The reason is that behavioral performance may follow more normally-distributed shapes than heavily right-tailed distributions (Beck et al., 2014). But when performance is defined and operationalized in terms of output as done in our study, performance is likely to follow non-normally-distributed curves characterized by heavy right-tails containing stars (Aguinis et al., 2018). It is possible that our results only apply to contexts where worker performance is defined and operationalized as output. Thus, future research can examine the extent to which our findings also apply to behavior- rather than output-based performance.

Concluding remarks

Our overall empirical finding was that HRM creates greater financial value by obtaining more stars. We also offered several theoretical contributions to HRM and the star literature. First, our results offered a nonlinear model of HRM's value, where HRM produces significant yet diminishing returns by increasingly focusing on obtaining the most productive stars. Second, regarding when, we provided evidence that diminishing returns from HRM are stronger when output differences among top stars are relatively small. Third, regarding why, we explained that small output differences among top stars may create various costs which diminish the returns from obtaining the most productive stars. Fourth, our explanation of HRM's nonlinear pattern also contributed to the star literature by helping integrate a number of specific explanations for stars' curvilinear influence proposed in past research. From a practical view, we highlighted the need to use utility analysis procedures that more fully consider the presence of stars because extant procedures often significantly underestimate the value brought by obtaining more stars. By considering stars more fully, valuations of HRM are more accurate and also comparable with valuations of other business areas that recognize the reality that often few products and services contribute disproportionately to a firm's bottom line. In closing, we hope our article will stimulate HRM research and applications that fully consider the prevalence of stars and their relative value to firms.

Notes

1. The nine distinct combinations of parameter values are as follows. **(1)**: $N_s = 50$, $r = 0.3$, $\phi = 0.3989$, $p = 0.5$, $N = 100$, $C = \$75.26$; **(2)**: $N_s = 50$, $r = 0.4$, $\phi = 0.3989$, $p = 0.5$, $N = 100$, $C = \$75.26$; **(3)**: $N_s = 50$, $r = 0.5$, $\phi = 0.3989$, $p = 0.5$, $N = 100$, $C = \$75.26$; **(4)**: $N_s = 32$, $r = 0.5$, $\phi = 0.3572$, $p = 0.32$, $N = 100$, $C = \$75.26$; **(5)**: $N_s = 15$, $r = 0.5$, $\phi = 0.2323$, $p = 0.15$, $N = 100$, $C = \$75.26$; **(6)**: $N_s = 15$, $r = 0.5$, $\phi = 0.2323$, $p = 0.15$, $N = 100$, $C = \$752.63$; **(7)**: $N_s = 15$, $r = 0.5$, $\phi = 0.2323$, $p = 0.15$, $N = 100$, $C = \$1{,}505.25$; **(8)**: $N_s = 3$, $r = 0.5$, $\phi = 0.2323$, $p = 0.15$, $N = 20$, $C = \$75.26$; and **(9)**: $N_s = 1$, $r = 0.5$, $\phi = 0.2323$, $p = 0.15$, $N = 7$, $C = \$75.26$. Table 3 includes a detailed rationale for the use of these parameter values.
2. As a robustness check, we reconducted the analyses reported in Table 5 by excluding non-US samples (i.e. samples #43–44, 166, 194–195, 199–204 and 206 described in Appendix A in the Online Supplement). We also reconducted the

analyses in Table 5 with non-US samples included, but without using income values updated for the non-US samples (from non-US sources). Results from the additional analyses were essentially the same as those in Table 5, as detailed in Appendix C in the Online Supplement.

3. Table 12 also reports financial gains based on samples following normality-based shapes (i.e. Weibull or normal), where output differences among stars at the top will tend to be even smaller than among exponential-tail samples. Consistent with our expectation, in this comparison, we found a stronger diminishing pattern among the normality-based shapes than among exponential-tail ones.

Acknowledgements

We thank Yoshio Yanadori and three anonymous reviewers from *The International Journal of Human Resource Management* for their highly constructive feedback on previous drafts. Shorter versions of this manuscript including some of the data were presented at the meetings of the Academy of Management, Chicago, August 2018, and Society for Industrial and Organizational Psychology, Chicago, April 2018.

ORCID

Harry Joo http://orcid.org/0000-0001-5494-5665
Herman Aguinis http://orcid.org/0000-0002-3485-9484
Joowon Lee http://orcid.org/0000-0002-4669-3012
Hannah Kremer http://orcid.org/0000-0001-5245-3966
Isabel Villamor http://orcid.org/0000-0003-0845-7016

Data availability statement

The data that support the findings of this study are available from the corresponding author, Herman Aguinis, upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Aguinis, H., Ji, Y. H., & Joo, H. (2018). Gender productivity gap among star performers in STEM and other scientific fields. *Journal of Applied Psychology*, *103*(12), 1283–1306. https://doi.org/10.1037/apl0000331

Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first-century organizations. *Personnel Psychology*, *67*(2), 313–350. https://doi.org/10.1111/peps.12054

Aguinis, H., O'Boyle, E., Gonzalez-Mulé, E., & Joo, H. (2016). Cumulative advantage: Conductors and insulators of heavy-tailed productivity distributions and productivity stars. *Personnel Psychology*, *69*(1), 3–66. https://doi.org/10.1111/peps.12095

Albrecht, S. L., & Marty, A. (2020). Personality, self-efficacy and job resources and their associations with employee engagement, affective commitment and turnover intentions. *The International Journal of Human Resource Management*, *31*(5), 657–681. https://doi.org/10.1080/09585192.2017.1362660

Amankwah-Amoah, J., Nyuur, R. B., & Ifere, S. (2017). A question of top talent? The effects of lateral hiring in two emerging economies. *The International Journal of Human Resource Management*, *28*(11), 1527–1546. https://doi.org/10.1080/09585192.2015.1116456

Ammann, M., Horsch, P., & Oesch, D. (2016). Competing with superstars. *Management Science*, *62*(10), 2842–2858. https://doi.org/10.1287/mnsc.2015.2266

Anderson, S. M., & Muchinsky, M. (1991). An examination of the robustness of the general utility function. *Educational and Psychological Measurement*, *51*(1), 49–65. https://doi.org/10.1177/0013164491511004

Asgari, E., Hunt, R., Lerner, D., Townsend, D., Hayward, M. L., & Kiefer, K. (2021). Red giants or black holes? The antecedent conditions and multi-level impacts of star performers. *Academy of Management Annals*, *15*(1), 223–265. https://doi.org/10.5465/annals.2019.0061

Beck, J. W., Beatty, A. S., & Sackett, P. R. (2014). On the distribution of job performance: The role of measurement characteristics in observed departures from normality. *Personnel Psychology*, *67*(3), 531–566. https://doi.org/10.1111/peps.12060

Becker, B. E., & Huselid, M. A. (1992). Direct estimates of *SDy* and the implications for utility analysis. *Journal of Applied Psychology*, *77*(3), 227–233. https://doi.org/10.1037/0021-9010.77.3.227

Bedeian, A. G., & Armenakis, A. A. (1998). The cesspool syndrome: How dreck floats to the top of declining organizations. *Academy of Management Perspectives*, *12*(1), 58–67. https://doi.org/10.5465/ame.1998.254978

Behfar, K., & Okhuysen, G. A. (2018). Perspective—Discovery within validation logic: Deliberately surfacing, complementing, and substituting abductive reasoning in hypothetico-deductive inquiry. *Organization Science*, *29*(2), 323–340. https://doi.org/10.1287/orsc.2017.1193

Bendapudi, N., & Leone, R. P. (2001). How to lose your star performer without losing customers, too. *Harvard Business Review*, *79*(10), 104–115.

Bishara, A. J., & Hittner, J. B. (2012). Testing the significance of a correlation with nonnormal data: Comparison of Pearson, Spearman, transformation, and resampling approaches. *Psychological Methods*, *17*(3), 399–417. https://doi.org/10.1037/a0028087

Boudreau, J. W. (1991). Utility analysis for decisions in human resource management. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 621–745). Consulting Psychologists Press.

Boudreau, J. W. (1983). Economic considerations in estimating the utility of human resource productivity improvement programs. *Personnel Psychology*, *36*(3), 551–576. https://doi.org/10.1111/j.1744-6570.1983.tb02235.x

Brogden, H. E. (1949). When testing pays off. *Personnel Psychology*, *2*(2), 171–185. https://doi.org/10.1111/j.1744-6570.1949.tb01397.x

Burke, M. J., & Frederick, J. T. (1984). Two modified procedures for estimating standard deviations in utility analyses. *Journal of Applied Psychology*, *69*(3), 482–489. https://doi.org/10.1037/0021-9010.69.3.482

Burke, M. J., & Frederick, J. T. (1986). A comparison of economic utility estimates for alternative *SDy* estimation procedures. *Journal of Applied Psychology*, *71*(2), 334–339. https://doi.org/10.1037/0021-9010.71.2.334

Cabrera, E. F., & Raju, N. S. (2001). Utility analysis: Current trends and future directions. *International Journal of Selection and Assessment*, *9*(1&2), 92–102. https://doi.org/10.1111/1468-2389.00166

Cappelli, P., & Keller, J. R. (2017). The historical context of talent management. In D. G. Collings, K. Mellahi, & W. F. Cascio (Eds.), *The Oxford handbook of talent management* (pp. 23–42). Oxford University Press.

Call, M. L., Campbell, E. M., Dunford, B. B., Boswell, W. R., & Boss, R. W. (2020). Shining with the Stars? Unearthing how group star proportion shapes non-star performance. *Personnel Psychology*. Advance online publication. https://doi.org/10.1111/peps.12420.

Carretero-Gómez, J. M., & Cabrera, E. F. (2012). An empirical evaluation of training using multi-attribute utility analysis. *Journal of Business and Psychology*, *27*(2), 223–241. https://doi.org/10.1007/s10869-011-9241-6

Cascio, W. F., & Aguinis, H. (2019). *Applied psychology in talent management* (8th ed.). Thousand Oaks.

Cascio, W. F., & Boudreau, J. W. (2011). *Investing in people: Financial impact of human resource initiatives* (2nd ed.). Pearson.

Cascio, W. F., & Montealegre, R. (2016). How technology is changing work and organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*(1), 349–375. https://doi.org/10.1146/annurev-orgpsych-041015-062352

Cascio, W. F., & Ramos, R. A. (1986). Development and application of a new method for assessing job performance in behavioral/economic terms. *Journal of Applied Psychology*, *71*(1), 20–28. https://doi.org/10.1037/0021-9010.71.1.20

Cascio, W. F., & Silbey, V. (1979). Utility of the assessment center as a selection device. *Journal of Applied Psychology*, *64*(2), 107–118. https://doi.org/10.1037/0021-9010.64.2.107

Chen, J. S., & Garg, P. (2018). Dancing with the stars: Benefits of a star employee's temporary absence for organizational performance. *Strategic Management Journal*, *39*(5), 1239–1267. https://doi.org/10.1002/smj.2758

Cortina, J. M., Aguinis, H., & DeShon, R. P. (2017). Twilight of dawn or of evening? A century of research methods in the Journal of Applied Psychology. *Journal of Applied Psychology*, *102*(3), 274–290. https://doi.org/10.1037/apl0000163

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., & McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing*, *30*(5), 696–713. https://doi.org/10.1016/j.jbusvent.2015.01.001

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. University of Illinois Press.

Cronshaw, S. F., Alexander, R. A., Wiesner, W. H., & Barrick, M. R. (1987). Incorporating risk into selection utility: Two models for sensitivity analysis and risk simulation. *Organizational Behavior and Human Decision Processes*, *40*(2), 270–286. https://doi.org/10.1016/0749-5978(87)90016-1

Dunnette, M. D., Rosse, R. L., Houston, J. S., Hough, L. M., Toquam, J., Lammlein, S., King, K. W., Bosshardt, M. J., & Keyes, M. (1982). *Development and validation of an industry-wide electric power plant operator selection system*. Edison Electric Institute.

Grigoriou, K., & Rothaermel, F. T. (2014). Structural microfoundations of innovation: The role of relational stars. *Journal of Management*, *40*(2), 586–615. https://doi.org/10.1177/0149206313513612

Groysberg, B., & Lee, L. E. (2008). The effect of colleague quality on top performance: The case of security analysts. *Journal of Organizational Behavior*, *29*(8), 1123–1144. https://doi.org/10.1002/job.558

Groysberg, B., & Lee, L. E. (2009). Hiring stars and their colleagues: Exploration and exploitation in professional service firms. *Organization Science*, *20*(4), 740–758. https://doi.org/10.1287/orsc.1090.0430

Groysberg, B., Polzer, J. T., & Elfenbein, H. A. (2011). Too many cooks spoil the broth: How high-status individuals decrease group effectiveness. *Organization Science*, *22*(3), 722–737. https://doi.org/10.1287/orsc.1100.0547

Gula, B., Vaci, N., Alexandrowicz, R. W., & Bilalić, M. (2021). Never too much—The benefit of talent to team performance in the National Basketball Association: Comment on Swaab, Schaerer, Anicich, Ronay, and Galinsky (2014). *Psychological Science*, *32*(2), 301–304. https://doi.org/10.1177/0956797620960396

Han, S., & Ravid, S. A. (2020). Star turnover and the value of human capital—Evidence from Broadway shows. *Management Science*, *66*(2), 958–978. https://doi.org/10.1287/mnsc.2018.3177

Harder, J. W. (1992). Play for pay: Effects of inequity in a pay-for-performance context. *Administrative Science Quarterly*, *37*(2), 321–335. https://doi.org/10.2307/2393227

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica*, *68*(3), 575–603. https://doi.org/10.1111/1468-0262.00124

Hazer, J. T., & Highhouse, S. (1997). Factors influencing managers' reactions to utility analysis: Effects of *SDy* method, information frame, and focal intervention. *Journal of Applied Psychology*, *82*(1), 104–112. https://doi.org/10.1037/0021-9010.82.1.104

Hoffman, C. C., & Thornton, G. C. (1997). Examining selection utility where competing predictors differ in adverse impact. *Personnel Psychology*, *50*(2), 455–470. https://doi.org/10.1111/j.1744-6570.1997.tb00916.x

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*(3), 340–362. https://doi.org/10.1016/0001-8791(86)90013-8

Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*(6), 869–879. https://doi.org/10.1037/0021-9010.85.6.869

Joo, H., Aguinis, H., & Bradley, K. J. (2017). Not all nonnormal distributions are created equal: Improved theoretical and measurement precision. *Journal of Applied Psychology*, *102*(7), 1022–1053. https://doi.org/10.1037/apl0000214

Kehoe, R. R., Lepak, D. P., & Bentley, F. S. (2018). Let's call a star a star: Task performance, external status, and exceptional contributors in organizations. *Journal of Management*, *44*(5), 1848–1872. https://doi.org/10.1177/0149206316628644

Kehoe, R. R., & Tzabbar, D. (2015). Lighting the way or stealing the shine? An examination of the duality in star scientists' effects on firm innovative performance. *Strategic Management Journal*, *36*(5), 709–727. https://doi.org/10.1002/smj.2240

Kowalski, C. J. (1972). On the effects of non-normality on the distribution of the sample product–moment correlation coefficient. *Applied Statistics*, *21*(1), 1–12. https://doi.org/10.2307/2346598

Kwon, K., & Rupp, D. E. (2013). High-performer turnover and firm performance: The moderating role of human capital investment and firm reputation. *Journal of Organizational Behavior*, *34*(1), 129–150. https://doi.org/10.1002/job.1804

Landy, F. J., Farr, J. L., & Jacobs, R. R. (1982). Utility concepts in performance measurement. *Organizational Behavior and Human Performance*, *30*(1), 15–40. https://doi.org/10.1016/0030-5073(82)90232-X

Li, Y., Li, N., Li, C., & Li, J. (2020). The boon and bane of creative "stars": A social network exploration of how and when team creativity is (and is not) driven by a star teammate. *Academy of Management Journal*, *63*(2), 613–635. https://doi.org/10.5465/amj.2018.0283

Liu, K. (2014). Human capital, social collaboration, and patent renewal within USpharmaceutical firms. *Journal of Management*, *40*(2), 616–636. https://doi.org/10.1177/0149206313511117

Macnamara, B. N., Moreau, D., & Hambrick, D. Z. (2016). The relationship between deliberate practice and performance in sports: A meta-analysis. *Perspectives on Psychological Science*, *11*(3), 333–350. https://doi.org/10.1177/1745691616635591

Marchese, M. C., & Muchinsky, P. M. (1993). The validity of the employment interview: A meta-analysis. *International Journal of Selection and Assessment*, *1*(1), 18–26. https://doi.org/10.1111/j.1468-2389.1993.tb00080.x

Mariotti, J. L. (2008). *The complexity crisis*. Simon & Schuster.

Marsh, L. C., & Cormier, D. R. (2001). *Spline regression models*. Sage.

Massey, C., & Thaler, R. H. (2013). The loser's curse: Decision making and market efficiency in the National Football League draft *Management Science, 59*(7), 1479–1495. https://doi.org/10.1287/mnsc.1120.1657

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*(4), 599–616. https://doi.org/10.1037/0021-9010.79.4.599

Minbaeva, D., & Collings, D. G. (2013). Seven myths of global talent management. *The International Journal of Human Resource Management*, *24*(9), 1762–1776. https://doi.org/10.1080/09585192.2013.777539

Morris, S. S., Alvarez, S. A., & Barney, J. B. (2021). Dancing with the stars: The practical value of theory in managing star employees. *Academy of Management Perspectives*, *35*(2), 248–264. https://doi.org/10.5465/amp.2017.0223

Naylor, J. C., & Shine, L. C. (1965). A table for determining the increase in mean criterion score obtained by using a selection device. *Journal of Industrial Psychology*, *3*(2), 33–42.

O'Boyle, E., & Aguinis, H. (2012). The best and the rest: Revisiting the norm of normality of individual performance. *Personnel Psychology*, *65*(1), 79–119. https://doi.org/10.1111/j.1744-6570.2011.01239.x

O'Boyle, E., & Kroska, S. (2017). Star performers. *The Oxford handbook of talent management* (pp. 43–65). Oxford University Press.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science*, *58*(6), 1122–1140. https://doi.org/10.1287/mnsc.1110.1470

Ock, J., & Oswald, F. L. (2018). The utility of personnel selection decisions: Comparing compensatory and multiple-hurdle selection models. *Journal of Personnel Psychology*, *17*(4), 172–182. https://doi.org/10.1027/1866-5888/a000205

Oprea, B. T., Barzin, L., Vîrgă, D., Iliescu, D., & Rusu, A. (2019). Effectiveness of job crafting interventions: A meta-analysis and utility analysis. *European Journal of Work and Organizational Psychology*, *28*(6), 723–741. https://doi.org/10.1080/1359432X.2019.1646728

Overbeck, J. R., Correll, J. C., & Park, B. (2005). The internal sorting process of group status: The problem of too many stars. In M. Thomas-Hunt, E. Mannix, & M. A. Neale (Eds.), *Research on managing groups and teams* (Vol. *7*, pp. 169–199). Elsevier Press.

Prato, M., & Ferraro, F. (2018). Starstruck: How hiring high-status employees affects incumbents' performance. *Organization Science*, *29*(5), 755–774. https://doi.org/10.1287/orsc.2018.1204

Sackett, P. R., & Yang, H. (2000). Correction for range restriction: An expanded typology. *Journal of Applied Psychology*, *85*(1), 112–118. https://doi.org/10.1037/0021-9010.85.1.112

Saetre, A. S., & Van de Ven, A. H. (2021). Generating theory by abduction. *Academy of Management Review*. Advance online publication. https://doi.org/10.5465/amr.2019.0233

Schmidt, F. L., & Hoffmann, B. (1973). Empirical comparison of three methods of assessing utility of a selection device. *Industrial and Organizational Psychology*, *1*, 13–22.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work–force productivity. *Journal of Applied Psychology*, *64*(6), 609–626. https://doi.org/10.1037/0021-9010.64.6.609

Scullen, S. E., & Meyer, B. C. (2014). More applicants or more applications per applicant? A big question when pools are small. *Journal of Management*, *40*(6), 1675–1699. https://doi.org/10.1177/0149206312438774

Seijts, G., Espinoza, J. A., & Carswell, J. (2020). Utility analysis of character assessment in employee placement. *Leadership & Organization Development Journal*, *41*(5), 703–720. https://doi.org/10.1108/LODJ-07-2019-0314

Simonsohn, U. (2018). Two lines: A valid alternative to the invalid testing of U-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science*, *1*(4), 538–555. https://doi.org/10.1177/2515245918805755

Sturman, M. C. (2000). Implications of utility analysis adjustments for estimates of human resource intervention value. *Journal of Management*, *26*(2), 281–299. https://doi.org/10.1177/014920630002600206

Sturman, M. C., Trevor, C. O., Boudreau, J. W., & Gerhart, B. (2003). Is it worth it to win the talent war? Evaluating the utility of performance-based pay. *Personnel Psychology*, *56*(4), 997–1035. https://doi.org/10.1111/j.1744-6570.2003.tb00248.x

Swaab, R. I., Schaerer, M., Anicich, E. M., Ronay, R., & Galinsky, A. D. (2014). The too-much-talent effect: Team interdependence determines when more talent is too much or not enough. *Psychological Science*, *25*(8), 1581–1591. https://doi.org/10.1177/0956797614537280

Taylor, E. C., & Bendickson, J. S. (2021). Star performers, unit performance and unit turnover: A constructive replication. *Human Resource Management Journal*. Advance online publication. https://doi.org/10.1111/1748-8583.12336

Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory*. Transaction Publishers.

Turetsky, A. (2017). Stars and skews rise and fall: A mathematical multilevel adventure. *Academy of Management Proceedings*, *2017*(1), 10158. https://doi.org/10.5465/AMBPP.2017.10158abstract

Tzabbar, D., & Kehoe, R. R. (2014). Can opportunity emerge from disarray? An examination of exploration and exploitation following star scientist turnover. *Journal of Management*, *40*(2), 449–482. https://doi.org/10.1177/0149206313513613

Vaci, N., Cocić, D., Gula, B., & Bilalić, M. (2019). Large data and Bayesian modeling—Aging curves of NBA players. *Behavior Research Methods*, *51*(4), 1544–1564. https://doi.org/10.3758/s13428-018-1183-8

Van Iddekinge, C. H., Aguinis, H., LeBreton, J. M., Mackey, J. D., & DeOrtentiis, P. S. (2021). Assessing and interpreting interaction effects: A reply to Vancouver, Carlson, Dhanani, and Colton (2021). *Journal of Applied Psychology*, *106*(3), 476–488. https://doi.org/10.1037/apl0000883

Veksler, D. (2010). *Some lesser known truths about programming*. Retrieved from https://bit.ly/3hwQMsY

Weekley, J. A., Frank, B., O'Connor, E. J., & Peters, L. H. (1985). A comparison of three methods of estimating the standard deviation of performance in dollars. *Journal of Applied Psychology*, *70*(1), 122–126. https://doi.org/10.1037/0021-9010.70.1.122

Whyte, G., & Latham, G. (1997). The futility of utility analysis revisited: When even an expert fails. *Personnel Psychology*, *50*(3), 601–610. https://doi.org/10.1111/j.1744-6570.1997.tb00705.x

Wroten, S. P. (1984). Overcoming the futilities of utility applications: Measures, models, and management [Paper presentation]. Presentation, the Annual Meeting of the American Psychological Association, Toronto.

Zucker, L. G., & Darby, M. R. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(23), 12709–12716. https://doi.org/10.1073/pnas.93.23.12709