BUSINESS RESEARCH
BRQ
QUARTERLY

www.elsevier.es/brq

ELSEVIER
DOYMA

ARTICLE

Revisiting some ''established facts'' in the field of management☆

Herman Aguinis

Department of Management and Entrepreneurship, Kelley School of Business, Indiana University, 1309 E. 10th Street, Bloomington, IN 47405-1701, USA

Abstract   Although management is now becoming a mature scientific field and much theoretical and methodological progress has been made in the past few decades, management scholars are not immune to received doctrines and things we ''just know to be true.'' This article revisits an admittedly selected set of these ''established facts'' including how to deal with outliers, conducting field experiments with real entrepreneurs in real settings, the file-drawer problem in meta-analysis, and the distribution of individual performance. For each ''established fact,'' I describe its nature, the negative consequences associated with it, and best-practice recommendations in terms of how to address each. I hope this article will serve as a catalyst for future research challenging ''established facts'' in other substantive and methodological domains in the field of management.

1. Introduction

Management is now becoming a mature scientific field. Although its beginnings were heavily influenced by other disciplines such as psychology, economics, and sociology (Agarwal and Hoetker, 2007; Molloy et al., 2011), the field of management now develops its own theories (Colquitt and Zapata-Phelan, 2007; Shepherd and Sutcliffe, 2011). Moreover, the field also develops its own methodological approaches mainly described in the Academy of Management sponsored journal *Organizational Research Methods*. In addition, although the field of management has become increasingly specialized, as indicated by groups of scholars who focus mainly on the individual and team levels of analysis (e.g., organizational behavior, human resource management) and those who focus on the firm and industry levels of analysis (e.g., business policy and strategy, entrepreneurship), there is now a trend toward the development of more comprehensive and integrative theories that address organizational phenomena from multiple levels of analysis (e.g., Aguinis et al., 2011a; Foss, 2010, 2011; Van de Ven and Lifschitz, 2013). Given the progress

attained over the past few decades, the evidence-based management movement now offers important theory-based insights that can be used to improve management practice (Rousseau, 2012). In short, as we approach the 25th anniversary of the foundation of the of the Spanish Asociación Científica de Economía y Dirección de la Empresa (ACEDE) in the year 2015, we can conclude that much progress has been made since the publication of Gordon and Howell's (1959) report sponsored by the Carnegie Ford Foundation scolding business schools for their lack of scholarly rigor.

2. ''Established facts'' in the field of management: facts or urban legends?

Former Academy of Management President Bill Starbuck asserted that ''professors of management are people of superior abilities...'' (Barnett, 2007: 126). However, in spite of the scientific progress made by the field of management and similar to the general population, management scholars are not immune to received doctrines and things we ''just know to be true.'' In many cases, these issues are ''taught in undergraduate and graduate classes, enforced by gatekeepers (e.g., grant panels, reviewers, editors, dissertation committee members), discussed among colleagues, and otherwise passed along among pliers of the trade far and wide and from generation to generation'' (Lance, 2011: 281). Moreover, these ''established facts'' have in many cases reached the status of myth and urban legends, similar to those about alligators living in the sewage system of the city of New York, or about King Juan Carlos I of Spain riding a motorcycle and helping a stranded motorist (Brunvand, 2012).

The existence of these myths and urban legends is expected as part of a scientific field's growing pains (Lance and Vandenberg, 2009). Moreover, the reason for their existence is that there are kernels of truth underlying each of these ''established facts.'' However, in all cases, the kernels of truth have been forgotten, exaggerated, or somehow twisted. Many of us have been at the receiving end of these ''established facts'' when a journal reviewer, dissertation committee members, or professor in a doctoral seminar has indicated that, for example, we should implement a particular methodological procedure but the rationale is not fully explicated. Admittedly, many of us have also been at the giving end of these ''established facts'' in conversations with peers and doctoral students, and also in our roles of journal reviewers. These issues include both substantive to methodological topics and range from micro- to macro-level topics. Next, I revisit an admittedly selected set of these ''established facts'' by explaining their nature, the negative consequences resulting from each, and best-practice recommendations regarding how to address each. As a preview, Table 1 includes a summary of the issues addressed in the remainder of this article. The ''established facts'' refer to outliers being regarded as data problems that must be fixed, the impossibility of conducting field experiments with real entrepreneurs in real settings, the belief that the file-drawer problem biases meta-analytic conclusions, and the belief that individual performance is best modeled using a normal distribution.

3. Outliers are data errors that must be fixed

Outliers are data points that deviate markedly from others. Thus, an outlier can be an individual, team, firm, or any other unit. The existence of outliers is one of the most enduring and pervasive methodological challenges in management research because their presence often has an important and disproportionate impact on substantive conclusions regarding relationships among variables. The important impact of outliers on substantive conclusions has been noted in many management subfields, ranging from organizational behavior and human resource management (Orr et al., 1991) to strategy (e.g., Hitt et al., 1998).

Aguinis et al. (2013) conducted a literature review on outliers involving all articles published between 1991 through 2010 in *Academy of Management Journal*, *Journal of Applied Psychology*, *Personnel Psychology*, *Strategic Management Journal*, *Journal of Management*, and *Administrative Science Quarterly*. As part of their review, they identified 232 articles that mentioned the issue of outliers. One of the main conclusions of this review was that management scholars view outliers as ''problems'' that must be ''fixed.'' Usually, this is done by removing particular cases from the analyses. Moreover, Aguinis et al.'s (2013) review also uncovered that it is common for management researchers to either be vague or not transparent in how outliers are defined and in how a particular outlier identification technique was chosen and used. In sum, there seems to be an ''established fact'' that outliers are a nuisance and must be removed – and the particular process used to do so is often not reported openly and transparently.

The current state of the science regarding how management scholars address outliers has important negative implications (Aguinis and Joo, in press). First, deleting outliers from a dataset simply because they are distant from other units can result in large opportunity costs in terms of uncovering interesting relationships. In other words, some outliers may not be problems that must be fixed; rather, they may be interesting observations worth studying further. Second, lack of transparency in how outliers are defined, identified, and handled diminishes the potential replicability of substantive results, which is required for the advancement of science (Brutus et al., 2013).

So, what should management researchers do regarding outliers? Aguinis et al. (2013) offered two general guidelines. First, choices and procedures regarding the treatment (i.e., definition, identification, and handling) of outliers should be described in detail to ensure transparency – including a rationale for the particular procedures that have been implemented. The second principle is that researchers should clearly and explicitly acknowledge the type of outlier in which they are interested, and then use an identification technique that is congruent with the outlier definition.

In addition, Aguinis et al. (2013) offered more specific recommendation on a sequential process for defining, identifying, and handling three different types of outliers. The first category consists of error outliers, or data points that lie at a distance from other data points because they are the result of inaccuracies. If error outliers are found, the recommendation is to either adjust the data points to their correct values or remove such observations from the dataset. In addition, it is necessary to explain in detail the reasoning

Table 1 Summary of selected ''established facts'' in the field of management, their negative consequences, and best-practice recommendations.

Established facts	Negative consequences of established facts	Best-practice recommendations
Outliers are data errors that must be fixed	Deleting an observation from a dataset simply because it is distant from other units can result in large opportunity costs in terms of uncovering interesting relationships Lack of transparency in how outliers are defined, identified, and handled decreases the potential replicability of substantive results, which is required for the advancement of science	Offer a detailed description of choices and procedures regarding the treatment (i.e., definition, identification, and handling) of error, interesting, and influential outliers, including a rationale for the particular procedures that have been implemented Clearly and explicitly acknowledge the type of outlier in which there is an interest and then use an identification technique that is congruent with the outlier definition Include information on outliers in a separate section titled ''Outlier Detection and Management''
It is virtually impossible to conduct experimental research using real entrepreneurs in real settings	External validity (i.e., degree of confidence in the generalizability of results) is maximized at the expense of internal validity (i.e., degree of confidence in the nature of causal relationships) by implementing passive observation designs (e.g., no random assignment of participants to conditions, no manipulation of variables) Internal validity is maximized at the expense of internal validity by implementing experimental designs often involving convenience samples in artificial settings	Use of eLancing as a tool to conduct field experiments – experiments involving real entrepreneurs in naturally occurring settings Follow the following three steps: (1) sign up for a simple job using any of the eLancing marketplaces to experience eLancing from the perspective of a potential research participant; (2) conduct a simple experiment that replicates a study that has already been done in the past; and (3) design and conduct an original experiment addressing substantive hypotheses and questions
The file-drawer problem biases meta-analytic conclusions	There is not much that meta-analysts can do in terms of avoiding the file drawer problem because they do not have access to the original data but only to the supposedly upwardly biased resulting effect-size estimates The file drawer problem is believed to have an important upwardly biasing effect on the resulting effect-size estimates which, in turn, lead to theory derailments and practices that may not be as effective as expected The file drawer problem puts into question the credibility and trustworthiness of all meta-analytic reviews	Because the file drawer problem does not seem to pose a serious threat to meta-analytic conclusions based on non-experimental research (i.e., usually using correlation coefficients as the effect size metric), concerns, caveats, and cautionary notes often included in published meta-analytic reviews based on the file drawer problem do not seem to be warranted Because the file drawer does not seem to pose a serious threat to meta-analytic conclusions based on non-experimental research, there is no need to ''fix'' an upward bias that does not exist
Individual performance is normally distributed	Deviations from normality are assumed to be errors and subsequent adjustments are made (e.g., dropping outliers, data transformation procedures) in order to make the sample ''better reflect'' the ''true'' underlying normal curve Typical data-analytic techniques that assume normality are unable to make accurate predictions regarding non-normally distributed outcomes (i.e., individual performance)	Rather than assuming that it does, empirically assess whether individual performance follows a non-normal distribution in the sample in hand Do not force normality through outlier manipulation or case deletion Consider implementing methodological techniques that properly and accurately estimate models where the outcome follows a power law (e.g., Poisson processes, agent based modeling, Bayesian analysis)

behind the classification of the observation as an error outlier. For example, was it a coding error? A data entry error? The second category represents interesting outliers, which are accurate data points that lie at a distance from other data points and may contain valuable or unexpected knowledge. If interesting outliers are found, the recommendation is to study them further – using both quantitative and qualitative approaches. The third category refers to influential outliers, which are accurate data points that lie at a distance from other data points, are not error or interesting outliers, and also affect substantive conclusions. Recommendations for how to handle influential outliers include (a) model respecification (e.g., adding an interaction or quadratic term to a regression model), (b) case deletion, and (c) using robust statistical techniques. Regardless of the approaches chosen, it is important to report results with and without the chosen handling technique and to provide an explanation for any differences in the results. Moreover, information on how each of the three types of outliers has been addressed can be included in a section titled ''Outlier Detection and Management.''

In sum, outliers, although typically not acknowledged or discussed openly in published journal articles, are pervasive in most empirical management research, ranging from the micro to the macro level of analysis and spanning all types of methodological and statistical approaches. The latest research regarding outliers challenges the established fact that outliers are errors that must be fixed. Accordingly, future empirical research should follow a standardized and systematic sequence of defining, identifying, and handling error, interesting, and influential outliers and providing information on the precise process that was followed to enhance transparency and replicability.

4. It is not possible to conduct experimental research using real entrepreneurs in real settings

Understanding the nature of causal relationships is at the heart of the scientific enterprise (Aguinis and Edwards, 2014). Making causal inferences from management research is not only important in terms of the field's theoretical advancement, but also in terms of deriving evidence-based recommendations for practice (Aguinis and Vandenberg, 2014). Stated differently, it is difficult to give sound evidence-based advice to practitioners without an understanding of underlying causal relationships among variables. Given the importance of entrepreneurship as an economic engine in the 21st century, understanding causal effects involved in why and when entrepreneurs engage in certain behaviors and make certain decisions is of particular interest both for research and practice (Short et al., 2010).

Aguinis and Lawal (2012) conducted a review of the 175 empirical articles published in *Journal of Business Venturing* (JBV) from January 2005 through November 2010. The goal of their review focusing on one of the most impactful journals in the field of entrepreneurship was to identify the relative frequency with which researchers refer to various methodological challenges. One of the most frequently mentioned challenges was authors' lack of confidence regarding the precise nature of causal relationships.

This result is not surprising because Aguinis and Lawal (2012) reported that 74.9% of studies published in JBV used non-experimental (i.e., passive observation) designs. Moreover, this result is also not surprising given that the same challenge has been identified in other literature reviews addressing methodological issues (e.g., Aguinis et al., 2009; Podsakoff and Dalton, 1987; Scandura and Williams, 2000). Non-experimental designs do not involve random assignment of participants to conditions or the manipulation of variables and, hence, results are ambiguous in terms of the precise nature and direction of effects (Gregoire et al., 2010). Moreover, when experimental designs are used, samples often do not include real entrepreneurs, but students and other types of convenience samples, and the study is often conducted in artificial settings (e.g., university laboratory or vignette study describing a hypothetical situation).

It seems that there is an ''established fact'' involving an unavoidable trade-off between external (i.e., generalizability of results) and internal (i.e., confidence in causal relationships) validity (Aguinis and Edwards, 2014). In other words, researchers seem to believe that they are put in an often inescapable catch-22 situation in which they can either use samples of real entrepreneurs in real settings, but cannot manipulate variables and therefore confidence about causality is lacking, or use samples of students and other convenience samples in laboratory settings, yielding results that are stronger about causal inferences but weaker in terms of generalizability. That is, experimental research often puts into question the external validity of results because it is not possible to know whether participants would behave in the same way in a natural as compared to an artificial (i.e., laboratory) or simulated setting.

Aguinis and Lawal (2012) offered a possible solution for the ''established fact'' that researchers can only maximize internal validity (i.e., confidence regarding causal relationships) at the expense of external validity (i.e., confidence regarding the generalizability of results) and vice versa. Specifically, they proposed the use of eLancing as a tool to conduct *field experiments* – experiments involving real entrepreneurs in naturally occurring settings.

eLancing, or Internet freelancing, is a fairly novel type of work arrangement that uses websites, called ''marketplaces,'' where individuals interested in being hired and clients looking for individuals to perform some type of work meet (Aguinis and Lawal, 2012). Examples of eLancing marketplaces include eLance.com, freelancer.com, guru.com, Amazon Mechanical Turk (mturk.com), oDesk.com, and microworkers.com, among many others. eLancing is producing a revolution in how work is done and regarding entrepreneurial activities around the world because it allows individuals from any location around the world to sign up and complete work for a client who literally can also be anywhere in the world (Aguinis and Lawal, 2013). There are entrepreneurs who are turning to eLancing marketplaces to acquire resources that they may not be able to access otherwise. There are also entrepreneurs and aspiring entrepreneurs who offer their services in a number of different arenas – eLancing being one of them – and use eLancing to raise funds.

Aguinis and Lawal (2012) described how to use eLancing to conduct field experiments. Specifically, a researcher can issue a call for work (i.e., which is actually a field

experiment) and require, for example, that study participants be from a certain region, industry, and with specific experience and educational characteristics (including demographic characteristics). eLancing allows researchers to manipulate independent variables (e.g., number and type of team members) and then measure the effect of that precise manipulation on key outcome variables (e.g., team performance). Moreover, by implementing random assignment of individuals to conditions (e.g., small versus large entrepreneurial teams, entrepreneurial teams with varying numbers of marketing experts), researchers can draw conclusions about the direction and strength of causal relationships.

From a practical perspective, eLancing allows researchers to recruit study participants 24/7 from around the world. Moreover, the cost of recruiting and compensating study participants can be as low as a few cents of a U.S. dollar per task. For example, based on jobs posted on elance.com on July 14, 2011, Aguinis and Lawal (2012) reported that there were 1200 audio recording tasks available for a pay ranging from $0.01 to $20 per task.

Aguinis and Lawal (2012) offered the following recommendations on how to use eLancing to conduct field experiments using real participants in natural settings. The first recommendation is for researchers to sign up for a simple job using any of the eLancing marketplaces to experience eLancing from the perspective of a potential research participant. Doing this will allow researchers to understand how a participant signs up for a job (i.e., potential field experiment), what are the documents involved (e.g., agreement to conduct certain work by a certain time), and how performance management and compensation systems are implemented. The second recommendation is that, this time from the perspective of an eLancing client (i.e., experimenter), a researcher can conduct a simple experiment that replicates a study that has already been done in the past. This second step will allow researchers to become familiar with the eLancing environment, post a call for work (i.e., experiment), manipulate variables (e.g., change the nature of the task, change the composition of teams, change the amount and type of information and knowledge given to various entrepreneurial teams or team members), and how to collect data (e.g., how to gather data using online surveys, chat rooms, and other online data collection tools available). After completing these two initial steps, most researchers will be in a position to design and conduct an original field experiment addressing substantive hypotheses and questions.

In sum, there is an ''established fact'' that researchers often face an inescapable dilemma in which they are forced to engage in an unavoidable trade-off between external validity (i.e., degree of confidence regarding the generalizability of results) versus internal validity (i.e., degree of confidence regarding causal relationships). Recent research challenges this ''established fact.'' Specifically, researchers can take advantage of eLancing to conduct field experiments involving real participants in real settings, thereby minimizing the external versus internal validity trade-off. Moreover, eLancing offers researchers the possibility of avoiding this trade-off by conducting field experiments in a way that are logistically and practically feasible. Finally, although the use of eLancing has been described in the context of

entrepreneurship research, this technological enablement can be used to conduct research in marketing (e.g., to study how various manipulations affect consumer behavior), organizational behavior and human resource management (e.g., to study how various types of compensation systems affect subsequent motivation and performance), and many other research domains.

5. The file-drawer problem biases meta-analytic conclusions

Meta-analysis is currently considered the most powerful and informative methodological approach for conducting a systematic literature review (Aguinis et al., 2011b). A meta-analysis has two fundamental goals. First, it aims at understanding the nature and size of a relationship across a large number primary-level studies. Second, it aims at understanding the variability of a relationship across primary-level studies as well as the factors that explain this variability – what are called moderating effects (Aguinis et al., 2011c).

Meta-analysis has been used to understand key issues in strategic management studies such as the relationship between firm resources and performance (e.g., Crook et al., 2008) as well as organizational behavior and human resource management such as the relationship between individual personality and leadership (e.g., Judge et al., 2002). Dalton et al. (2012) conducted a literature search and reported that for the period 1980–2010, there are 5,183 articles with the expression ''meta-analysis'' or its derivatives in the PsycINFO database, 10,905 in the EBSCO Academic/Business Source Premier database, and 15,627 in the MedLine database. It is clear, then, that meta-analysis is a widely used approach to synthesize a body of empirical research not only in management, but also in many other scientific fields.

In spite of its widespread use and influence, meta-analyses are often viewed with suspicion due to the so-called ''file drawer problem'' (Greenwald, 1975; Rosenthal, 1979). There seems to be an ''established fact'' that statistically non-significant results are less likely to be published and, hence, less likely to be included in meta-analytic reviews, thereby resulting in an upwardly biased sample of primary-level effect-size estimates and upwardly biased meta-analytically derived summary effect sizes (Dalton et al., 2012). Moreover, there seems to be an established fact that the file drawer problem is an important cause for concern that compromises meta-analytic results and substantive conclusions (e.g., Mone et al., 1996; Viswesvaran et al., 1993). The file drawer problem seems to be an established fact not only in the field of management, but in other fields as well. For example, in neuroscience, Fiedler (2011: 164) issued the warning that ''a file-drawer bias...facilitates the selective publication of strong correlations while reducing the visibility of weak research outcomes'' (p. 167). Moreover, he argued that ''[w]hat we are dealing with here is a general methodological problem that has intrigued critical scientists under many different labels: the file-drawer bias in publication'' (Fiedler, 2011: 164).

As described by Dalton et al. (2012), there are several ''established facts'' directly related to the file drawer

problem. First, the file drawer problem is believed to be pervasive and almost unavoidable. Stated differently, the reason for the existence of the file drawer problem is that it is reflected in effect-size estimates reported in primary-level research and, because meta-analysts do not have access to the original data but only to the supposedly upwardly biased resulting effect-size estimates, the file drawer is an insurmountable problem for a meta-analyst. Second, the file drawer problem is believed to have an important upwardly biasing effect on the resulting effect-size estimates. Moreover, biased effect-size estimates lead to theory derailments and practices that may not be as effective as expected. For example, practitioners may implement selection, training, and other interventions believing incorrectly that the effectiveness of such interventions will be greater than they actually are due to an assumed overestimation of meta-analytically derived effect sizes. This overestimation and incorrect belief are supposedly caused by the file drawer problem.

Dalton et al. (2012) offered and implemented a novel protocol involving a five-study research program to revisit established facts around the file drawer problem. To reiterate, the file drawer problem rests on the assumption that research with statistically non-significant results is less likely to be published and that meta-analyses do not include a random sample of effect size estimates, but, rather, only an upwardly biased sample. However, correlation coefficients used as input in meta-analyses are not obtained from studies that include a single hypothesis or report a single correlation coefficient only. On the contrary, meta-analyses published in the field of management use primary-level effect size estimates usually obtained from published articles' full correlation matrix (note that this is not necessarily the case in other fields that rely more heavily on experimental designs such as the biomedical sciences, where ds rather than rs are used as the effect size estimate). Accordingly, Dalton et al. (2012) revisited established facts around the file drawer problem based on an assessment of correlation matrices of published and non-published research. Thus, the issue investigated by Dalton et al. (2012) was whether the correlations reported in non-published studies are fundamentally different in terms of statistical significance and magnitude compared to those reported in published studies. If differences are found, then concerns about the file drawer problem would seem warranted.

Dalton et al. (2012) conducted five studies to address this issue. In Study 1, they examined 37,970 correlations included in 403 matrices published in *Academy of Management Journal* (AMJ), *Journal of Applied Psychology* (JAP), and *Personnel Psychology* (PPsych) between 1985 and 2009 and found that 46.81% of those correlations are not statistically significant. In Study 2, they examined 6,935 correlations used as input in 51 meta-analyses published in AMJ, JAP, PPsych, and elsewhere between 1982 and 2009 and found that 44.31% of those correlations are not statistically significant. In Study 3, they investigated 13,943 correlations reported in 167 matrices in non-published manuscripts and found that 45.45% of those correlations are not statistically significant. In Study 4, they examined 20,860 correlations reported in 217 matrices in doctoral dissertations and found that 50.78% of those correlations are not statistically significant. Finally, in Study 5, Dalton et al. (2012) compared

the average magnitude of a sample of 1,002 correlations from Study 1 (published articles) versus 1,224 from Study 4 (dissertations) and found that they were virtually identical (i.e., 0.2270 and 0.2279, respectively). Taken together, Dalton et al.'s (2012) results indicated that between 40% and 50% of primary-level effect sizes in published and non-published sources that are potentially included and actually included in meta-analytic reviews are not statistically significant. Moreover, there was a very high degree of similarity in the magnitude of correlations reported in primary studies compared to correlations reported in doctoral dissertations.

In sum, Dalton et al.'s (2012) work has important implications for meta-analysis in particular and empirical research in general because it challenges the ''established fact'' that the file drawer problem poses a serious threat to the validity of meta-analytically derived conclusions. Specifically, concerns, caveats, and cautionary notes often included in published meta-analytic reviews based on the file drawer problem do not seem to be warranted. In other words, Dalton et al. (2012) found no evidence to support the long-lamented belief that the file drawer problem produces an upward bias in meta-analytically derived effect sizes. Consequently, the methodological practice of estimating the extent to which results are not vulnerable to the file drawer problem may be eliminated in many cases. Stated differently, if, as Dalton et al.'s (2012) results suggest, the file drawer problem is in fact not producing a bias, then there is no need to ''fix'' an upward bias that does not exist.

6. Individual performance is normally distributed

Most theories in the field of management build upon, directly or indirectly, the output of individual workers. For example, theories about human capital, turnover, compensation, downsizing, leadership, knowledge creation and dissemination, top management teams, teamwork, corporate entrepreneurship, and microfoundations of strategy involve the performance of individual workers (e.g., Felin and Ferterly, 2007). In fact, a central goal of organizational behavior and human resource management theories is to understand and predict individual performance. From the perspective of strategic management studies, a better understanding of individual performance is crucial in terms of making progress regarding microfoundations of strategy, which are the foundations of a field based on individual actions and interactions (Foss, 2010, 2011). For example Mollick (2012: 1001–1002) noted that the overwhelming focus on macro-level process variables in explaining firm performance, rather than compositional variables (i.e., workers), ''has prevented a thorough understanding of which individuals actually play a role in determining firm performance [and] to expect that not all variation among individuals contributes equally to explaining performance differences between firms.''

There is an ''established fact'' that individual performance clusters around a mean and then fans out into symmetrical tails. That is, individual performance is assumed to follow a normal distribution (Hull, 1928; Schmidt and Hunter, 1983; Tiffin, 1947). When individual performance data do not conform to the normal distribution, the

conclusion is that the error ''must'' lie within the sample, not the population. Subsequent adjustments are made (e.g., dropping outliers, data transformation procedures) in order to make the sample ''better reflect'' the ''true'' underlying normal curve. In addition, researchers often use data-analytic techniques that assume normality and are unable to make accurate predictions regarding non-normally distributed outcomes (i.e., individual performance) (Cascio and Aguinis, 2008).

In contrast to this established fact, there seem to be many illustrations of situations in which the performance distribution is not normal, but follows a power law. In contrast to normal distributions, power law distributions are typified by unstable means, infinite variance, and a greater proportion of extreme events. In other words, in contrast to a normal distribution, a power law distribution suggests that a small proportion of individuals accounts for a disproportionate amount of output.

There are numerous examples of situations that suggest the presence of star performers (i.e., outliers) that deny the appropriateness of modeling performance using a normal distribution (Aguinis and O'Boyle, in press). For example, since reassuming the role of Starbucks CEO in 2008, Howard Schultz has achieved a market capitalization of $33 billion, more than $11 billion annual sales, and net annual profits of $1.7 billion (Starbucks Corporation, 2012). While the U.S. economy is still struggling and the average growth of S&P 500 companies was −0.4 percent in 2011, Starbucks' share price increased by more than 40 percent. Demonstrating a similarly high degree of performance that cannot be modeled by using a normal distribution, thirty years earlier, a young Japanese programmer named Shigeru Miyamoto developed a bizarre game involving a gorilla throwing barrels for a near-bankrupt company named Nintendo. The success of *Donkey Kong* helped fund Nintendo's launch of a home gaming system where Miyamoto continued to work and develop some of the most successful franchises in gaming history, including *Mario Brothers* and the *Legend of Zelda* (Suellentrop, 2013). More recently, in Bangalore, India, dropout rates in the public school system were soaring due to students' malnutrition until an engineer named Shridhar Venkat overhauled the failing lunch program with a series of logistical and supply chain adaptations (Vedantam, 2012). Venkat's continued enhancements of the program have so significantly improved both children's health and school attendance that the Bangalore Public School System is now a Harvard Business School case study.

As noted by Aguinis and O'Boyle (in press), Schultz, Miyamoto, and Venkat typify star performers who consistently generate exorbitant output levels that influence the success or failure of their organizations and even society as a whole. Also, although it is likely that such star performers have existed throughout history, their presence is particularly noticeable across many industries and organizations that make up the twenty-first-century workplace and they occupy roles ranging from front-line workers to top management. Moreover, their addition can signal the rise of an organization and their departure can portend decline and even organizational death (Bedeian and Armenakis, 1998). These elite performers are not identified based on some bundle of traits or combination of ability and motivation. Rather, their presence is noticeable based on their output and what

makes them special is that their production is so clearly superior. Moreover, their presence challenges the notion that performance is normally distributed because such star performers – responsible for such a large proportion of the output – fall outside of a normal curve.

O'Boyle and Aguinis (2012) engaged in a systematic five-study research program to assess the relative fit of a normal compared to a power law distribution. Study 1 included the distribution of performance of 488,717 researchers who have produced 950,616 publications across 54 academic disciplines between January 2000 and June 2009. Study 2 involved the performance of 18,449 individuals in the entertainment industry, with performance rated by a large voting body or more objective performance measures such as the number of times an entertainer received an award, nomination, or some other indicator (e.g., Grammy nominations, New York Times best-selling list). Study 3 involved an examination of the performance of politicians – 42,745 candidates running for office in 42 different types of elections in Australia, Canada, Denmark, Estonia, Finland, Holland, Ireland, New Zealand, the United Kingdom, and the U.S. Study 4 included 25,283 athletes in collegiate and professional sports (e.g., number of goals during the 2009–2010 season for teams in the English Premier League, homeruns and strikeouts in United States' Major League Baseball, points in United States' National Hockey League). Finally, Study 5 investigated the negative performance of 57,246 athletes (e.g., English Premier League yellow cards, National Basketball Association career turnovers, and Major League Baseball first-base errors). In short, results based on five separate studies and involving 198 samples including 632,599 researchers, entertainers, politicians, and amateur and professional athletes were highly consistent. Of a total of 198 samples, 186 (93.94%) followed a Paretian (i.e., power law) distribution more closely than a Gaussian (i.e., normal) distribution.

Revisiting the norm of normality for the distribution of individual performance has important implications for theory and practice. These implications affect numerous research and practice domains such as performance measurement and performance management, preemployment testing, leadership, human capital, turnover, compensation, downsizing, leadership, teamwork, corporate entrepreneurship, and microfoundations of strategy (Aguinis and O'Boyle, in press; O'Boyle and Aguinis, 2012). For example, a power law perspective regarding the performance distribution provides a means to bridging micro-macro domains by reconciling the human capital paradox in microfoundations of strategy research of how plentiful and average workers at the individual level metamorphosize into rare and inimitable human capital at the firm level. As a second illustration, a power law perspective suggests that leadership theories may need to abandon the assumption that in order to increase overall productivity, all workers must improve by the same degree (i.e., normality assumption). Regarding implications for practice, a power law perspective suggests that firms experiencing financial difficulties should pay special attention to star performers as budget cuts, downsizing, and other cost cutting measures may signal that the organization is in decline, leading to preemptive star departure. In addition, star departure can create a downward spiral of production when ''marplots and meddlers'' deliberately

replace stars with inferior workers (Bedeian and Armenakis, 1998).

In sum, there is an ''established fact'' that individual performance follows a normal distribution. However, recent research challenges this ''established fact'' and suggests that individual performance follows, in many cases, a power law distribution. The presence of star performers and their disproportionate contribution in terms of many types of outcomes provides evidence in this regard. As a consequence of revisiting this established fact, many management theories firmly rooted in the manufacturing sector, corporate hierarchy, and the human capital of the ''necessary many'' may not apply to today's workplace that operates globally and is driven by the ''vital few.'' So, rather than assuming that it does, it is now necessary to empirically assess whether individual performance follows a non-normal distribution in the sample in hand. Moreover, the common practice of forcing normality through outlier manipulation or case deletion should be avoided. Rather, in many cases it will be useful to implement methodological techniques that properly and accurately estimate models where the outcome follows a power law (e.g., Poisson processes, agent based modeling, Bayesian analysis) (Kruschke et al., 2012).

7. Concluding remarks

Management is now a mature scientific field. Much progress has been made in terms of theory as well as methodological developments since the publication of Gordon and Howell's (1959) report admonishing business schools for their lack of academic rigor. Thus, management researchers should be sufficiently confident in the scientific status of the field to engage in the task of challenging prevailing theories and methodological practices that may have been established over time and turned into myths and urban legends. As noted by philosopher of science Karl Popper (1963: 36), ''the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.'' I hope this article will serve as a catalyst for future research challenging ''established facts'' in other substantive and methodological domains in the field of management.

References

Agarwal, R., Hoetker, G., 2007. A Faustian bargain? The growth of management and its relationship with related disciplines. Academy of Management Journal 50, 1304–1322.

Aguinis, H., Boyd, B.K., Pierce, C.A., Short, J.C., 2011a. Walking new avenues in management research methods and theories: bridging micro and macro domains. Journal of Management 37, 395–403.

Aguinis, H., Dalton, D.R., Bosco, F.A., Pierce, C.A., Dalton, C.M., 2011b. Meta-analytic choices and judgment calls: implications for theory building and testing, obtained effect sizes, and scholarly impact. Journal of Management 37, 5–38.

Aguinis, H., Edwards, J.R., 2014. Methodological wishes for the next decade and how to make wishes come true. Journal of Management Studies 51, 143–174.

Aguinis, H., Gottfredson, R.K., Joo, H., 2013. Best-practice recommendations for defining, identifying, and handling outliers. Organizational Research Methods 16, 270–301.

Aguinis, H., Joo, H., 2013. Debunking myths and urban legends about how to identify influential outliers. In: Lance, C.E., Vandenberg, R.J. (Eds.), Statistical and Methodological Myths and Urban Legends: Received Doctrine, Verity, and Fable in the Organizational and Social Sciences. , 2nd ed. Routledge, New York (in press).

Aguinis, H., Lawal, S.O., 2012. Conducting field experiments using eLancing's natural environment. Journal of Business Venturing 27, 493–505.

Aguinis, H., Lawal, S.O., 2013. eLancing: a review and research agenda for bridging the science–practice gap. Human Resource Management Review 23, 6–17.

Aguinis, H., O'Boyle, E., 2013. Star performers in twenty-first-century organizations. Personnel Psychology, http://dx.doi.org/10.1111/peps.12054 (in press).

Aguinis, H., Pierce, C.A., Bosco, F.A., Dalton, D.R., Dalton, C.M., 2011c. Debunking myths and urban legends about meta-analysis. Organizational Research Methods 14, 306–331.

Aguinis, H., Pierce, C.A., Bosco, F.A., Muslin, I.S., 2009. First decade of organizational research methods: trends in design, measurement, and data-analysis topics. Organizational Research Methods 12, 69–112.

Aguinis, H., Vandenberg, R.J., 2014. An ounce of prevention is worth a pound of cure: Improving research quality before data collection. Annual Review of Organizational Psychology and Organizational Behavior 1, 1–27.

Barnett, M.L., 2007. (Un)learning and (mis)education through the eyes of Bill Starbuck: an interview with Pandora's playmate. Academy of Management Learning and Education 6, 114–127.

Bedeian, A.G., Armenakis, A.A., 1998. The cesspool syndrome: How dreck floats to the top of declining organizations. Academy of Management Executive 12, 58–67.

Brunvand, J.H., 2012. Encyclopedia of Urban Legends Updated and Expanded Edition. ABC-CLIO, Santa Barbara, CA.

Brutus, S., Aguinis, H., Wassmer, U., 2013. Self-reported limitations and future directions in scholarly reports: analysis and recommendations. Journal of Management 39, 48–75.

Cascio, W.F., Aguinis, H., 2008. Staffing twenty-first-century organizations. Academy of Management Annals 2, 133–165.

Colquitt, J.A., Zapata-Phelan, C.P., 2007. Trends in theory building and theory testing: a five-decade study of the Academy of Management Journal. Academy of Management Journal 50, 1281–1303.

Crook, T.R., Ketchen, D.J., Combs, J.G., Todd, S.Y., 2008. Strategic resources and performance: a meta-analysis. Strategic Management Journal 29, 1141–1154.

Dalton, D.R., Aguinis, H., Dalton, C.A., Bosco, F.A., Pierce, C.A., 2012. Revisiting the file drawer problem in meta-analysis: an empirical assessment of published and non-published correlation matrices. Personnel Psychology 65, 221–249.

Felin, Ferterly, 2007. The knowledge-based view, nested heterogeneity, and new value creation: philosophical considerations on the locus of knowledge. Academy of Management Review 32, 195–218.

Fiedler, K., 2011. Voodoo correlations are everywhere – not only in neuroscience. Perspectives on Psychological Science 6, 163–171.

Foss, N.J., 2010. Micro-foundations for management research: what, why, and whither? Cuadernos de Economía y Dirección de la Empresa 13, 11–34.

Foss, N.J., 2011. Why micro-foundations for resource-based theory are needed and what they may look like. Journal of Management 37, 1413–1428.

Gordon, R.A., Howell, J.E., 1959. Higher Education for Business. Columbia University Press, New York.

Greenwald, A.G., 1975. Consequences of prejudice against the null hypothesis. Psychological Bulletin 82, 1–20.

Gregoire, D.A., Shepherd, D.A., Lambert, L.S., 2010. Measuring opportunity-recognition beliefs illustrating and validating an experimental approach. Organizational Research Methods 13, 114–145.

Hitt, M.A., Harrison, J.S., Ireland, R.D., Best, A., 1998. Attributes of successful and unsuccessful acquisitions of US firms. British Journal of Management 9, 91–114.

Hull, C.L., 1928. Aptitude Testing. World Book Company, Chicago, IL.

Judge, T.A., Bono, J.E., Ilies, R., Gerhardt, M., 2002. Personality and leadership: a qualitative and quantitative review. Journal of Applied Psychology 87, 765–780.

Kruschke, J.K., Aguinis, H., Joo, H., 2012. The time has come: Bayesian methods for data analysis in the organizational sciences. Organizational Research Methods 15, 722–752.

Lance, C.E., 2011. More statistical and methodological myths and urban legends. Organizational Research Methods 14, 279–286.

Lance, C.E., Vandenberg, R.J. (Eds.), 2009. Statistical and Methodological Myths and Urban Legends: Received Doctrine, Verity, and Fable in the Organizational and Social Sciences. Routledge, New York.

Mollick, E., 2012. People and process, suits, and innovators: the role of individuals in firm performance. Strategic Management Journal 33, 1001–1015.

Molloy, J.C., Ployhart, R.E., Wright, P.M., 2011. The myth of ''the'' micro-macro divide: bridging system-level and disciplinary divides. Journal of Management 37, 581–609.

Mone, M.A., Mueller, G.C., Mauland, W., 1996. The perceptions and usage of statistical power in applied psychology and management research. Personnel Psychology 49, 103–120.

O'Boyle, E., Aguinis, H., 2012. The best and the rest: revisiting the norm of normality of individual performance. Personnel Psychology 65, 79–119.

Orr, J.M., Sackett, P.R., DuBois, C.L.Z., 1991. Outlier detection and treatment in I/O psychology: a survey of researcher beliefs and an empirical illustration. Personnel Psychology 44, 473–486.

Podsakoff, P.M., Dalton, D.R., 1987. Research methodology in organizational studies. Journal of Management 13, 419–441.

Popper, K., 1963. Conjectures and Refutations: The Growth of Scientific Knowledge. Harper Torchbooks, New York.

Rosenthal, R., 1979. The ''file drawer problem'' and tolerance for null results. Psychological Bulletin 86, 638–641.

Rousseau, D.M. (Ed.), 2012. The Oxford Handbook of Evidence-based Management. Oxford University Press, New York.

Scandura, T.A., Williams, E.A., 2000. Research methodology in management: current practices, trends, and implications for future research. Academy of Management Journal 43, 1248–1264.

Schmidt, F.L., Hunter, J.E., 1983. Individual differences in productivity: an empirical test of estimates derived from studies of selection procedure utility. Journal of Applied Psychology 68, 407–414.

Shepherd, D.A., Sutcliffe, K.M., 2011. Inductive top-down theorizing: a source of new theories of organization. Academy of Management Review 36, 361–380.

Short, J.C., Ketchen, D.J., Combs, J.G., Ireland, R.D., 2010. Research methods in entrepreneurship: opportunities and challenges. Organizational Research Methods 13, 6–15.

Starbucks Corporation, 2012. FY(2011) Annual Report. SJBQ412OTH-00801.

Suellentrop, C., 2013. Divining what's next for video games. The New York Times, Retrieved from http://www.nytimes.com/2013/03/12/arts/video-games on March 30, 2013.

Tiffin, J., 1947. Industrial Psychology, 2nd ed. Prentice-Hall, New York.

Van de Ven, A.H., Lifschitz, A., 2013. Rational and reasonable microfoundations of markets and institutions. Academy of Management Perspectives 27, 156–172.

Vedantam, S., 2012. Indian Engineers Build A Stronger Society With School Lunch Program. National Public Radio: April 6.

Viswesvaran, C., Barrick, M.R., Ones, D.S., 1993. How definitive are conclusions based on survey data: estimating robustness to nonresponse. Personnel Psychology 46, 551–567.