© Academy of Management Learning & Education, 2014, Vol. 13, No. 3, 446–466. http://dx.doi.org/10.5465/amle.2013.0193

Evidence-Based Management and the Trustworthiness of Our Cumulative Scientific Knowledge: Implications for Teaching, Research, and Practice

SVEN KEPES
ANDREW A. BENNETT
MICHAEL A. McDANIEL
Virginia Commonwealth University

The promise of evidence-based management (EBMgt) is that educators and practitioners can access cumulative scientific information to guide their teaching and decision making. We argue that the EBMgt movement may be unable to live up to this promise to the extent that our cumulative scientific knowledge is not trustworthy. We review why the management discipline may not have trustworthy evidence and how this widens the divide among educators, practitioners, and researchers. Next, we review the implications of untrustworthy cumulative knowledge, focusing on how educators can critically assess evidence and teach this process in the classroom. We close with recommendations for improving the trustworthiness of our literature to enhance teaching and practice from an evidence-based perspective. Suggestions include increasing the reproducibility and replication of primary studies, changing the editorial review process, and focusing on the production and dissemination of practically relevant and actionable knowledge.

The concept of evidence-based management (EBMgt), "the systematic, evidence-informed practice of management" (Rousseau, 2012a: 3), is part of a broader evidence-based practice movement. The general movement started in the medical sciences in the late 1960s and early 1970s (Cochrane, 1972; Feinstein, 1967) and has since expanded to various scientific fields, including education (Buskist & Groccia, 2011), social work (Bellamy, Bledsoe, & Traube, 2006), criminal justice (Mears & Barnes, 2010), and management (Rousseau, 2006). The promise of EBMgt is that it provides collective "scientific evidence and validated local facts" as the starting point for managers to make decisions (van Aken & Romme, 2012: 43) and should improve both what and how we teach management. Thus, EBMgt may help to overcome the research–practice gap (Pfeffer & Sutton, 2006, 2007). Effective use of EBMgt when teaching is also important because educated and competent managers are rare resources who can provide value to an organization (Rousseau, 2006). Therefore, educators who teach from an evidence-based perspective can enhance the competitive advantage of organizations through educating managers to distinguish between sound evidence-based recommendations and some of the dangerous half-truths and "total nonsense that often passes for sound advice" (Pfeffer & Sutton, 2006:

13). However, to fulfill the promise of EBMgt, all the evidence must be trustworthy. Across scientific disciplines, recent research has questioned the trustworthiness of published findings.

Here we examine the trustworthiness of our scientific knowledge and highlight why trustworthy information is critical for achieving the promise of EBMgt (Briner, Denyer, & Rousseau, 2009; Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000; Satterfield Spring, Brownson, Mullen, Newhouse, Walker, & Whitlock 2009). We discuss the EBMgt process and the implications for teaching from an evidence-based perspective, focusing on critically assessing evidence and having open, honest, and transparent discussions of shortcomings of our research evidence. In addition, we suggest ways to improve the trustworthiness of our cumulative evidence and translate it into actionable knowledge for educators and practitioners. This two-pronged approach (improving our teaching and our research) should improve the effectiveness of evidence-based practice.

EVIDENCE-BASED MANAGEMENT

Rousseau (2006) and Pfeffer and Sutton (2006) introduced the term *evidence-based management* to the academic discipline of management. The promise of EBMgt is that it will close the research–practice gap, ensuring that current and future managers make decisions based on the best available empirical evidence (both general and local) instead of personal preferences or conventional wisdom (Pfeffer & Sutton, 2006; Rousseau, 2006). Using trustworthy scientific information should improve decision-making outcomes (Briner et al., 2009; Pfeffer & Sutton, 2006) by minimizing the use of ineffective management practices (Rousseau & McCarthy, 2007) as well as by aligning management practices with organizational goals (Rousseau, 2006).

A primary role of scholars in applied fields such as management is to find trustworthy empirical evidence, translate it for educators and practitioners, and make it publically available (Reay, Berta, & Kohn, 2009; Shapiro, Kirkman, & Courtney, 2007). This is a fundamental part of the infrastructure required for effective EBMgt (Briner et al., 2009). When this infrastructure is available and students and practitioners trust it, EBMgt provides empirical justification for when and why specific practices should be used (Cronin & Klimoski, 2011). This will facilitate closer ties between educators, researchers, and practitioners because research will be focused on creating positive outcomes outside of academia (Abrahamson & Eisenman, 2001), which is already occurring in specific areas in human resource management (Gibbons & Woock, 2007) and industrial-organizational psychology (Briner et al., 2009). Examples include practice-focused research in areas such as personnel selection and goal setting (Rousseau, 2012b).

The Process of Evidence-Based Practice

The process of evidence-based practice follows five well-defined steps: (1) ask the question, (2) acquire evidence, (3) appraise the evidence, (4) apply the evidence, and (5) analyze the effectiveness and efficiency of the decision and, if necessary, make adjustments (Jelley, Caroll, & Rousseau, 2012; Straus, Glasziou, Richardson, & Haynes, 2011). The first step typically originates from a current problem or issue in the workplace and then develops into formulating an answerable question and choosing search terms and keywords. Next, using the search terms and keywords, one gathers scientific evidence, practitioner expertise, and organizational information to answer the question using on-line databases and other resources. After acquiring the available evidence, one needs to critically appraise it. Given the potential concerns with the trustworthiness of the cumulative scientific evidence, this step is especially vital, and thus, a focal point of the article. After appraisal, one should integrate all acquired evidence and consider the unique circumstances in the organization (Briner & Rousseau, 2011; Satterfield et al., 2009). Next, the integrated evidence is applied by making a decision to solve the problem (i.e., to answer the question formulated in the first step). In the last and final step, the effectiveness of the decision is analyzed and, if necessary, adjustments can be made.

Trust as a Prerequisite to Teaching and Practicing Evidence-Based Management

The EBMgt process illustrates that trustworthy cumulative knowledge is part of the infrastructure required for effective EBMgt (Briner et al., 2009). Without such information, practitioners may not have sound empirical justification for when and why specific practices should be used (Cronin & Klimoski, 2011). Consistent with prior scientific and practitioner use of the term *trustworthy* (Hintermann, Alberini, & Markandya, 2010; Kepes & McDaniel, 2013; Pfeffer & Sutton, 2006; Smith, 2013), we use it to

indicate evidence that is valid, reliable, generalizable, has been replicated, was derived in a transparent manner, and to indicate that our published literature is representative of all evidence.

Scholars and practitioners need evidence about what has worked in the past and prescriptions for what should work in the future (Rousseau & McCarthy, 2007). Given the research–practice gap, it is likely that both researchers and practitioners consider their particular knowledge as more valid (Briner & Rousseau, 2011). Practitioners may have misgivings about research because some academics do not provide scientific evidence that is organizationally defensible or readily transferable to practitioners (Cronin & Klimoski, 2011). Also, practitioners perceive, often accurately, that academics are disconnected from real management problems (Giluk & Rynes, 2012). Practitioners may also discount science-based recommendations because they lack knowledge about science, the scientific method, and statistics (Giluk & Rynes, 2012), and they may prefer to rely on their own experiences as the primary basis for their decisions.

Evidence-based management has the potential to overcome these and related problems. To do so and to live up to its promise, the empirical evidence underlying EBMgt should be trustworthy. For instance, if our published literature excludes (e.g., suppresses) small magnitude effects, our literature will overestimate typical effect sizes (Kepes, Banks, McDaniel, & Whetzel, 2012; Rothstein, Sutton, & Borenstein, 2005). Consequently, practitioners' applications of research findings without a critical assessment may be less successful than anticipated based on research findings. Thus, even with an excellent design, good psychometric properties, and a sound analytic approach, practitioners may still view our evidence as "pointless" (Pearce & Huang, 2012a: 259), not applicable (Straus et al., 2011), and untrustworthy (Giluk & Rynes, 2012). Yet, trust in our scientific knowledge is needed for practitioners to use and apply it (Amabile et al., 2001). To overcome these issues, we need to be honest and transparent about the shortcomings of our scientific research and the threats to its trustworthiness.

ASSESSING THE TRUSTWORTHINESS OF OUR CUMULATIVE KNOWLEDGE

Unfortunately, the management discipline does not have a process that routinely assesses the trustworthiness of the scientific knowledge it produces. One plausible explanation is that the reward structure inhibits scientific progress and publication bias prohibits compiling all data that exists (Kepes & McDaniel, 2013). Our current reward structure encourages the publication of articles that discover something new over an emphasis on replicable results (Nosek, Spies, & Motyl, 2012; Pfeffer, 2007; Pillutla & Thau, 2013). Articles on trendy or novel topics that yield statistically significant findings are more likely to get published, particularly in top-tier journals (Hartshorne & Schachner, 2012; Kepes & McDaniel, 2013). In addition, our field emphasizes theory development (Hambrick, 2007; McKinley, 2010) and statistically significant results (Fanelli, 2012; Sterling & Rosenbaum, 1995) over robust[1] and replicable findings. Thus, there is a "quest for 'what's new' rather than 'what's true' " (Pfeffer, 2007: 1339).

By contrast, trustworthiness is enhanced when studies are replicated. In fact, one of the hallmarks of any scientific discipline is its ability to self-correct; to disconfirm incorrect or misleading theories and empirical findings (Merton, 1973). However, this self-correcting ability has been questioned, especially in the social sciences (Ferguson & Heene, 2012; Kepes & McDaniel, 2013; Nosek et al., 2012). Moreover, even the popular press has called the self-correction process in science a myth (Estes, 2012), asking whether there "is something wrong with the scientific method" (Lehrer, 2010: 52; see also, e.g., *The Economist*, 2013), partly because the submission of replication studies is typically discouraged by our journals (Makel, Plucker, & Hegarty, 2012; Neuliep & Crandall, 1990; Yong, 2012). Because of these and other structural problems within the scientific process in management, a "disconnect between what is good for scientists and what is good for science" exists (Nosek et al., 2012: 616; see also, e.g., Adler & Harzing, 2009; Kepes & McDaniel, 2013). Without trustworthy cumulative knowledge, it is difficult to teach or practice from an evidence-based perspective effectively because the infrastructure required to fulfill the promise of EBMgt could be compromised. The management discipline is not alone in facing this problem. In the medical sciences, Ioannidis (2005b: e124) proclaimed that "most published research findings are false." Thus, one may ask whether these

[1] A sensitivity analysis examines the extent to which results and conclusions are altered as a result of changes in data or analysis approach (Greenhouse & Iyengar, 2009). Results and conclusions that remain largely stable are termed *robust*. Management research seldom conducts sensitivity analyses (Kepes & McDaniel, 2013).

warnings and statements are true, and, if true, what can be done to address them. Next, we review the evidence regarding the trustworthiness of our cumulative knowledge.

"Without trustworthy cumulative knowledge, it is difficult to teach or practice from an evidence-based perspective effectively."

Evidence Regarding the Trustworthiness of Our Cumulative Knowledge

Across scientific disciplines, nonsignificant results do not appear in the published scientific literature as often as they occur (Fanelli, 2012; Simonsohn, Nelson, & Simmons, 2014; Sterling & Rosenbaum, 1995). In other words, journal articles should contain more nonsignificant findings than they do. Evidence for the suppression of results from the publically available literature (i.e., publication bias; Rothstein et al., 2005), and thus the distortion of cumulative knowledge, is abundantly available in the medical sciences (e.g., Chalmers, 1990; Ioannidis, 2005b; Sutton, 2005). With regard to *primary studies* in the social sciences, the most visible example of this phenomenon may be the recent controversy about Bem's (2011) article, which supported the presence of psychic effects (LeBel & Peters, 2011; Ritchie, Wiseman, & French, 2012; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). However, there are other examples in the social sciences as well (e.g., Barends, Janssen, ten Have, & ten Have, 2014; Blanton et al., 2009; Doyen, Klein, Pichon, & Cleeremans, 2012; Francis, 2012; Schimmack, 2012; Vul, Harris, Winkielman, & Pashler, 2009).

Evidence for publication bias in *meta-analytic studies* is also well-documented in the social sciences, including education (Banks, Kepes, & Banks, 2012a), economics (Doucouliagos & Stanley, 2009), and psychology (Wagenmakers et al., 2011). In their review of meta-analytic studies in psychology, Ferguson and Brannick (2012) concluded that the results of approximately 40% were affected by publication bias, the suppression of small magnitude, nonsignificant results from the publically available literature. This affects the accuracy of meta-analytic results (Kepes et al., 2012), which are the primary way to advance and support EBMgt (Briner et al., 2009; Le, Oh, Shaffer, & Schmidt, 2007; Rousseau & McCarthy, 2007).[2]

In recent years, the cumulative knowledge in management literatures has also been assessed. McDaniel, Rothstein, and Whetzel (2006) concluded that the results summarizing the effectiveness of commercial employment tests for some test vendors are consistent with an inference of the suppression of small magnitude correlations such that the reported results are likely overestimated. More recently, Banks, Kepes, and McDaniel (2012b) found evidence that the published literature on conditional reasoning tests for aggression was consistent with an inference of publication bias in the direction of overestimating the validity of the test. Kepes and colleagues (2012) reported that the validity estimate on the relation between structured interviews and job performance using data from journal articles was almost 45% higher than the validity estimate on the same relation using data from sources other than journal articles. Other management-relevant publication bias findings have also been reported, ranging from microtopics, such as the efficacy of Pygmalion interventions (e.g., Kepes, Banks, & Oh, 2014; Renkewitz, Fuchs, & Fiedler, 2011) to macrotopics, such as the relation between top management team diversity and firm performance (e.g., Homberg & Bui, 2013; O'Boyle, Banks, & Rutherford, in press). Overall, a growing body of evidence indicates that the results of some meta-analytic reviews in management contain misleading or erroneous results,[3] which damages the trustworthiness of our literature. Because meta-analytic reviews are at the top of the evidence hierarchy in scientific research (Earl-Slater, 2001; Greenhalgh, 1997; Guyatt et al., 1995) and are a cornerstone of EBMgt (Briner et al., 2009; Le et al., 2007; Rousseau & McCarthy, 2007), teaching and

[2] We note that meta-analytic reviews, although the primary form of systematic reviews, are only one particular type of systematic review, and that systematic reviews are a cornerstone of EBMgt. However, given that virtually all systematic reviews in the organizational sciences are meta-analytic reviews, we focus mainly on such systematic reviews.

[3] We note that Dalton and colleagues (2012) concluded that publication bias is not of concern in the management and I-O psychology literatures. Yet, Dalton and colleagues are alone in that conclusion. Furthermore, their conclusion has been questioned based on conceptual and methodological grounds (Kepes et al., 2013; Kepes & McDaniel, 2013), and we assert that it does not inform our knowledge concerning the occurrence or severity of publication bias in the management and related literatures.

practicing EBMgt can be problematic if the results of such reviews are not trustworthy.[4]

RECOMMENDATIONS FOR TEACHING

The goal of evidence-based education is to help students and practitioners learn the process of asking questions, acquiring, assessing, and applying the evidence, and when applicable, analyzing the effectiveness of the application and making adjustments (Keim, Howse, Bracke, & Mendoza, 2008). Thus far, we have argued that some scientific evidence in management may not be trustworthy because of several problems within the scientific process in our field. As such, teaching from an evidence-based perspective requires being honest and transparent about the shortcomings of our research and the threats to its trustworthiness. We acknowledge that since most educators are involved in both teaching and research, being open about potential problems associated with their own research or the research of their colleagues could be difficult. Nonetheless, research shows that honesty and transparency has the potential to increase perceived authenticity and decrease avoidance behaviors (Gillath, Sesko, Shaver, & Chun, 2010), which could reduce the research–practice gap in our field. In the next section, we first discuss teaching from the evidence-based perspective in general, before providing a detailed overview of the step-by-step EBMgt process and describing how this process can be taught in a classroom environment.

Teaching From an Evidence-Based Perspective

Evidence-based medicine (EBMed) is often used as the "blueprint" or exemplar for EBMgt (Barends, ten Have, & Huisman, 2012; Pfeffer & Sutton, 2006; Rousseau, 2006), and the adoption of practices from the medical and related sciences can provide some guidance. Given that much of the literature in EBMed focuses on educating graduate students, we also borrow from medical programs with an emphasis on undergraduate education (e.g., nursing; Melnyk & Fineout-Overholt, 2011). Teaching from an evidence-based perspective can occur using two different approaches: the push or the pull approach. The push method centers around "pushing" the best available scientific evidence on a topic of interest to students (Straus et al., 2011). This approach is important because it helps students develop sufficient knowledge, skills, and abilities to make automatic decisions (e.g., system 1 processing; Kahneman, 2011) or incorporate learned evidence into the heuristics used when making decisions (Gigerenzer & Gaissmaier, 2011). An example that is very common is lecturing: telling students the cumulative scientific evidence on a specific topic, such as stating that "general cognitive ability is the best predictor of job performance." Therefore, this approach to teaching EBMgt requires the educator to have completed the critical appraisal process before "pushing" the evidence to the students.

The second approach to teaching EBMgt is the "pull" approach. Here, students are "pulled" toward the scientific evidence and other information and engage in the EBMgt process themselves (i.e., ask, acquire, appraise, apply, and analyze and adjust). Having students decide on a question or topic of interest can be accomplished by providing them with an organizational example (e.g., a case study) or inviting a local business professional to speak with the class. If less time is available, instructors can also pose specific questions to the students and require them to engage in the next steps of the EBMgt process on their own. An example would be asking the class, "if you were developing a personnel selection test to hire new employees, what would be the best construct to measure?"[5] Although the definitions above make a distinction between these two primary approaches, they are not exclusive of each other. Many excellent educators use both. However, teaching from a "pull" approach ensures that students become familiar with and engage in the EBMgt process.

Depending on the curriculum and prerequisites of a course, educators might need to review the scientific method, correlation and regression sta-

[4] We note that the extent to which publication bias is a problem in our scientific literature is likely to vary across research topics and literature areas. Some literature areas are likely to be severely or moderately affected by this bias, while others may exhibit negligible or no bias (e.g., Dickersin, 2005; Rothstein et al., 2005; Kepes et al., 2014). The degree to which individual literature areas are affected by publication bias is currently unknown.

[5] We note that, at this early phase in the evidence-based practice process, the question tends to be relatively broad. As one progresses through the evidence-based practice process, the question needs to become more refined and focused to properly guide the search for applicable and useful evidence.

tistics, introduce students to academic research articles, explain the hierarchy of evidence (Greenhalgh, 1997), and describe the appropriateness of different research designs for distinct research questions and contexts (Berwick, 2008; Petticrew & Roberts, 2003). Drawing upon the process of teaching EBMed (Sackett et al., 2000), instructors of EBMgt in the classroom begin with a case statement, or a description of a situation or problem involving an individual or organization. The case statement can include nonpublished information such as organizational data or expert opinions in addition to the problem or situation. This aligns with the "Patient A comes into the hospital" opening common to medical pedagogy, and can be integrated with case studies and other problem-focused course material typically found in management classes (e.g., simulations, video case studies, and other "hands-on" methods; Leung & Bartunek, 2012).[6] Educators can choose to have students work through the EBMgt process individually or in small groups. It is suggested that such case statement activities occur in class so that the educator can act as a learning coach and help students find relevant evidence, assess its trustworthiness, and develop a solution (Sackett et al., 2000; Straus et al., 2011). In addition, allocating time for students to share their findings during class makes learning cumulative, allows students to learn from the successes and mistakes of others, and refines the evidence-based practice skills of both students and educators (Straus et al., 2011).

Asking the Question

This first step of the EBMgt process often originates from a real-life problem that occurs within organizations. One should identify the problem specifically as well as formulate the problem into an answerable question. When asking a question, it is important to use the appropriate scientific terminology as this facilitates the next step in the evidence-based practice process (acquiring the evidence). The PICOC framework (population, intervention, comparison, outcome, context; e.g., Barends et al., 2012) can help one narrow the question and identify the proper search terms and keywords for the subsequent step in the evidence-based practice process (i.e., acquire the evidence). This framework helps one to consider if there is a specific sample population of interest, if a particular intervention is needed, if a comparison group should be part of the evidence to be collected, and if there are specific outcomes and contexts that need to be considered. For example, one would include the keywords "executives" (population), "stress management intervention" (intervention required), and "manufacturing" (industry context) if these elements are part of the case statement.

Teaching students how to "ask the question." In the "ask the question" step, educators should ask students to formulate a specific question derived from the information in the case statement. Their first time doing this might require collecting questions from each student or group of students and then determining with the entire class the specific question that all will use. To generate search terms, instructors should help students recognize variables of interest that can be used as keywords. The PICOC framework should be used to determine if the case statement requires investigating contextual issues such as a specific population, intervention, or industry.

Acquiring the Evidence

Once the search terms and keywords are determined, the acquisition of the best available, most relevant information can begin. This process occurs in multiple ways, ranging from on-line searches of electronic academic databases to the collection of available organizational and contextual information and asking experts for relevant information.

Teaching students how to "acquire the evidence." Most students in this "information age" understand the process of using a search engine to gather information. However, conducting an effective electronic search can be difficult without proper training (for a good review on the search process and search skills training, see Goodman, Gary, & Wood, this issue). Educators should have students begin this step by completing a search using both a regular web search engine and academic databases (e.g., PsycNet). Prompting students to examine the source of the material provides an excellent opportunity to help students begin differentiating between academic publications, trade publications, and the popular press. Educators can help students acquire information from local sources by encouraging them to inter-

[6] An example of such a case statement and other materials for an EBMgt class activity is available in the on-line supplemental materials.

view local business professionals or having a question-and-answer session with a guest speaker during class time. To help students gather organizational information, educators can suggest they find a video or transcript of a CEO's speech or a strategic plan from the organization, as these might provide additional, relevant information.

Appraising the Evidence

Given that some of the information found might not be trustworthy, individuals must critically appraise the information they acquire. The critical appraisal should include all information, including what is found within published works such as academic journal articles, textbooks, articles in trade journals and the popular press, as well as nonpublished information such as organizational data and practitioner expertise. Next, we describe the critical appraisal process for both published and nonpublished information, as well as how educators can teach this process.

Critically appraising published information. When appraising published scientific evidence, one needs to assess potential threats to the trustworthiness of the research, including issues not necessarily covered by the trinity of research design, measurement, and analysis (Pedhazur & Pedhazur Schmelkin, 1991). When appraising evidence from a primary study, one has to evaluate it according to its trustworthiness, which includes traditional scientific factors such as validity, reliability, and generalizability, as well as contextual factors such as quality, importance, and relevance (Briner et al., 2009; Greenhalgh, 1997; Guyatt et al., 1995; Jelley et al., 2012; Straus et al., 2011). Guidelines, standards, and checklists can facilitate the critical appraisal. For instance, the Journal Article Reporting Standards (JARS) of the American Psychological Association (APA, 2008) and other guidelines and instruments (e.g., AGREE II; Brouwers et al., 2010; QUADAS-2; Whiting et al., 2011) for the evaluation of primary quantitative research should be used during this assessment. When appraising qualitative studies, different standards and guidelines apply when assessing issues such as quality, importance, relevance, and methodological soundness (Horsburgh, 2003). An assessment of the nature of the study context and the data analysis may be especially important to determine whether the results are trustworthy (Pratt, 2008). The Center for Evidence-based Management also has questionnaires for the critical appraisal for various types of primary research on their website (see http://www.cebma.org/teaching-materials/).

One example of a question that should be addressed during the critical assessment relates to the transparency of published studies; therefore, one should ask, "Who benefits from this research?" Considerations of conflicts of interest and related issues (Wagner & Steinzor, 2006) can help assess whether the evidence found in a primary study is trustworthy enough to be used when making evidence-based decisions. Unfortunately, few management journals require statements of conflict of interest. Thus, educators may want to assess whether authors of articles or the organizations with which the authors are affiliated benefit financially or otherwise from the published research (e.g., does the research document the efficacy of a commercial practice or product, and are the authors associated with the organization that developed the practice or product?).

Another example question that should be asked during the critical assessment is "Are there replications of the finding(s)?" Because every individual study has limitations and every result could be due to factors other than those offered by the researcher (e.g., measurement error, sampling error, threats to internal and external validity, construct and predictive validity), one can have greater confidence in the evidence if there are replications of the research that yield similar results and conclusions (Jasny, Chin, Chong, & Vignieri, 2011). Unfortunately, replication studies are rare in our literature (Barends et al., 2014; Makel et al., 2012; Yong, 2012). This poses a problem because it is difficult to determine the trustworthiness of a particular finding from a single study (Hunter & Schmidt, 2004). For example, the first study on a particular topic often overestimates the magnitude of the results, and thus, the importance or practical impact of the conclusions (time-lag bias; Banks et al., 2012b; Kepes et al., 2012).[7] Greater confidence can be placed in the evidence if results can be averaged across replications and if the summary mean estimate is not substantially affected by the first few studies.

[7] The time-lag bias refers to the situation where the time to complete and publish a study is affected by the study's results (Ioannidis, 1998; Trikalinos & Ioannidis, 2005). If the time-lag bias is present, a study's time to publication tends to be shorter when it contains statistically significant results than when it does not (see Banks et al., 2012b; Kepes et al., 2012).

Because meta-analytic reviews play a fundamental role in the development, assessment, and dissemination of our cumulative knowledge (Briner et al., 2009; Le et al., 2007; Rousseau & McCarthy, 2007), they should be critically assessed as well. The assessment of such reviews should consider meta-analytic professional standards and guidelines as described in the Meta-Analysis Reporting Standards (MARS; APA, 2008), QUOROM; Moher et al., 1999), PRISMA (Liberati et al., 2009), AMSTAR (Shea et al., 2009), the Cochrane Collaboration guidelines (Higgins & Green, 2011), and general best-practice research synthesis recommendations (e.g., Cooper, Hedges, & Valentine, 2009; Kepes, McDaniel, Brannick, & Banks, 2013). Unfortunately, studies across scientific disciplines, from the medical sciences (Sacks, Berrier, Reitman, Ancona-Berk, & Chalmers, 1987) to the organizational sciences (Kepes et al., 2013), have indicated that many, if not the majority, of published meta-analyses do not adhere to these standards.

In meta-analytic reviews, one should evaluate if the literature search process is comprehensive, transparent, and replicable. Similarly, one may want to evaluate if the primary samples are comparable (e.g., are the measures sufficiently similar?). Aggregating effect sizes from samples that are not comparable can lead to misleading meta-analytic results. Similarly, one should assess whether sensitivity analyses were conducted to assess the robustness of the results due to outliers, publication bias, the time-lag bias, data imputations, and so on (Greenhouse & Iyengar, 2009; Kepes et al., 2013). As past assessments of the meta-analytic review process have indicated, these issues are typically not addressed in meta-analytic reviews in the management literatures. Yet, they are important when assessing the trustworthiness of meta-analytic results.

The critical appraisal is similar for textbooks (Sackett et al., 2000; Straus et al., 2011). Theoretically, textbooks should provide summaries of the best cumulative evidence in a scientific discipline. Unfortunately, just the opposite could be true (Goodman & O'Brian, 2012; Pearce, 2012a; Sackett et al., 2000). In fact, our textbooks tend to focus more on general discussions of theories that may have "considerable appeal to many people [despite the fact that] the prevailing view in the academic literature is that the specific hypotheses of these theories are not supported by empirical evidence" (Gerhart & Rynes, 2003: 53). As a result, our textbooks often fail to "distill the high-caliber evidence that does exist into principles on which learners or managers can base their professional practice" (Rousseau & McCarthy, 2007: 86). Thus, some of our textbooks may not be suitable to teach from an EBMgt perspective. Although select textbooks are beginning to use an evidence-based framework to discuss management topics (e.g., Pearce, 2012b), they may still be less than adequate because they do not always use the most current and trustworthy research findings (Dipboye, this issue). That is why Sackett and colleagues (2000) advised against the use of textbooks when teaching from an evidence-based perspective (see also Straus et al., 2011).

Critically appraising nonpublished information. In addition to scientific evidence, other sources of information are available (e.g., practitioner expertise, stakeholder preferences, and evidence from the local context; Briner et al., 2009; Satterfield et al., 2009). This information can also be used when making evidence-based decisions and may be especially useful in situations when no scientific evidence is available (Pearce, 2012a). Just as there are concerns about the trustworthiness of published information, other information (e.g., practitioner expertise, stakeholder preferences, and evidence from the local context) is not necessarily trustworthy as a basis for making sound conclusions and evidence-based decisions, either (Latham & Locke, 2009: 90; see also Pfeffer & Sutton, 2006, 2011). Thus, nonscientific information, including firm-specific big data, should also be critically evaluated (Baba & HakemZadeh, 2012; Davenport & Harris, 2007).

There are instances in which stakeholder preferences may not match practitioner expertise (Guskey, 2007) or are in conflict with sound empirical evidence (Pfeffer & Sutton, 2006, 2011; Silver, 2012). One example of a mismatch between organizational data and scientific evidence is the definition and construct validity of Gallup's Employment Engagement Survey. Although many organizations use the Gallup survey, the scientific community has found that the items of this questionnaire address the work environment but not employee engagement (Macey & Schneider, 2008). Thus, managers who use the Gallup survey might change select working conditions (e.g., setting clear job expectations) without improving the level at which their workforce is engaged, defined as employees being excited and invigorated by their work.

To assist in the critical appraisal of nonpublished information, some items and questions from

the previously discussed standards, guidelines, and instruments (e.g., APA, 2008) can be helpful. For example, organizational data are sometimes constrained by measurement deficiencies in the variables assessed (e.g., the typical low reliability of supervisor's evaluations of performance). Another potential problem with organizational information is that big data are usually noisy, unstructured, and may require complex data analysis capabilities that are not necessarily available in organizations (Davenport & Harris, 2007; Ouellette, 2013; Silver, 2012). If proper measurement of or analysis with these data cannot be verified, it is difficult to conclude that the information is trustworthy.

Teaching students how to "appraise the evidence." Once students have learned how to properly conduct an effective electronic search for information (Goodman et al., this issue), the next step is teaching how to appraise the acquired evidence. In medical school, students complete a critically appraised topic (CAT; Sauve et al., 1995), which is a concise summary of the critically appraised best available evidence with very short, bottom-line recommendations to the problem (Sackett et al., 2000; Straus et al., 2011). The use of CATs has successfully transferred to applied programs such as nursing (Jones, Crookes, & Johnson, 2011) and library science (Glynn, 2006). A CAT is very practitioner-focused because it summarizes key scientific evidence and translates it into useful measures of efficacy, risk, and accuracy (Sadigh, Parker, Kelly, & Cronin, 2012; Sauve et al., 1995) and could also be used in management and related courses.

To begin appraising management evidence, students should rank the acquired evidence using a hierarchy of evidence (e.g., Earl-Slater, 2001; Greenhalgh, 1997; Guyatt et al., 1995). In addition, they need to consider the research question they are trying to answer because distinct research designs can be more or less appropriate for answering their particular question (Berwick, 2008; Petticrew & Roberts, 2003). An adapted hierarchy, as seen in Figure 1, provides students with a visual representation of a ranking system of evidence available. At the top of the hierarchy are systematic reviews. In management and related disciplines, these are typically found in the form of meta-analytic reviews, but other types of systematic reviews may also be available. The next levels include primary studies that use an experimental design. Level II are randomized controlled trials. Although this research design is extremely rare in

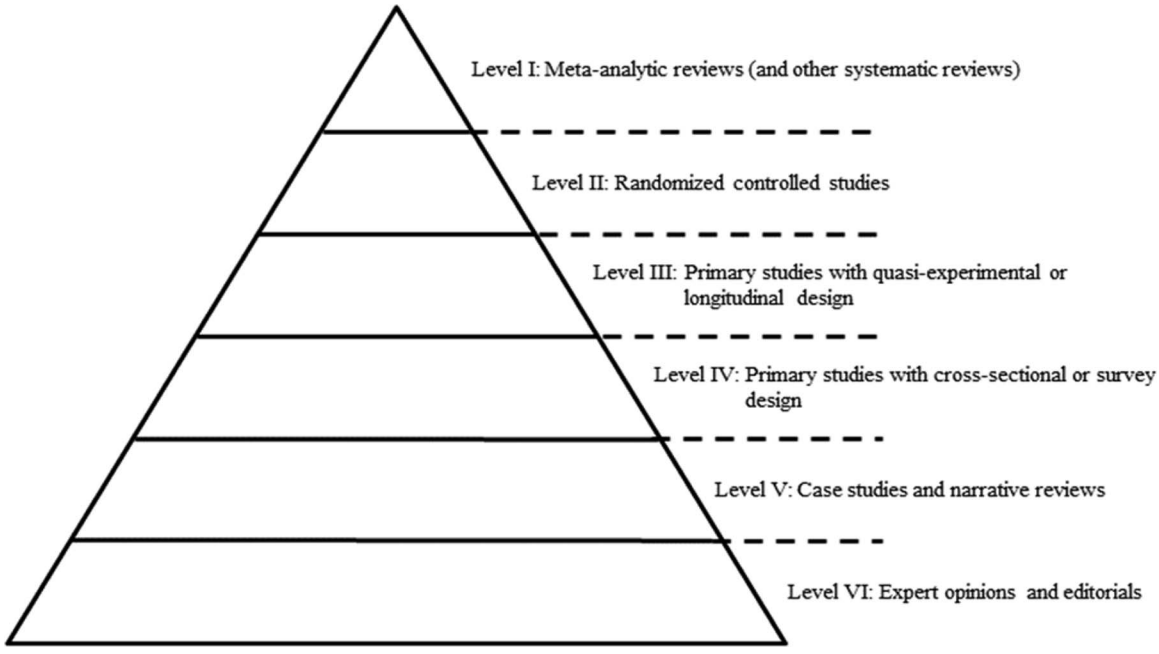

FIGURE 1

Hierarchy of Evidence to Assess Evidence in the Management Literature. *Note:* This is a general hierarchy of evidence. The appropriateness of different research designs is affected by, for instance, the particular research question and the context (Berwick, 2008; Petticrew & Roberts, 2003).

the organizational sciences (Reay et al., 2009), it does exist and provides the strongest causal inference. Level III studies have a quasi-experimental or a longitudinal design, which have the potential to allow for stronger causal inferences than cross-sectional designs. Quasi-experimental designs include a control and a treatment group but participants are not randomly selected for each group. Longitudinal research is the study of change with multiple observations of the same variable (Ployhart & Vandenberg, 2010). Level IV in the hierarchy contains cross-sectional and nonexperimental primary studies. Most research studies in management fall within this category (Barends et al., 2012). In addition, most organizational data also falls within this category. Level V are case studies or narrative literature reviews, describing in detail a specific organization or a collection of multiple examples without much statistical analysis. The lowest level in the hierarchy are the experiences and opinions of experts or respected authority.

After students have used the hierarchy of evidence to assess the general strength of acquired evidence while keeping in mind the appropriateness of distinct research designs, including the designs of the primary studies if a meta-analytic review is appraised (Berwick, 2008; Petticrew & Roberts, 2003), they should appraise it using the standards and guidelines discussed previously (e.g., APA, 2008). Table 1 summarizes some of the key points from these standards and guidelines, and this shorter format might be more useful for an undergraduate class. Educators can provide a series of written or verbal prompts based on the questions in Table 1 to help students appraise the trustworthiness of the evidence. For example, when asking students to assess if the measures and research methodology of an acquired study are appropriate, educators could encourage their students to read sample items from a questionnaire of an acquired study and ask them to see if they could come up with a better way to measure and collect the information. Educators can also ask students to examine the generalizability of the findings by assessing if the results or suggestions made by the authors are congruent with the population in the case statement. Sample questions could be: "Would you expect a similar result if you

TABLE 1
An Illustrative Set of Questions to Guide the Critical Assessment

Primary studies	
Validity	Do research design and measures match question being asked? Are results valid?
Generalizability	Has study been replicated? Can results transfer to a different population?
Reliability	Can this study be reproduced?
Relevance	Are results relevant? Does this study answer a useful question?
Quality	Is research design appropriate for research question? Do results show how something changed over time and attribute change to correct cause?
Transparency	Who benefits from this study? Who funded research? Were there incentives for authors to find results?
Meta-analytic (systematic) reviews	
Validity	Is number of studies in meta-analytic distribution large enough to produce credible results?
Generalizability	Are multiple sample populations included (e.g. different occupations, ages, genders)?
Reliability	Are primary samples comparable? Did these studies use similar measures and research designs?
Quality	Are all decision rules detailed? Can decision rules be replicated?
Literature search	Was literature search process transparent? Could process be replicated?
Comprehensiveness	Are any relevant published studies missing? Are unpublished studies included in analysis?
Sensitivity analyses	Did study include a publication bias analysis? Were results reported with/without outliers?
Textbooks	
Validity	Does textbook include theories with little empirical support or disproven theories (e.g., Maslow's hierarchy of needs)?
Quality	Does evidence cited come from trustworthy cumulative research (rather than one case study)?
References	Are in-line references provided next to evidence discussed?
Relevance	Is information provided graded by strength of evidence (e.g., strong evidence vs. weak evidence)?
Transparency	Did authors discuss their literature search and evidence-grading process? Could this be replicated?
Timeliness	Is most up-to-date information provided? A schedule or timeline to update chapters?

were a manager at a hospital? A construction firm?" In essence, this encourages students to use the PICOC framework to determine if the acquired and appraised evidence matches the case statement and the question they asked in the first step of the evidence-based practice process.

Applying the Evidence

The objective of this step is to integrate all available information, scientific evidence as well as information from the local context, before generating the best possible solution for the problem. This process is a more explicit form of decision making (e.g., system 2 processing; Stanovich & West, 2000) that should reduce human biases and lead to better, more evidence-based decisions (Pfeffer & Sutton, 2006, 2011).

Teaching students how to "apply the evidence." The fourth step when teaching the evidence-based management process requires that students summarize the appraised evidence and integrate it with additional elements of EBMgt: organizational information, practitioner expertise, and the local context (Briner et al., 2009). Educators may choose to make this integration process explicit by having students write brief summaries of all acquired information in a short bullet point format before comparing, contrasting, and integrating the scientific evidence with any other acquired information. Students conclude this step by determining a solution to the problem provided in the case statement.

Analyzing and Adjusting

Once a decision has been reached, the developed solution should be put into action. Following this, one should monitor the effectiveness of the application. Ideally, individuals should gather and analyze data to assess the effectiveness of the solution. This last step begins an iterative process in which one makes adjustments to the application of the solution so that the outcomes better align with the intended solution.

Teaching students how to "analyze and adjust." The last step involves applying the solution, which might be difficult or time consuming in some classes. However, using role-play or business simulations would allow for the analysis of the effectiveness of the solution. Educators could also ask students to collect data as part of a course project, and then help them analyze the collected data. In simulations or with a data collection in class, educators can ask students if adjustments should be made to their original solution and if they want to consider any adjustments to their solution to better address the problem from the case statement.

RECOMMENDATIONS FOR RESEARCH

Given that researchers play a critical role in the EBMgt process by providing the infrastructure (i.e., cumulative scientific evidence) required for EBMgt, it is important that the infrastructure be as trustworthy as possible. We therefore ask: How can we ensure that our cumulative knowledge, particularly our meta-analytic reviews, are as trustworthy as possible? To facilitate the effective critical appraisal of our scientific evidence, and thus the teaching and practice of EBMgt, we may have to makes changes in our scientific process. As when discussing recommendations for teaching, evidence-based medicine is often used as an exemplar for our recommendations.

Reproducibility of the Reported Results

Reproducibility refers to the ability of other researchers to obtain the same results when they reanalyze the same data. Studies from the medical sciences have illustrated that the degree of reproducibility is very low (e.g., Begley & Ellis, 2012; Ioannidis et al., 2009). In management and psychology, reproducibility is seldom assessed and rarely published. However, emerging evidence suggests that published findings may not be reproducible (e.g., Francis, 2012, 2013; Schimmack, 2012; Wagenmakers et al., 2011). As an example, Blanton and colleagues (2009) tried to reproduce the results of several articles that assessed the predictive validity of the Implicit Association Test (IAT; e.g., McConnell & Leibold, 2001; Ziegert & Hanges, 2005). However, Blanton and colleagues' (2009) reanalysis of the data from one these papers (McConnell & Leibold, 2001) revealed that the results are opposite to the ones originally reported.[8] This has clear implications for the quality of IAT-based EBMgt.

[8] We note that Blanton and colleagues (2009) also obtained results that question some of the reported results in one other IAT study (Ziegert & Hanges, 2005). In addition, of the eight datasets on the IAT that Blanton and colleagues (2009) tried to obtain, only half of the authors complied with their request. Based on the findings reported by Wicherts and colleagues (2011), one has to wonder if the data sets that were not obtained contained similarly weak or even weaker evidence for the claims contained in the original articles.

For instance, in his *Harvard Law Review* article, Kang (2005: 1514) stated that "there is now persuasive evidence that implicit bias against a social category, as measured by instruments such as the IAT, predicts disparate behavior toward individuals mapped to that category." In a later paper, Kang and Banaji (2006) relied on McConnell and Leibold's (2001) study when arguing that that antidiscrimination law should be revised to include implicit biases. However, based on more recent evidence regarding the predictive validity of the IAT (Blanton et al., 2009; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013), the negative repercussions to human resource management practice could have been substantial if the originally published research had been used to revise existing discrimination law. We should thus be more transparent with our data and ensure that published results are reproducible.

Replication Studies

Once a particular research finding is reproducible, we should assess whether it is replicable. Thus, while reproducibility is more concerned with the internal validity of a particular study, *replicability* is concerned with the external validity or generalizability (Asendorpf et al., 2013). Unfortunately, replication studies are relatively rare in the organizational sciences and seldom welcome in our journals (Kepes & McDaniel, 2013). In the medical sciences, Ioannidis (2005a) investigated the results of highly cited articles and found that out of 45 articles that claimed that the intervention was effective, the findings of only 20 (44%) were successfully replicated in subsequent studies. In their review of the psychology literature, Makel and colleagues (2012) estimated that only around 1% of all published studies are replications. We suggest that the number is not much different in management. For instance, when assessing the trustworthiness of the change management literature, Barends and colleagues (2014) noted that the typical published study has low validity and is never replicated. Given that replications are considered by many to be the "scientific gold standard" (Jasny et al., 2011: 1225), this should be of concern.

Meta-analytic reviews are often viewed as a primary means of synthesizing quantitative research findings from primary studies and communicating the summary results to the research and practitioner communities (Le et al., 2007).[9] Yet, even the trustworthiness of meta-analytic results have been questioned (Kepes et al., 2012; Rothstein et al., 2005). If the published results of our primary studies on a particular relation of interest are not representative of all studies on that relation, meta-analytic mean estimates are unlikely to be trustworthy. For instance, the initial meta-analysis on the relation between conditional reasoning tests of aggression and counterproductive work behaviors estimated the validity coefficient to be .44 (James et al., 2005). A subsequent meta-analysis yielded a substantially smaller mean estimate (.16; Berry, Sackett, & Tobares, 2010), which was further reduced to near .08 once the influence of publication bias was considered (Banks et al., 2012b). Given that meta-analytic reviews are at the top of the evidence hierarchy in our scientific research and a cornerstone of effective EBMgt, an assessment of the robustness of their results is especially important to the creation and dissemination of trustworthy cumulative knowledge. Variability in the trustworthiness of meta-analytic parameter estimates certainly exists; although some results may not be trustworthy, others could be. We thus echo prior calls for a rigorous assessments of the robustness of previously published meta-analytic results (e.g., Kepes & McDaniel, 2013; Kepes et al., 2013; McDaniel et al., 2006).

Robustness Checks, Objective Standards, Guidelines, and Checklists

Changes in the objectivity of the editorial review process could also help to ensure that our published results are trustworthy. As Hambrick (2007) and others (e.g., LeBel & Peters, 2011; McKinley, 2010) have highlighted, our top journals appear more concerned with the development of new and interesting theory than with methodological rigor and accuracy. This is in sharp contrast with the medical sciences where the lack of a strong theoretical framework is typically not a reason for the rejection of any paper (Byrne, 2000).[10] Authors may

[9] Given that practitioners across countries and culture are unlikely to read our journal articles (e.g., Giluk & Rynes, 2012; Rynes et al., 2002; Tenhiälä, Giluk, Kepes, Simón, Oh, & Kim, in press), this view may not reflect actual practice.

[10] Problems with the methodological approach, particularly related to the study design and the measures, are by far the most common reasons for article rejection in the medical sciences (Byrne, 2000).

engage in questionable research practices to fit the data to "interesting" hypotheses (Bedeian, Taylor, & Miller, 2010; Kepes & McDaniel, 2013; Simmons, Nelson, & Simonsohn, 2011). As an example, O'Boyle, Banks, and Gonzalez-Mule (in press) have documented the extensive use of questionable research practices (e.g., HARKing, hypothesizing after the results are known, Kerr, 1998) in transforming dissertation research into journal articles (see also Pigott, Valentine, Polanin, Williams, & Canada, 2013).

Some have advocated for a concentrated effort to strengthen the methods-oriented belief system of researchers, including editors and reviewers (e.g., Kepes & McDaniel, 2013; LeBel & Peters, 2011; McKinley, 2010). Authors should demonstrate that their results do not hinge on arbitrary methodological and analytic decisions (Simmons et al., 2011). If possible, different control variables, alternative operationalizations of constructs, and alternate analysis approaches should be used to demonstrate the robustness of obtained results. The mandatory use of a power analysis to assess the probability that the chosen analysis will reject the null hypothesis when the null hypothesis is false seems warranted as well (Simmons et al., 2011). The use of objective reviewing and publishing standards (e.g., APA, 2008), guidelines (Simmons et al., 2011), or checklists (Nosek et al., 2012) could help in the implementation. Although available standards (e.g., APA, 2008) mention several of the issues we present, our journals tend not to adhere to them. According to Cooper and VandenBos (2013), the APA left it up to their journal editors to implement the standards. Apparently, most editors elected not to enforce them. Integrating our own standards in checklists into the editorial review process should make the critical appraisal easier and increase the trustworthiness of our published results.

Alternative Approaches to the Editorial Review Process

Given the importance of the editorial review process for our sciences, it is odd that little is known about it (Jefferson, Alderson, Wager, & Davidoff, 2002). Also, what is known about the process raises serious questions about its validity, value, and the trustworthiness of the published results (Miner, 2003; Pfeffer, 2007; Starbuck, 2005). The current editorial review system may introduce biases, HARKing, and other questionable research practices into the publication process (Bedeian et al., 2010; Kepes & McDaniel, 2013; O'Boyle et al., in press; Rupp, 2011; Starbuck, 1994). For instance, reviewers and editors may ask authors to drop nonsupported hypotheses and tangential or "uninteresting" results. Alternative theories, models, or hypotheses are also sometimes introduced during the editorial review process (Rupp, 2011). Researchers in our field should stop their tendency to write and publish manuscripts with significant results while suppressing the availability of insignificant findings (Fanelli, 2012; Sterling & Rosenbaum, 1995) as well as engaging in questionable research practices to fit data to a theory (Bedeian et al., 2010; Kepes & McDaniel, 2013; O'Boyle et al., in press). To do this, alternative approaches to the editorial review process need to be considered that increase the transparency of the process and the trustworthiness of our published literature.

Two-Stage Review Processes

One approach to improving the editorial review process is to separate the overall review process into two stages (Kepes & McDaniel, 2013; Smulders, 2013). In the first, authors would only submit the Introduction, Theory, and Methods sections. Reviewers and editors could solely focus on the theory development, the design of the study, the measurement and the analysis approach, and the potential contribution to the literature. Article reviews without any knowledge of the results reduces the biases that affect reviewers when evaluating the contribution of manuscripts (Epstein, 1990; Mahoney, 1977).

If the manuscript survives the first review stage, the author would submit the entire manuscript, including the Results and Discussion sections. This second stage of the review process is straightforward. Only three aspects would have to be evaluated: (1) did the author carry out the methodological and analytic approach described during the first stage of the review process, (2) is the description of the results accurate, and (3) are the conclusions in the discussion section accurate and actionable? Thus, this stage of the review process would be very objective and could be conducted in a relatively short time frame. The benefits of this two-stage process outweigh costs as the evaluation of soundness and contribution to the literature (stage one) would be unaffected by the results of the study, which affect decisions during the editorial review process (Epstein, 1990; Greenwald, 1975;

Mahoney, 1977). Because this process reduces the impact of statistical significance on acceptance decisions, it should reduce authors' motivation to engage in questionable research practices. That is, if authors could be confident that their results, whether statistically significant or not, have no bearing on the decision to accept or reject a manuscript, they may feel less pressure to engage in HARKing and other questionable research practices to obtain statistically significant results and get their manuscript published.

"Because this process reduces the impact of statistical significance on acceptance decisions, it should reduce authors' motivation to engage in questionable research practices."

Multistage Review Processes

An approach used in some journals, such as the prestigious *Atmospheric Chemistry and Physics* journal,[11] is the multistage open peer-review process (Pöschl, 2012). In the first stage, after an initial screening by the editor, submitted manuscripts are posted on-line as discussion papers. For several weeks, comments and answers of traditional reviewers, the authors, and the scientific community are posted on-line with the paper. The next stage mirrors the traditional editorial review process. This multistage review process offers numerous advantages (see Pöschl, 2012), including rapid dissemination of a manuscript's results, including open discussion with transparent and direct feedback. The high degree of transparency during the open review process should increase the perceived trustworthiness of the final product, the published article.

Other Means to Improve the Trustworthiness of Our Published Research Findings

Other recommendations include changes in the assessment of article quality (e.g., Adler & Harzing, 2009), the reward structure (e.g., Nosek et al., 2012), the way we archive and share our data (e.g., Hanson, Sugden, & Alberts, 2011) and the establishment of research registries (e.g., Kepes et al., 2012). Many of these and related recommendations may take a substantial amount of time to implement. Consistent with previous calls (Ferguson & Brannick, 2012; Kepes et al., 2012), we suggest the establishment of mandatory research registries[12] in management because they have been shown to increase the availability of research findings (Zarin, Tse, Williams, Califf, & Ide, 2011), and thus, the trustworthiness of cumulative scientific knowledge in the medical sciences (Dickersin, 1990).

Translating the Scientific Evidence

To better aid the EBMgt movement, we need better translations of our trustworthy cumulative knowledge because educators are unlikely to critically appraise all the evidence we produce and publish (Rousseau & McCarthy, 2007), and many practitioners across cultures do not read or understand our published research (Pearce & Huang, 2012b; Rynes, Colbert, & Brown, 2002; Tenhiälä et al., in press). Indeed, practitioners and academics live in separate worlds and speak different languages (Giluk & Rynes, 2012), which contributes to the research–practice gap. We need a sound translation process of our trustworthy scientific evidence to educators and practitioners; otherwise, they may not be able to separate weak from strong research evidence. According to Cummings (2007: 357), actionable knowledge "must transcend purely scientific concerns and address specific problems facing practitioners, the actions they can take to solve them, and the change process for making all this happen."

This is a key purpose of EBMgt. Recommendations and guidelines for how to translate academic research into practically relevant and actionable research are available (Gruber, 2006; Latham, 2007; Pfeffer & Sutton, 2007; Sommer, 2006). We thus encourage researchers to create more trustworthy actionable knowledge (Pearce & Huang, 2012a). A vital and integral part of this translation process is

[11] According to Thomson Reuters Journal Citation Reports, the *Atmospheric Chemistry and Physics* journal has the second highest impact factor in the Meteorology and Atmospheric Sciences (Journal Citation Reports Science Edition, 2011).

[12] Several research registries exist in the medical sciences; the largest registry in use is ClinicalTrials.gov (Dickersin & Rennie, 2003, 2012; Kepes & McDaniel, 2013; Stetz & Subramony, 2013). In 2005, the *International Committee of Medical Journal Editors* made the registration of clinical trials before they are conducted a requirement for publication in the associated journals (Laine et al., 2007).

the more effective use of graphic displays to effectively communicate our results. Unfortunately, we rarely use such displays effectively in our articles (Kepes et al., 2013).

Recent developments in our discipline such as the *Society for Industrial and Organizational Psychology*'s (www.siop.org) Leading Edge Consortium series, which brings together scientists and practitioners to discuss a single practitioner-relevant topic, and the new *Academy of Management Discoveries* journal could be important steps.[13] However, for such developments to be effective and succeed, the reward structure needs to provide incentives for translational research, which it currently does not (Adler & Harzing, 2009; Nosek et al., 2012).

In addition to changes in the scientific publication process, we recommend a close examination of our approach to teaching. Unfortunately, there are many areas that may not use the best available evidence for practice and in the classroom. As an example, only around 25% of the core courses in management are evidence-based at the graduate level, and even fewer classes actually discuss evidence-based concepts and practices (Charlier, Brown, & Rynes, 2011), which shows that an examination of our teaching must include a critical evaluation of our textbooks. Indeed, some of our textbooks do not always contain trustworthy evidence, which makes them inadequate to teach from an EBMgt approach (Goodman & O'Brian, 2012; Pearce, 2012a).

CONCLUSIONS

The purpose of this article was to provide a review of the trustworthiness of the our cumulative scientific knowledge and highlight the effect that an untrustworthy knowledge base may have on teaching from an EBMgt perspective. Several scientific disciplines have questioned the accuracy of their published findings. Given the importance of EBMgt, it is somewhat surprising that an assessment of the trustworthiness of the management literature appears to be missing. Also, because practitioners often do not embrace the scientific process, they do not use much of our published evidence (Giluk & Rynes, 2012). Although EBMgt does not solely depend on trustworthy scientific evidence, such evidence is an integral and fundamental part of it (Briner et al., 2009; Pfeffer & Sutton, 2006; Rousseau, 2006; Satterfield et al., 2009).

We note that some of our recommendations may not adequately address all potential threats to the trustworthiness of our cumulative knowledge and ensure that EBMgt continues to grow and be successful. Our recommendations require changes in the way we teach and conduct research. There are costs to many of our suggestions, but these changes are needed to ensure the success of EBMgt. Psychologists are generally supportive of changes in the scientific process (Fuchs, Jenny, & Fiedler, 2012), and we offer that it may not be different in management.

Do we have the will to resist our hunt for the fashionable (Witten & Tibshirani, 2013), newsworthy (Hartshorne & Schachner, 2012), and statistically significant (Ferguson & Heene, 2012), and, instead, focus on bringing the most trustworthy scientific knowledge to the classroom and the workplace? This is happening in the medical sciences, after decades of criticism of the practices of physicians and no increase in life expectancies despite an explosion in healthcare costs (Barends et al., 2012). After talking for almost a decade about the necessity and virtues of EBMgt, we are confident that we can follow the medical sciences and make the necessary changes to the generation and dissemination of trustworthy cumulative knowledge. However, because the implementation of some of the recommendations may take a substantial amount of time (it took around 2 decades in the medical sciences; Barends et al., 2012), we recommend a stronger emphasis on the teaching of EBMgt and thus the use of critically appraised evidence in our classrooms.

[13] The mission statement of the *Academy of Management Discoveries* (see http://aom.org/amd/) includes several of the recommendations we discussed in this manuscript. The degree of adherence and implementation remains to be seen. Furthermore, we hope that additional recommendations mentioned by us or other researchers will get incorporated into the mission statement.

REFERENCES

Abrahamson, E., & Eisenman, M. 2001. Why management scholars must intervene strategically in the management knowledge market. ***Human Relations,*** 54: 67–75.

Adler, N. J., & Harzing, A.-W. 2009. When knowledge wins: Transcending the sense and nonsense of academic rankings. ***Academy of Management Learning and Education,*** 8: 72–95.

Amabile, T. M., Patterson, C., Mueller, J., Wojcik, T., Odomirok, P. W., Marsh, M., & Kramer, S. J. 2001. Academic-practitioner collaboration in management research: A case of cross-

profession collaboration. ***Academy of Management Journal,*** 44: 418–431.

American Psychological Association. 2008. Reporting standards for research in psychology: Why do we need them? What might they be? ***American Psychologist,*** 63: 839–851.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., van Aken, M. A. G., Weber, H., & Wicherts, J. M. 2013. Recommendations for increasing replicability in psychology. ***European Journal of Personality,*** 27: 108–119.

Baba, V. V., & HakemZadeh, F. 2012. Toward a theory of evidence based decision making. ***Management Decision,*** 50: 832–867.

Banks, G. C., Kepes, S., & Banks, K. P. 2012a. Publication bias: The antagonist of meta-analytic reviews and effective policy making. ***Educational Evaluation and Policy Analysis,*** 34: 259–277.

Banks, G. C., Kepes, S., & McDaniel, M. A. 2012b. Publication bias: A call for improved meta-analytic practice in the organizational sciences. ***International Journal of Selection and Assessment,*** 20: 182–196.

Barends, E., Janssen, B., ten Have, W., & ten Have, S. 2014. Effects of change interventions: What kind of evidence do we really have? ***The Journal of Applied Behavioral Science,*** 50: 5–27.

Barends, E. G. R., ten Have, S., & Huisman, F. 2012. Learning from other evidence-based practices: The case of medicine. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management:*** 25–42. New York: Oxford University Press.

Bedeian, A. G., Taylor, S. G., & Miller, A. N. 2010. Management science on the credibility bubble: Cardinal sins and various misdemeanors. ***Academy of Management Learning and Education,*** 9: 715–725.

Begley, C. G., & Ellis, L. M. 2012. Drug development: Raise standards for preclinical cancer research. ***Nature,*** 483: 531–533.

Bellamy, J. L., Bledsoe, S. E., & Traube, D. E. 2006. The current state of evidence-based practice in social work: A review of the literature and qualitative analysis of expert interviews. ***Journal of Evidence-Based Social Work,*** 3: 23–48.

Bem, D. J. 2011. Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. ***Journal of Personality and Social Psychology,*** 100: 407–425.

Berry, C. M., Sackett, P. R., & Tobares, V. 2010. A meta-analysis of conditional reasoning tests of aggression. ***Personnel Psychology,*** 63: 361–384.

Berwick, D. M. 2008. The science of improvement. ***Journal of the American Medical Association,*** 299: 1182–1184.

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. 2009. Strong claims and weak evidence: Reassessing the predictive validity of the IAT. ***Journal of Applied Psychology,*** 94: 567–582.

Briner, R. B., Denyer, D., & Rousseau, D. M. 2009. Evidence-based management: Concept cleanup time? ***Academy of Management Perspectives,*** 23: 19–32.

Briner, R. B., & Rousseau, D. M. 2011. Evidence-based I–O psychology: Not there yet but now a little nearer? ***Industrial and Organizational Psychology: Perspectives on Science and Practice,*** 4: 76–82.

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., Fervers, B., Graham, I. D., Grimshaw, J., Hanna, S. E., Littlejohns, P., Makarski, J., & Zitzelsberger, L. 2010. AGREEII: Advancing guideline development, reporting and evaluation in health care. ***Canadian Medical Association Journal,*** 182: E839–E842.

Buskist, W., & Groccia, J. E. 2011. Evidence-based teaching: Now and in the future. ***New Directions for Teaching and Learning,*** 128: 105–111.

Byrne, D. W. 2000. Common reasons for rejecting manuscripts at medical journals: A survey of editors and peer reviewers. ***Science Editor,*** 23: 39–44.

Chalmers, I. 1990. Underreporting research is scientific misconduct. ***Journal of the American Medical Association,*** 263: 1405–1408.

Charlier, S. D., Brown, K. G., & Rynes, S. L. 2011. Teaching evidence-based management in MBA programs: What evidence is there? ***Academy of Management Learning and Education,*** 10: 222–236.

Cochrane, A. L. 1972. ***Effectiveness and efficiency: Random reflections on health services.*** London, UK: Nuffield Provincial Hospitals Trust.

Cooper, H., Hedges, L. V., & Valentine, J. C. 2009. ***The handbook of research synthesis and meta-analysis*** (2nd ed.). New York, NY: Russell Sage Foundation.

Cooper, H., & VandenBos, G. R. 2013. Archives of scientific psychology: A new journal for a new era. ***Archives of Scientific Psychology,*** 1: 1–6.

Cronin, M. A., & Klimoski, R. 2011. Broadening the view of what constitutes "evidence". ***Industrial and Organizational Psychology,*** 4: 57–61.

Cummings, T. G. 2007. Quest for an engaged Academy. ***Academy of Management Review,*** 32: 355–360.

Dalton, D. R., Aguinis, H., Dalton, C. M., Bosco, F. A., & Pierce, C. A. 2012. Revisiting the file drawer problem in a meta-analysis: An assessment of published and nonpublished correlation matrices. ***Personnel Psychology,*** 65: 221–249.

Davenport, T. H., & Harris, J. G. 2007. ***Competing on analytics: The new science of winning.*** Boston, MA: Harvard Business School Press.

Dickersin, K. 1990. The existence of publication bias and risk factors for its occurrence. ***Journal of the American Medical Association,*** 263: 1385–1389.

Dickersin, K., & Rennie, D. 2003. Registering clinical trials. ***Journal of the American Medical Association,*** 290: 516–523.

Dickersin, K., & Rennie, D. 2012. The evolution of trial registries and their use to assess the clinical trial enterprise. ***Journal of the American Medical Association,*** 307: 1861–1864.

Doucouliagos, H., & Stanley, T. D. 2009. Publication selection bias in minimum-wage research? A meta-regression analysis. ***British Journal of Industrial Relations,*** 47: 406–428.

Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. 2012. Be-

havioral priming: It's all in the mind, but whose mind? ***PLOS ONE,*** 7: e29081.

Earl-Slater, A. 2001. Critical appraisal and hierarchies of the evidence. ***British Journal of Clinical Governance,*** 6: 59–63.

The Economist. 2013. Trouble at the lab: Unreliable research. ***The Economist,*** 409: 26–30.

Epstein, W. M. 1990. Confirmational response bias among social work journals. ***Science, Technology and Human Values,*** 15: 9–38.

Estes, S. 2012. The myth of self-correcting science. ***Atlantic.*** Retrieved from http://www.theatlantic.com/health/archive/2012/12/the-myth-of-self-correcting-science/266228/. March 14, 2013.

Fanelli, D. 2012. Negative results are disappearing from most disciplines and countries. ***Scientometrics,*** 90: 891–904.

Feinstein, A. R. 1967. ***Clinical judgment.*** Baltimore, MD: Williams and Wilkins.

Ferguson, C. J., & Brannick, M. T. 2012. Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. ***Psychological Methods,*** 17: 120–128.

Ferguson, C. J., & Heene, M. 2012. A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. ***Perspectives on Psychological Science,*** 7: 555–561.

Francis, G. 2012. Too good to be true: Publication bias in two prominent studies from experimental psychology. ***Psychonomic Bulletin and Review,*** 19: 151–156.

Francis, G. 2013. Publication bias in "Red, rank, and romance in women viewing men," by Elliot et al. (2010). ***Journal of Experimental Psychology: General,*** 142: 292–296.

Fuchs, H. M., Jenny, M., & Fiedler, S. 2012. Psychologists are open to change, yet wary of rules. ***Perspectives on Psychological Science,*** 7: 639–642.

Gerhart, B., & Rynes, S. L. 2003. ***Compensation: Theory, evidence, and strategic implications.*** Thousand Oaks, CA: Sage.

Gibbons, J., & Woock, C. 2007. Evidence-based human resources: A primer and summary of current literature. ***The Conference Board,*** Research Report E-0015-07-RR.

Gigerenzer, G., & Gaissmaier, W. 2011. Heuristic decision making. ***Annual Review of Psychology,*** 62: 451–482.

Gillath, O., Sesko, A. K., Shaver, P. R., & Chun, D. S. 2010. Attachment, authenticity, and honesty: Dispositional and experimentally induced security can reduce self- and other-deception. ***Journal of Personality and Social Psychology,*** 98: 841–855.

Giluk, T. L., & Rynes, S. L. 2012. Research findings practitioners resist. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management.*** New York: Oxford University Press.

Glynn, L. 2006. A critical appraisal tool for library and information research. ***Library Hi Tech,*** 24: 387–399.

Goodman, J. S., & O'Brian, J. 2012. Teaching and learning using evidence-based principles. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management:*** 309–336. New York, NY: Oxford University Press.

Greenhalgh, T. 1997. How to read a paper. Getting your bearings (deciding what the paper is about). ***British Medical Journal,*** 315: 243–246.

Greenhouse, J. B., & Iyengar, S. 2009. Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), ***The handbook of research synthesis and meta-analysis:*** 417–433 (2nd ed.). New York, NY: Russell Sage Foundation.

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

Gruber, D. A. 2006. The craft of translation: An interview with Malcolm Gladwell. ***Journal of Management Inquiry,*** 15: 397–403.

Guskey, T. R. 2007. Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning. ***Educational Measurement: Issues and Practice,*** 26: 19–27.

Guyatt, G. H., Sackett, D. L., Sinclair, J. C., Hayward, R., Cook, D. J., & Cook, R. J. 1995. Users' guides to the medical literature: IX. A method for grading health care recommendations. ***The Journal of the American Medical Association,*** 274: 1800–1804.

Hambrick, D. C. 2007. The field of management's devotion to theory: Too much of a good thing? ***Academy of Management Journal,*** 50: 1348–1352.

Hanson, B., Sugden, A., & Alberts, B. 2011. Making data maximally available. ***Science,*** 331: 649.

Hartshorne, J., & Schachner, A. 2012. Tracking replicability as a method of post-publication open evaluation. ***Frontiers in Computational Neuroscience,*** 6: 8.

Higgins, J. P. T., & Green, S., (Eds.), 2011. ***Cochrane handbook for systematic reviews of interventions,*** version 5.1.0. The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org.

Hintermann, B., Alberini, A., & Markandya, A. 2010. Estimating the value of safety with labour market data: Are the results trustworthy? ***Applied Economics,*** 42: 1085–1100.

Homberg, F., & Bui, H. T. M. 2013. Top management team diversity: A systematic review. ***Group and Organization Management,*** 38: 455–479.

Horsburgh, D. 2003. Evaluation of qualitative research. ***Journal of Clinical Nursing,*** 12: 307–312.

Hunter, J. E., & Schmidt, F. L. 2004. ***Methods of meta-analysis: Correcting error and bias in research findings*** (2nd ed.). Newbury Park, CA: Sage.

Ioannidis, J. P. 1998. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. ***The Journal of the American Medical Association,*** 279: 281–286.

Ioannidis, J. P. 2005a. Contradicted and initially stronger effects in highly cited clinical research. ***The Journal of the American Medical Association,*** 294: 218–228.

Ioannidis, J. P. 2005b. Why most published research findings are false. ***PLOS Medicine,*** 2: e124.

Ioannidis, J. P., Allison, D. B., Ball, C. A., Coulibaly, I., Cui, X., Culhane, A. C., Falchi, M., Furlanello, C., Game, L., Jurman,

G., Mangion, J., Mehta, T., Nitzberg, M., Page, G. P., Petretto, E., & van Noort, V. 2009. Repeatability of published microarray gene expression analyses. ***Nature Genetics,*** 41: 149–155.

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., Frost, B. C., Russell, S. M., Sablynski, C. J., Mitchell, T. R., & Williams, L. J. 2005. A conditional reasoning measure for aggression. ***Organizational Research Methods,*** 8: 69–99.

Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. 2011. Data replication & reproducibility. Again, and again, and again . . . . . ***Science,*** 334: 1225.

Jefferson, T., Alderson, P., Wager, E., & Davidoff, F. 2002. Effects of editorial peer review: A systematic review. ***Journal of the American Medical Association,*** 287: 2784–2786.

Jelley, R. B., Caroll, W. R., & Rousseau, D. M. 2012. Reflections of teaching evidence-based management. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management:*** 337–355. New York, NY: Oxford University Press.

Jones, S. C., Crookes, P. A., & Johnson, K. M. 2011. Teaching critical appraisal skills for nursing research. ***Nurse Education in Practice,*** 11: 327–332.

Journal Citation Reports Science. 2011. ***ISI Web of Knowledge Database***, Thomson Reuters.

Kahneman, D. 2011. ***Thinking, fast and slow***. New York: Farrar, Strauss & Giroux.

Kang, J. 2005. Trojan horses of race. ***Harvard Law Review,*** 118: 1489–1593.

Kang, J., & Banaji, M. R. 2006. Fair measures: A behavioral realist revision of "affirmative action". ***California Law Review,*** 94: 1063–1118.

Keim, S. M., Howse, D., Bracke, P., & Mendoza, K. 2008. Promoting evidence based medicine in preclinical medical students via a federated literature search tool. ***Medical Teacher,*** 30: 880–884.

Kepes, S., Banks, G., & Oh, I.-S. 2014. Avoiding bias in publication bias research: The value of "null" findings. ***Journal of Business and Psychology,*** 29: 183–203.

Kepes, S., Banks, G. C., McDaniel, M. A., & Whetzel, D. L. 2012. Publication bias in the organizational sciences. ***Organizational Research Methods,*** 15: 624–662.

Kepes, S., & McDaniel, M. A. 2013. How trustworthy is the scientific literature in industrial and organizational psychology? ***Industrial and Organizational Psychology: Perspectives on Science and Practice,*** 6: 252–268.

Kepes, S., McDaniel, M. A., Brannick, M. T., & Banks, G. C. 2013. Meta-analytic reviews in the organizational sciences: Two meta-analytic schools on the way to MARS (the Meta-analytic Reporting Standards). ***Journal of Business and Psychology,*** 28: 123–143.

Kerr, N. L. 1998. HARKing: Hypothesizing after the results are known. ***Personality and Social Psychology Review,*** 2: 196–217.

Laine, C., Horton, R., DeAngelis, C. D., Drazen, J. M., Frizelle, F. A., Godlee, F., Haug, C., Hébert, P. C., Kotzin, S., Marusic, A., Sahni, P., Schroeder, T. V., Sox, H. C., Van der Weyden, M. B., & Verheugt, F. W. A. 2007. Clinical trial registration: Looking back and moving ahead. ***New England Journal of Medicine,*** 356: 2734–2736.

Latham, G. P. 2007. A speculative perspective on the transfer of behavioral science findings to the workplace: "The times they are a-changin". ***Academy of Management Journal,*** 50: 1027–1032.

Latham, G. P., & Locke, E. A. 2009. Science and ethics: What should count as evidence against the use of goal setting? ***Academy of Management Perspectives,*** 23: 88–91.

Le, H., Oh, I.-S., Shaffer, J., & Schmidt, F. L. 2007. Implications of methodological advances for the practice of personnel selection: How practitioners benefit from meta-analysis. ***Academy of Management Perspectives,*** 21: 6–15.

LeBel, E. P., & Peters, K. R. 2011. Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. ***Review of General Psychology,*** 15: 371–379.

Lehrer, J. 2010. The truth wears off. ***New Yorker,*** 86: 52–57.

Leung, O., & Bartunek, J. M. 2012. Enabling evidence-based management: Bridging the gap between academics and practitioners. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management.*** New York, NY: Oxford University Press.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. ***British Medical Journal,*** 339: b2700.

Macey, W. H., & Schneider, B. 2008. The meaning of employee engagement. ***Industrial and Organizational Psychology: Perspectives on Science and Practice,*** 1: 3–30.

Mahoney, M. J. 1977. Publication prejudices: An experimental study of confirmatory bias in the peer review system. ***Cognitive Therapy and Research,*** 1: 161–175.

Makel, M. C., Plucker, J. A., & Hegarty, B. 2012. Replications in psychology research: How often do they really occur? ***Perspectives on Psychological Science,*** 7: 537–542.

McConnell, A. R., & Leibold, J. M. 2001. Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. ***Journal of Experimental Social Psychology,*** 37: 435–442.

McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. 2006. Publication bias: A case study of four test vendors. ***Personnel Psychology,*** 59: 927–953.

McKinley, W. 2010. Organizational theory development: Displacement of ends? ***Organization Studies,*** 31: 47–68.

Mears, D. P., & Barnes, J. C. 2010. Toward a systematic foundation for identifying evidence-based criminal justice sanctions and their relative effectiveness. ***Journal of Criminal Justice,*** 38: 702–710.

Melnyk, B. M., & Fineout-Overholt, E. 2011. ***Evidence-based practice in nursing and healthcare: A guide to best practice*** (2nd ed.). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.

Merton, R. K. 1973. ***The sociology of science: Theoretical and empirical investigations.*** Chicago, IL: The University of Chicago Press.

Miner, J. B. 2003. Commentary on Arthur Bedeian's "The manu-

script review process: The proper roles of authors, referees, and editors". ***Journal of Management Inquiry,*** 12: 339–343.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. ***The Lancet,*** 354: 1896–1900.

Neuliep, J. W., & Crandall, R. 1990. Editorial bias against replication research. ***Journal of Social Behavior and Personality,*** 5: 85–90.

Nosek, B. A., Spies, J. R., & Motyl, M. 2012. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. ***Perspectives on Psychological Science,*** 7: 615–631.

O'Boyle, E. H., Banks, G. C., & Gonzalez-Mule, E. In press. The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. ***Journal of Management***.

O'Boyle, E. H., Banks, G. C., Rutherford, M. W. In press. Publication bias in entrepreneurship research: An examination of dominant relations to performance. ***Journal of Business Venturing***.

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. 2013. Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. ***Journal of Personality and Social Psychology,*** 105: 171–192.

Ouellette, J. 2013. The mathematical shape of things to come: Scientific data sets are becoming more dynamic, requiring new mathematical techniques on par with the invention of calculus. ***Quanta Magazine.*** Retrieved from https://www.simonsfoundation.org/quanta/20131004-the-mathematical-shape-of-things-to-come/. October 14, 2013.

Pearce, J. L. 2012a. Creating evidence-based management textbooks. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management:*** 377–386. New York: Oxford University Press.

Pearce, J. L. 2012b. ***Organizational behavior: Real research for real managers*** (3rd ed.). Irvine, CA: Melvin & Leigh.

Pearce, J. L., & Huang, L. 2012a. The decreasing value of our research to management education. ***Academy of Management Learning and Education,*** 11: 247–262.

Pearce, J. L., & Huang, L. 2012b. Toward an understanding of what actionable research is. ***Academy of Management Learning and Education,*** 11: 300–301.

Pedhazur, E. J., & Pedhazur Schmelkin, L. 1991. ***Measurement, design, and analysis: An integrated approach.*** New York: Psychology Press.

Petticrew, M., & Roberts, H. 2003. Evidence, hierarchies, and typologies: Horses for courses. ***Journal of Epidemiology and Community Health,*** 57: 527–529.

Pfeffer, J. 2007. A modest proposal: How we might change the process and product of managerial research. ***Academy of Management Journal,*** 50: 1334–1345.

Pfeffer, J., & Sutton, A. 2011. Trust the evidence, not your instincts. ***New York Times.*** Retrieved from http://www.nytimes.com/2011/09/04/jobs/04pre.html. September 21, 2013.

Pfeffer, J., & Sutton, R. I. 2006. ***Hard facts, dangerous half-truths and total nonsense: Profiting from evidence-based management.*** Boston, MA: Harvard Business School Press.

Pfeffer, J., & Sutton, R. I. 2007. Suppose we took evidence-based management seriously: Implications for reading and writing management. ***Academy of Management Learning and Education,*** 6: 153–155.

Pigott, T. D., Valentine, J. C., Polanin, J. R., Williams, R. T., & Canada, D. D. 2013. Outcome-reporting bias in education research. ***Educational Researcher,*** 42: 424–432.

Pillutla, M. M., & Thau, S. 2013. Organizational sciences' obsession with "that's interesting!": Consequences and an alternative. ***Organizational Psychology Review,*** 3: 187–194.

Ployhart, R. E., & Vandenberg, R. J. 2010. Longitudinal research: The theory, design, and analysis of change. ***Journal of Management,*** 36: 94–120.

Pöschl, U. 2012. Multi-stage open peer review: Scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation. ***Frontiers in Computational Neurosciences,*** 6: 33.

Pratt, M. G. 2008. Fitting oval pegs into round holes: Tensions in evaluating and publishing qualitative research in top-tier North American journals. ***Organizational Research Methods,*** 11: 481–509.

Reay, T., Berta, W., & Kohn, M. K. 2009. What's the evidence on evidence-based management? ***Academy of Management Perspectives,*** 23: 5–18.

Renkewitz, F., Fuchs, H. M., & Fiedler, S. 2011. Is there evidence of publication biases in JDM research? ***Judgment and Decision Making,*** 6: 870–881.

Ritchie, S. J., Wiseman, R., & French, C. C. 2012. Failing the future: Three unsuccessful attempts to replicate Bem's "retroactive facilitation of recall" effect. ***PLOS ONE,*** 7: e33423.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. 2005. ***Publication bias in meta-analysis: Prevention, assessment, and adjustments.*** West Sussex, UK: Wiley.

Rousseau, D. M. 2006. Is there such a thing as "evidence-based management"? ***Academy of Management Review,*** 31: 256–269.

Rousseau, D. M. 2012a. Envisioning evidence-based management. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management.*** New York: Oxford University Press.

Rousseau, D. M. 2012b. Organizational behavior's contributions to evidence-based management. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management.*** New York: Oxford University Press.

Rousseau, D. M., & McCarthy, S. 2007. Educating managers from an evidence-based perspective. ***Academy of Management Learning and Education,*** 6: 84–101.

Rupp, D. E. 2011. Ethical issues faced by editors and reviewers. ***Management and Organization Review,*** 7: 481–493.

Rynes, S. L., Colbert, A. E., & Brown, K. G. 2002. HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. ***Human Resource Management,*** 41: 149–174.

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. 2000. ***Evidence-based medicine: How to practice and teach EBM*** (2nd ed.). London, UK: Churchill Livingstone.

Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A., & Chal-

mers, T. C. 1987. Meta-analyses of randomized controlled trials. ***New England Journal of Medicine,*** 316: 450–455.

Sadigh, G., Parker, R., Kelly, A. M., & Cronin, P. 2012. How to write a critically appraised topic (CAT). ***Academic Radiology,*** 19: 872–888.

Satterfield, J. M., Spring, B., Brownson, R. C., Mullen, E. J., Newhouse, R. P., Walker, B. B., & Whitlock, E. P. 2009. Toward a transdisciplinary model of evidence-based practice. ***Milbank Quarterly,*** 87: 368–390.

Sauve, S., Lee, H. N., Meade, M. O., Lung, J. D., Farkouh, M., Cook, D. J., & Sackett, D. L. 1995. The critically appraised topic: A practical approach to learning critical appraisal. ***Annals of the Royal Society of Physicians and Surgeons of Canada,*** 28: 396–398.

Schimmack, U. 2012. The ironic effect of significant results on the credibility of multiple-study articles. ***Psychological Methods,*** 17: 551–566.

Shapiro, D. L., Kirkman, B. L., & Courtney, H. G. 2007. Perceived causes and solutions of the translation problem in management research. ***Academy of Management Journal,*** 50: 249–266.

Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., Henry, D. A., & Boers, M. 2009. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. ***Journal of Clinical Epidemiology,*** 62: 1013–1020.

Silver, N. 2012. ***The signal and the noise: Why so many predictions fail—But some don't.*** New York: Penguin.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. ***Psychological Science,*** 22: 1359–1366.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. 2014. *P*-curve: A key to the file drawer. ***Journal of Experimental Psychology: General,*** 143: 534–547.

Smith, J. 2013, March 8. America's 100 most trustworthy companies. ***Forbes***: 3. Retrieved from http://www.forbes.com/sites/jacquelynsmith/2013/03/18/americas-100-most-trustworthy-companies/. September 17, 2013.

Smulders, Y. M. 2013. A two-step manuscript submission process can reduce publication bias. ***Journal of Clinical Epidemiology,*** 66: 946–947.

Sommer, R. 2006. Dual dissemination: Writing for colleagues and the public. ***American Psychologist,*** 61: 955–958.

Stanovich, K. E., & West, R. F. 2000. Individual differences in reasoning: Implications for the rationality debate? ***Behavioral and Brain Sciences,*** 23: 645–665.

Starbuck, W. H. 1994. On behalf of naïveté. In J. A. C. Baum & J. V. Singh (Eds.), ***Evolutionary dynamics of organizations:*** 205–220. New York: Oxford University Press.

Starbuck, W. H. 2005. How much better are the most-prestigious journals? The statistics of academic publication. ***Organization Science,*** 16: 180–200.

Sterling, T. D., & Rosenbaum, W. L. 1995. Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. ***American Statistician,*** 49: 108–112.

Stetz, T. A., & Subramony, M. 2013. Research registries and trustworthiness of industrial-organizational psychological research. ***Industrial and Organizational Psychology,*** 6: 302–305.

Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. 2011. ***Evidence-based medicine: How to practice and teach it*** (4th ed.). London, UK: Churchill Livingstone.

Sutton, A. J. 2005. Evidence concerning the consequences of publication and related biases. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), ***Publication bias in meta-analysis: Prevention, assessment, and adjustments:*** 175–192. West Sussex, UK: Wiley.

Tenhiälä, A., Giluk, T. L., Kepes, S., Simón, C., Oh, I.-S., & Kim, S. In press. The research-practice gap in human resource management: A cross-cultural study. ***Human Resource Management***.

Trikalinos, T. A., & Ioannidis, J. P. A. 2005. Assessing the evolution of effect sizes over time. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), ***Publication bias in metaanalysis: Prevention, assessment and adjustments:*** 241–259. West Sussex, UK: Wiley.

van Aken, J. E., & Romme, G. L. 2012. A design science approach to evidence-based management. In D. M. Rousseau (Ed.), ***The Oxford handbook of evidence-based management:*** 140–184. New York: Oxford University Press.

Vul, E., Harris, C., Winkielman, P., & Pashler, H. 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. ***Perspectives on Psychological Science,*** 4: 274–290.

Witten, D. M., & Tibshirani, R. 2013. Scientific research in the age of omics: The good, the bad, and the sloppy. ***Journal of the American Medical Informatics Association:*** 20: 125–127.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. 2011. Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). ***Journal of Personality and Social Psychology,*** 100: 426–432.

Wagner, W., & Steinzor, R. 2006. ***Rescuing science from politics: Regulation and the distortion of scientific research.*** New York: Cambridge University Press.

Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., & Bossuyt, P. M. M.; QUADAS-2 Group. 2011. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. ***Annals of Internal Medicine,*** 155: 529–536.

Wicherts, J. M., Bakker, M., & Molenaar, D. 2011. Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. ***PLOS ONE,*** 6: e26828.

Yong, E. 2012. Replication studies: Bad copy. ***Nature,*** 485: 298–300.

Zarin, D. A., Tse, T., Williams, R. J., Califf, R. M., & Ide, N. C. 2011. The ClinicalTrials.gov results database–update and key issues. ***New England Journal of Medicine,*** 364: 852–860.

Ziegert, J. C., & Hanges, P. J. 2005. Employment discrimination: The role of implicit attitudes, motivation, and a climate for racial bias. ***Journal of Applied Psychology,*** 90: 553–562.

Sven Kepes is an assistant professor of management at Virginia Commonwealth University. Kepes received his PhD from the University of Arkansas. His research interests include staffing, compensation, strategic human resource management, and research methods, particularly meta-analysis and publication bias. His research has been published in several journals, including the *Journal of Applied Psychology*, *Personnel Psychology*, and *Organizational Research Methods*.

Andrew A. Bennett is a doctoral candidate in the Department of Management at Virginia Commonwealth University. Bennett received his master's degree in organizational leadership from Gonzaga University. His research focuses on employee well-being and evidence-based management.

Michael A. McDaniel is a professor of management at Virginia Commonwealth University. McDaniel's research focusses on personnel selection, the application of meta-analysis to management research areas, and publication bias. His research has been published in several journals, including the *Journal of Applied Psychology*, *Personnel Psychology*, *Intelligence*, and *Organizational Research Methods*.