

# Can we trust the scientific literature in Industrial and Organizational Psychology?

**Paul Barrett**

Chief Research Scientist, Cognadev UK & SA Ltd.  
Honorary Prof. of Psychology, University of Auckland, NZ  
Adjunct Prof. of Psychology, University of Canterbury, NZ



## The Target Article

Kepes, S., & McDaniel, M.A. (2013). *How trustworthy is the scientific literature in I-O psychology? Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 252-268.

### *from the Abstract*

“This article calls for a review of the trustworthiness of the scientific literature in industrial–organizational (I–O) psychology and a reconsideration of common practices that may harm the credibility of our literature. We note that most hypotheses in I–O psychology journals are confirmed. Thus, we are either approaching omniscience or our journals are publishing an unrepresentative sample of completed research.”

## The Target Article

“... In a review of psychology journals, Sterling (1959) found that 97% of the articles rejected the null hypothesis. In a replication more than 36 years later, Sterling and Rosenbaum (1995) reported essentially identical results. We offer these findings as evidence supporting the inference of structural problems in the psychology scientific process that should also affect I–O psychology.” *p. 254*

## The Target Article

“... Note that we are not arguing that our literature is entirely composed of zero-magnitude population effect sizes that are falsely presented as nonzero. Rather, we are arguing that it is common for the magnitude of our population effect sizes to be mis-estimated, often overestimated (Kepes, Banks, McDaniel, & Whetzel, 2012). Thus, although authors and journals seek to improve science, our actions in the “chase for significance” may actually damage it.” *p. 255*

## The Target Article

“... If one is to trust Bedeian et al.’s (2010) survey of management faculty that examined knowledge of methodological flexibility in their colleagues during the previous year, instances of methodological flexibility are not rare events. For example, 60% of faculty knew of a colleague who “dropped observations or data points from analyses based on a gut feeling that they were inaccurate.” Fifty percent of faculty knew of a colleague who “withheld data that contradicted their previous research.” Other questionable practices are also fairly common (see Bedeian et al., 2010).” *p. 256*

**“... If methodological flexibility does not yield the desired support of a hypothesis, researchers can simply change the hypotheses to match the results (HARKing: hypothesizing after the results are known; Kerr, 1998). The Bedeian et al. (2010) survey reported that approximately 92% of faculties know at least one colleague who has engaged in HARKing in the last year. Thus, HARKing is likely to be a common practice in I–O psychology.” *p. 256***

“... In the most comprehensive review of **publication bias** in psychology, Ferguson and Brannick (2012) reported that publication bias was present in approximately **40%** of meta-analyses and that the degree of this bias was worrisome in about **25%** of meta-analyses. This bias has resulted in misestimating the magnitude of population effects in several I–O research domains. **For example, results were found consistent with the inference that the validities of some commercially available employment tests are overestimated, sometimes substantially** (McDaniel, Rothstein, & Whetzel, 2006).” *p. 257*

## The Target Article

“... Similarly, Kepes, Banks, and Oh (in press) analyzed four datasets from previously published I–O meta-analyses and reported that three of these datasets (work experience and performance, gender differences in transformational leadership, and Pygmalion\* interventions) are likely to have been affected by publication bias.

\* The Pygmalion effect is a type of self-fulfilling prophecy (SFP) in which raising manager expectations regarding subordinate performance boosts subordinate performance. Managers who are led to expect more of their subordinates lead them to greater achievement (from the abstract of Eden, D. (1992). Leadership and expectations: Pygmalion effects and other self-fulfilling prophecies in organizations. *Leadership Quarterly*, 3, 4, 271-305).



### 1 Create Research Registries

*“A research registry is a database in which researchers register studies that they plan to conduct (Banks & McDaniel, 2011; Berlin & Ghera, 2005).” p. 258*

### ② Change the Editorial Review Process

*“we recommend the implementation of a two-step review process (Liberati, 1992). In the **first stage**, authors would submit only part of their manuscript for review, specifically the introduction, method section, and the analysis approach.*

*Manuscripts that pass this stage would advance to the **second stage** in which the complete manuscript, including the results and discussion, would be provided. The editor and reviewers could then assess whether the results and conclusion sections are aligned with the introduction, theory, and method sections (i.e., verify that the authors actually did what they proposed during the initial submission)... ” p. 259*

### ② Change the Editorial Review Process (cont)

*“As part of the second stage of the review process, we also recommend that authors be required to submit the raw data and relevant documentation, including syntax and a summary of all measures. This practice benefits research in three ways.*

***First**, it gives reviewers and the editor the opportunity to check the data and syntax for potential mistakes, a practice that currently is almost never done (Schminke, 2009; Schminke & Ambrose, 2011).*

***Second**, this practice would ensure that the data are securely stored so that, at some future time (e.g., 5 years after the publication date), the journal may publically release the data.*

***Third**, authors would be aware that their data are subject to audit immediately by the editor and reviewers and possibly later by other researchers. ” p. 259*

### ③ Encourage null-effect/null-result publications

*“The suppression of small or null effect sizes is likely the primary cause of research distortion due to publication bias.*

*As mentioned previously, in the social sciences, including I–O psychology, **there is a severe lack of exact replications** (e.g., Makel et al., 2012; Pashler & Harris, 2012; Yong, 2012a). For the entire field of psychology, Makel et al. (2012) estimated that between 1900 and today only around **1%** of all published articles in psychology journals are replication studies. This is unfortunate because **a scientifically “true” effect is one “which can be regularly reproduced by anyone who carries out the appropriate experiment in the way prescribed”** (Popper, 1959, p. 23). ” p. 260-261*

### ④ Strengthen the methods-related belief system

*“One can divide our scientific knowledge into theory-relevant and method-relevant beliefs (LeBel & Peters, 2011). Although the distinction is not sharp, **theory-relevant beliefs** concern the how and why (Sutton & Staw, 1995); that is, the theoretical mechanisms that cause behaviors and other outcomes. By contrast, **method-relevant beliefs** concern the procedures and processes by which data are measured, collected, and analyzed (LeBel & Peters, 2011)...*

*Because psychology tends to be driven by a theory-relevant belief system, much more than a method-relevant belief system, researchers tend to interpret confirmatory results as theory relevant and disconfirmatory results as method relevant, “with the result that the researcher’s hypothesis is artificially buffered from falsification” (LeBel & Peters, 2011, p. 372).” p. 261*

## The Target Article: **Commentary.1**

Briner, R.B. & Walshe, N.D. (2013). **The causes and consequences of a scientific literature we cannot trust: An evidence-based practice perspective.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 269-272.

“We first discuss the fact that researchers are themselves practitioners. Next, we consider some of the reasons I–O and HR practitioners give for adopting dubious practices and illustrate how these closely parallel researchers own practice decisions. Finally, we address the question of how practitioners can be evidence based if the scientific evidence itself is untrustworthy.” *p. 269*

Briner, R.B. & Walshe, N.D. (2013). (cont) ...

**“Why do practitioners adopt practices that are known to be dubious?**

- ①** It would be career-limiting if I didn't pretty much use the same practices as everyone else.
- ②** The best organizations have the best practices so it makes sense to look at what they are doing and to do it too.
- ③** I'm not rewarded for introducing practises that work because they are not really evaluated – I'm rewarded for getting stuff done – and fast.” *p. 269*

Briner, R.B. & Walshe, N.D. (2013). (cont) ...

**They conclude:**

**“If the scientific evidence base is untrustworthy, how can practitioners rely on it or use it to help inform their practice?**

The short answer is that practitioners should not trust any source of evidence whether from their experience, the local organizational context, the views of people who may be affected by the decision, or the scientific literature. Rather, as expressed in the description of evidence-based management above, evidence needs to be critically evaluated for its reliability, validity, and relevance to the problem at hand. In addition, it needs to be combined with the other sources of evidence in order to reach a decision.” *p. 272*



## The Target Article: **Commentary.2**

Sliter, M., Yuan, Z., & Boyd, E. (2013). **Let's be honest: Evidence for why Industrial-Organizational Psychology is trustworthy.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 273-276.

“Although we acknowledge that the issues brought up by the authors might affect the representativeness of published I–O research, we do not believe they have presented any real evidence for a lack of trustworthiness in the extant published work. In fact, the authors never actually define what they mean by ‘trustworthiness.’” *p. 273*

## The Target Article: **Commentary.3**

Bennett, A.A. & Miao, C. (2013). **How do we know the truth? Extensions and examples from similar academic fields.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 276-278.

No substantive points – all tact and no substance.

## The Target Article: **Commentary.4**

Mazzola, J.J. & Deuling, J.K. (2013). **Forgetting what we learned as graduate students: HARKing and selective outcome reporting in I-O journal articles.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 279-284.

“in this commentary, we sought to test, through a pilot study, how much selective outcome reporting (i.e., dropping non-significant hypotheses) and HARKing (hypothesizing after the results are known; Kerr, 1998) might be present in the I–O field. To this aim, we conducted a brief pilot study comparing the percentage of supported hypotheses between published journals, where these practices are likely to take place, and dissertations, which are hypothesized and presented before the study and typically cannot be HARKed post hoc.” **p. 273**

Mazzola, J.J. & Deuling, J.K. (2013). cont ...

“Given that our concern is about testing hypotheses without selective reporting or HARKing, we believe it is only appropriate that we present a few a priori hypotheses of our own:

**Hypothesis 1:** Published journal articles will have a significantly higher percentage of reported hypotheses that are supported than dissertations.

**Hypothesis 2:** Published journal articles will have a significantly lower percentage of reported hypotheses that are unsupported than dissertations.” *p. 279*

## The Target Article: **Commentary.4**

Mazzola, J.J. & Deuling, J.K. (2013). cont ...

Table1: Percentages of Supported, Partially Supported, and Unsupported Hypotheses in a sample of journal articles and dissertations from 2010-2012. (*p. 280*)

	Total	Supported	Partial	Unsupported
Journal Articles	1,684	1,231 (73.1%)	256 (15.2%)	197 (11.7%)
Dissertations	1,227	404 (32.9%)	243 (19.8%)	580 (42.3%)

“These results seem to suggest that as researchers move into the academic field and are forced to “publish or perish,” they (ourselves included) seem to leave behind the careful research practices of hypothesis testing they learned as graduate students conducting dissertations. **Based on the overwhelming support (and medium-large effect sizes) for our hypotheses, it would seem that selective reporting and HARKing are quite common in the journals that we examined and, we suspect, in many of the other journals in the field.**” *p. 281*

Mazzola, J.J. & Deuling, J.K. (2013). cont ...

### **They conclude:**

“When did we become storytellers instead of truth seekers? Where is our sense of adventure? When did the model of a community of problem-solvers seeking truth and knowledge evolve into a community of screenwriters chasing the most plausible good story? In order for I–O psychology to move forward as a science, it is important that we allow the objective data do the majority of the storytelling, not the subjective researcher.” *p. 281*

## The Target Article: **Commentary.5**

Banks, G.C., & O'Boyle, E.H. (2013). **Why we need Industrial-Organizational Psychology to fix Industrial-Organizational Psychology.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 284-287.

No substantive points – all tact and no substance.

## The Target Article: **Commentary.6**

Gabriel, A.S., & Wessel, J.L. (2013). **A step too far? Why publishing raw datasets may hinder data collection.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 287-290.

“we believe that some of the suggestions made by Kepes and McDaniel may actually hinder our research efforts instead of help... Specifically, we believe that requiring raw data submission and publication will be problematic for three key reasons:

- (a) organizations may be less likely to “buy in” to projects,
- (b) sensitive populations may be less likely to consent to participating, and
- (c) responses may be distorted on sensitive variables.

We then argue for increased accountability among co-authors to help minimize unethical publishing practices.” *p. 273*



## The Target Article: **Commentary.7**

Leavitt, K.(2013). **Publication bias might make us untrustworthy, but the solutions may be worse.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 290-295.

“I propose that although the biases proposed by the authors are potentially problematic, the solutions proposed by Kepes and McDaniel would do more to inhibit meaningful scientific progress than they would do to curb the suppression bias...” **p. 291**

## The Target Article: **Commentary.7**

Leavitt, K.(2013)... cont.

“Bedeian, Taylor, and Miller’s (2010) survey of methodological flexibility found that withholding contradicting studies and tweaking research designs during data collection may be common practices in industrial–organizational (I–O) psychology. I suggest that both of these practices may actually reflect scholars’ development of tacit knowledge and drawing informal inductive inferences during the research process. ” *p. 293*

## The Target Article: **Commentary.8**

Derksen, M., & Rietzschel, E.F. (2013). **Surveillance is not the answer, and replication is not a test: Comment on Kepes and McDaniel, “How Trustworthy Is the Scientific Literature in I–O Psychology?”** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 295-298.

“Although we share Kepes and McDaniel’s (2013) concern about the state of affairs in I–O psychology, we think their emphasis on control and correction will, in the end, be counterproductive. Specifically, we argue that questionable research practices can best be remedied by encouraging an open academic culture, characterized by error management, rather than a culture of distrust, aimed at error prevention. Further, we address the call for more replication studies, ... as normal research” *p. 273*

## The Target Article: **Commentary.9**

Landis, R.S., & Rogelberg, S.G. (2013). Our scholarly practices are derailing our progress: The importance of "nothing" in the Organizational Sciences. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 299-302.

“The question becomes how can our field encourage the publication of high-quality research that produces null findings?

① Resocialization: Our field must come to value the importance of null research.

② Examine and potentially modify training models.

Methodological rigor is critical in all research and perhaps even more so in research examining the null ... We need to emphasize in our formal and informal training clear methods for reaching confidence in observed null findings. " *p. 301*

Landis, R.S., & Rogelberg, S.G. (2013)... cont.

### **③ Expand beyond conventional statistical tools**

Other approaches such as reporting point estimates for effect sizes along with confidence intervals and the use of Bayesian methods (Kruschke, Aguinis, & Joo, 2012) should be more fully embraced.

### **④ Editors and Reviewers keep an open mind**

through the review process and not view statistical significance as isomorphic with research impact.

### **⑤ Consider changes to our journal evaluation system**

.. editors are understandably focused on maintaining or improving citation counts associated with published articles as such improvements are typically associated with stronger impact factors. Unfortunately, this focus may lead to practices that have deleterious effects on our literature. *p. 301*

## The Target Article: **Commentary.10**

Stetz, T.A., & Subramony, M. (2013). **Research Registries and trustworthiness of Industrial– Organizational Psychological research** . *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 302-305.

No substantive points – just a few stats about the medical science data registries.

Schmidt, G.B., & Landers, R.N. (2013). **Solving the replication problem in psychology requires much more than a website.** *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 305-309.

“Overall, we agree with Kepes and McDaniel that the lack of replications in I–O psychology is a critical problem, but we believe the allocation of a few pages at the back of journals is not sufficient by itself to increase such replications. We propose instead that Kepes and McDaniel’s vision can be better achieved if journals establish formal processes to publish replications, with organization of such effort conducted online, explicitly recognizing the value of these replications to scientific progress.” *p. 308*

## The Target Article: **Commentary.12**

Jones, J.W., & Dages, K.D. (2013). *A new era of large-scale data sharing: A test publisher's perspective. Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6, 3, 309-312.

In the age of “Big Data” and use of business analytics to gain competitive advantage ... the value of data is at a premium. Vangent recently shared two large datasets with a university researcher that had a total estimated value of over \$500,000. Contracts were needed to ensure that these data were protected, would not be shared with business competitors, and would be used only for the approved purpose. Had this contract not been signed, the data could not have been shared. These contractual requirements also raise concerns regarding Kepes and McDaniel's recommendation of submitting raw data as part of the journal review process.” *p. 310*



### So, can we trust the scientific literature in Industrial and Organizational Psychology?

Who really knows?

Clearly the case made by the authors is compelling of a certain kind of selectivity in the reporting of results. But, to broadly categorize all of ‘the scientific literature in I-O psychology’ as untrustworthy is, I think, a step too far.

But in the accompanying document entitled:

#### **Untrustworthy Reporting of Results in I-O/Psychology**

I’ve presented a few examples which, for me, indicate where *trustworthiness* is needlessly compromised in various ways.

I think for the authors to pose their question and provide a negative answer is nevertheless indicative of an *ethical dilemma* affecting many students and I-O researchers.

What's also of concern to me are the recommendations to impose integrity on researchers by instituting systems and processes for external checks on their data, policing researchers *en-masse*, and visibly maintaining measures of surveillance on them.

Psychology is now apparently a profession in which, because of the actions of a *noticeable number* of its members, policing by external agencies, processes, and regulation of their activities is considered a necessity.

E.g. This is what Elsevier now have in place for article submissions ..

“We recently announced that integration of the plagiarism detection software **CrossCheck** within the Elsevier Editorial System (EES) would go live in October. I am pleased to confirm that for *Personality and Individual Differences* this new service will be available on the EES site from mid-October.

**CrossCheck** uses **iThenticate originality detection software** to identify text similarities which may indicate plagiarism by comparing manuscripts with both a web repository and the CrossCheck database. The integration of CrossCheck in EES benefits editors in several ways: ...”

**We may not trust our students,  
but it seems others no longer trust us.**

① I think this problem is far deeper than inferring it is endemic of a profession gone bad ...

Pulfrey, C., & Butera, F. (2013). [Why neoliberal values of self-enhancement lead to cheating in Higher Education: A motivational account](#). *Psychological Science*, EarlyView, , 1-11.

As they conclude in their abstract: “**These results point to the relevance of diagnosing societal values as social causes of cheating.**”

When I completed my degrees back in the early 80s, the notion that students or academics would engage in **Questionable Research Practices** was simply a ‘rather remote’ proposition. Yes, it happened – but it was so very rare. The thought of ever engaging in this kind of activity for personal gain was almost inconceivable. Now we seem forced to accept it as a common feature of modern academic life.

② And what are the likely consequences of cohorts of graduate students beginning their academic careers, holding a personal view that Questionable Research Practices {QRP} are now just part of the reality of gaining a foothold on the academic employment ladder?

**QRP:** *such as:*

- *personally withholding/not reporting/downplaying contra-hypothesis results for personal gain,*
- *HARKing,*
- *being 'economical with the truth',*
- *obeying their supervisor's/boss's demands to ignore certain data/results,*
- *avoiding critiquing their own results due to implicit threats of supervisor/ boss displeasure and harm to future job prospects*

③ With the corporatization of University Administration has come the imposition of 'performance management', loss of tenure, and government institutional funding formulae based upon 'outputs'.

The notion of the academic as a scholarly, independent-minded intellectual pursuing (*at their own preferred rate*) whatever they find interesting has been replaced by a person who must be politically and organizationally aware, organizationally compliant, under constant pressure to publish in the 'right' journals, and produce 'measurable' outputs – or else.

That transformation, along with the mind-set among many students that they must be awarded a good degree because they have paid for it, has produced sufficient numbers of students and academics HARKing and plagiarizing to the extent that external systems are now required to 'police' their integrity.

④ Do I have any answers, simple or complex? No.

Given university administration and government imposition of performance management and corporate-compliance mandates on their staff, with students under pressure themselves to gain the qualification for which they have paid a considerable sum of money, **the price of failure now far exceeds the risk of being ‘found out’ after engaging in QRPs or a bit of plagiarism.**

And even if found out, does the severity of sanction reflect our own creeping ambiguity over what we *should* feel is trivial or important?

In short, for psychologists *en masse*, has scientific integrity itself become ‘negotiable’ ?

⑤ cont... it's probably why we see articles appearing such as:

Fanelli, D., & Ioannidis, J.P.A. (2013). **US studies may overestimate effect sizes in softer research**. *Proceedings of the National Academy of Sciences*, 110, 37, 15031-15036, ... and ...

**Looks good on paper: A flawed system for judging research is leading to academic fraud**

<http://www.economist.com/news/china/21586845-flawed-system-judging-research-leading-academic-fraud-looks-good-paper?frsc=dg|a>

Chinese academics can now buy research articles and 'slots' in journals. "As China tries to take its seat at the top table of global academia, the criminal underworld has seized on a feature in its research system: the fact that research grants and promotions are awarded on the basis of the number of articles published, not on the quality of the original research. This has fostered an industry of plagiarism, invented research and fake journals that Wuhan University estimated in 2009 was worth \$150m, a fivefold increase on just two years earlier".



[www.cognadev.com](http://www.cognadev.com)



Paul Barrett, PhD.  
Chief Research Scientist

18B Balmoral Ave  
Hurlingham Sandton 2196  
Johannesburg  
South Africa

Tel: +27 11 884-0878  
Fax: +27 11 884-0910

New Zealand

Tel: +64 9 889-2630  
Fax: +64 9 280-6121  
Skype: [pbar088](#)

[paul@cognadev.com](mailto:paul@cognadev.com)