

# **The Validity of Conscientiousness is Moderately Overestimated**

Sven Kepes & Michael A. McDaniel  
Virginia Commonwealth University

# Goals

- This talk will address data supporting the conclusion that conscientiousness is less predictive of job performance than we are lead to believe based on the cumulative research.
- I frame these findings in the context of challenges to evidence-based practice in personnel selection.

# Evidence to Guide Practice

- IPAC is largely composed of practitioners who seek to apply the best evidence when implementing selection systems.
- Journals are typically viewed as sources of the best evidence for personnel selection (and all other disciplines).
- Personnel selection research is largely published in industrial and organizational psychology (I-O) and management journals.

# Growing Concerns

- There are growing concerns about the trustworthiness of the literature reported in I-O and management journals.
- The primary concern stems from the observation that hypotheses are almost always supported.

# Growing Concerns

- We argue that researchers are either:
  - 1) approaching omniscience, or
  - 2) there are forces at work that cause our journal articles to be unrepresentative of all completed research.

# What Journals Publish

- Statistically significant results are more likely to get published in our journals than studies with non-significant results.
- About 97% of all hypotheses presented are supported.
  - Researchers find what they want to find.
  - Journals publish significant findings.

# Journal Competition

- Any given journal is in a competitive market with other journals.
- Journals seek to publish “hot findings” because:
  - Articles with such findings get cited more often than other articles.
  - These citations enhance the reputation of a journal (through metrics such as the impact factor).

# Researcher and Academic Program Competition

- Researchers and the university academic programs in which they work are in a competitive market.
- The highest ranked programs are motivated to stay highly ranked. The lesser ranked programs are motivated to improve their ranking.
- Well-cited researchers wish to continue to be well-cited. Lesser cited researchers seek to become well-cited.



# Government & Industry Reputations

- Researchers in industry and government often wish to document the efficacy of an organization's I-O practices (e.g., selection systems).
- These researchers strive for publications with supported hypotheses in part to serve the commercial or reputational interests of their organizations.

# Questionable Research Practices

- Researchers have substantial methodological flexibility that can be marshaled to obtain supported hypotheses.
  - They can stop data collection, tweak the design or measurement, and discard the original data and collect new data.
  - They can drop outliers or other observations that diminish the magnitude of the research findings (e.g., validity).

# Questionable Research Practices

- Researchers can delete variables that “didn’t work out” and retain variables that “worked.”
- They can collect additional data needed to increase sample size to move a marginally statistically significant effect size (e.g.,  $p < .06$ ) to a significance level that is more acceptable to journals (e.g.,  $p < .05$ ).

# Questionable Research Practices

- Researchers can also terminate data collection once the preferred statistical significance value ( $p$ ) value is obtained.
- They can abandon the hypotheses that were not supported, reverse the direction of the hypotheses, or create new hypotheses after looking at the results.

# Questionable Research Practices

- Researchers can then report the results that best support the retained hypotheses in a nice, neat publishable package and never mention the discarded variables, analyses, observations, and hypotheses.

# How Much Distortion is in the Literature?

- O'Boyle, Banks & Gonzales-Mule (2013, AOM)
  - Best Paper Award - Research Methods Division
- For the management literature, they tracked changes in hypotheses, data, and results as a manuscript moved from defended dissertation to journal publication.

# How Much Distortion in The Literature?

- Dissertation to Journal:
  - 10% added subjects.
  - 20% dropped subjects.
  - 33% showed changes in the means, standard deviations, or interrelations of the included variables.
  - 21% of the unsupported dissertation hypotheses turned into statistically significant journal hypotheses.
  - Added hypotheses were nearly twice as likely to be significant as dropped hypotheses (70.0% vs 38.7%).

# Current Study



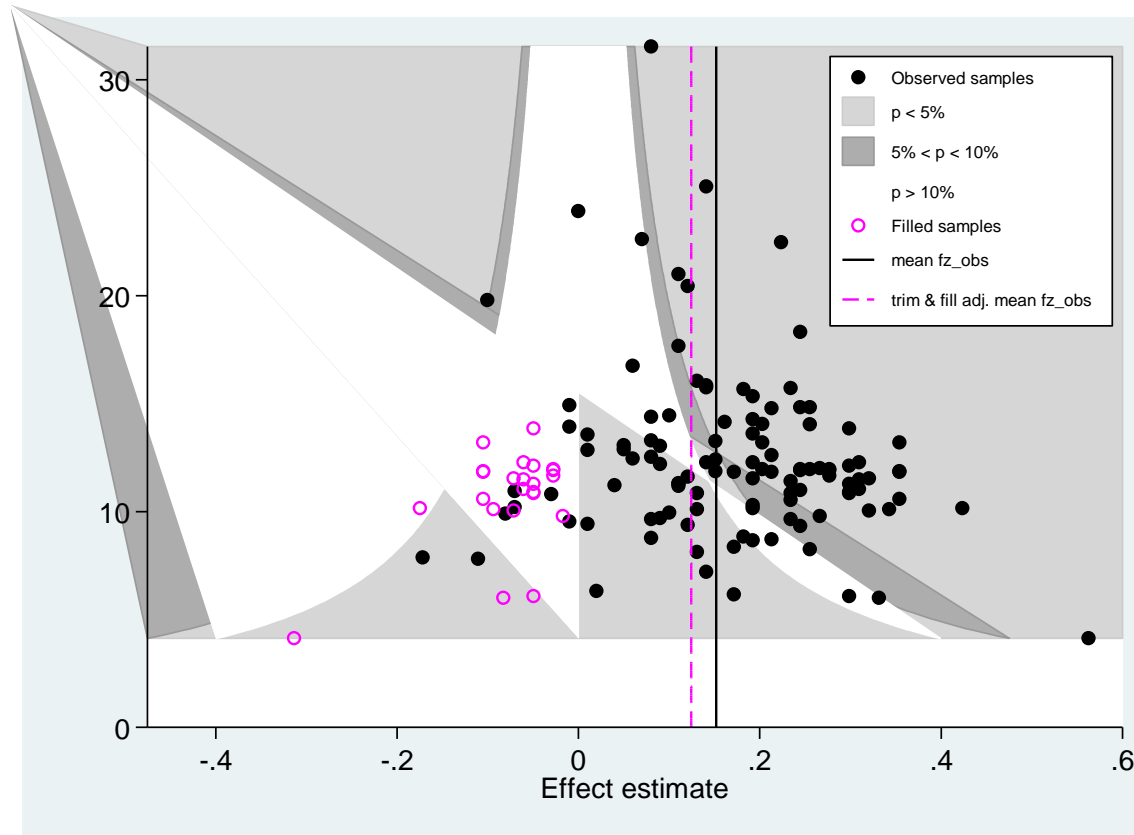
# Method

- The conscientiousness data analyzed are from Shaffer and Postlethwaite (2012) meta-analysis and include 113 validity coefficients for conscientiousness predicting job performance.
- We conducted publication bias analyses.
- Publication bias exists to the extent that data available to a reviewer on a topic is not representative of all data on the topic of interest.

# Highlights

- I can send you the paper with all the details and another paper that explains publication bias analyses.
- Here I will focus on some highlights with pictures.

# All Sources of Data



Mean observed validity = .16

Trim and fill validity = .13

Selection models validity = between .12 and .14

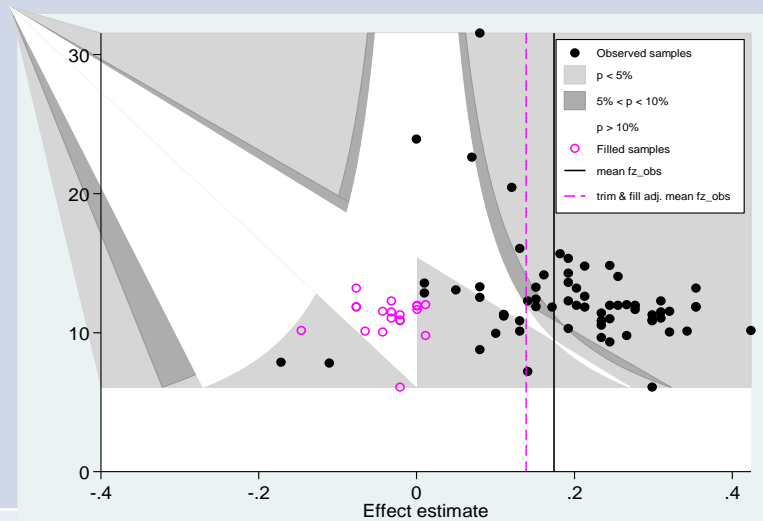
Validity appears to be overestimated between 13% and 19%

# Journal Data has Higher Validity Than Non-Journal Data

- A common, but insufficient, publication bias analysis is a comparison of data source (published vs. non-published).
  - Mean validity in journal data was .19
  - Mean validity in non-journal data was .12
- It is useful to look at potential bias within study source.

# Published vs. Non-published

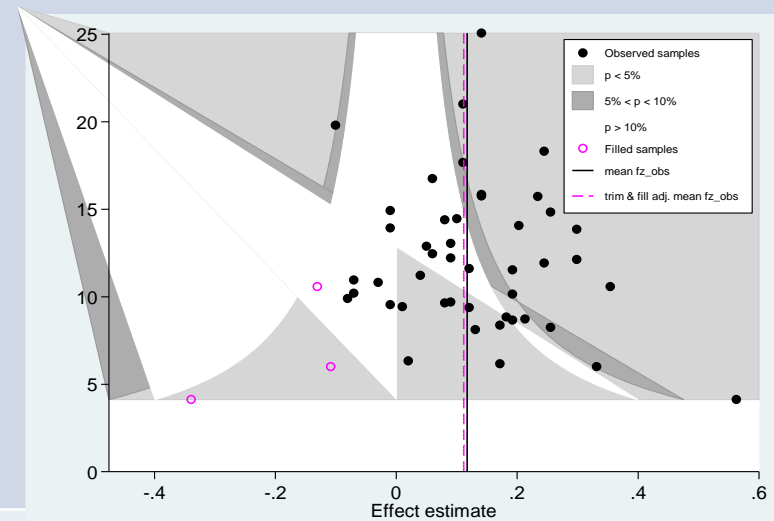
Published



Trim & Fill .19 vs .14 (26% drop)

Selection models .19 vs .16 or .17  
(11% to 16% drop)

Non-Published



Trim & Fill .12 vs .11 (8% drop)

Selection models .12 vs .10  
(17% drop)

Publication bias appears larger in the published data than in the unpublished data.

# Contextualization

- Shaffer and Postlethwaite (2012) compared personality measures that were contextualized vs. non-contextualized.
- Contextualized items reference work.
  - At work, I keep my commitments.
- Non-contextualized items don't reference a context.
  - I keep my commitments.

# Contextualization

## Validity and Publication Bias

- Contextualized conscientiousness measures have higher validity than non-contextualized measures (.19 vs .15).
- Little evidence of publication bias was found for contextualized measures.
- Most of the publication bias in the journal articles with for non-contextualized measures.

# Take-Aways

- Our research literature likely has some publication bias due to questionable research practices.
- Journals want “hot stuff” and authors tend to accommodate the journals by engaging in questionable research practices.



# Take-Aways

- Results are consistent with these inferences:
  - The literature on the validity of conscientiousness is affected by publication bias.
  - Journal articles tend to overestimate the validity of conscientiousness in the prediction of job performance.
  - Validity estimates for non-contextualized measures show the most publication bias, particularly in journals.
  - Validity estimates for contextualized are relatively free of publication bias.