A critique of IRAP research

Ian Hussey, Jamie Cummins & Chad Drake

Data & code:

osf.io/ke7zx

IRAP’s Predictive Validity

Updating a recent meta analysis

IRAP's Predictive Validity

Meta-analysis of association between IRAP & clinically-relevant criterion effects

- Vahey, Nicholson & Barnes-Holmes' (2015)

Widely-cited for sample size justifications

- 66 citations
  - 39 new IRAP papers in this period

"the Ns involved in the studies ... are often relatively small.

Indeed, it could be argued that this impacts upon on the credibility of IRAP research.

However, in a recent meta-analysis of IRAP studies, it was reported that even small *N* IRAP studies have sufficient statistical power" (McEnteggart, 2015)

"the Ns involved in the studies ... are often relatively small.

Indeed, it could be argued that **this impacts upon on the credibility of IRAP research**.

However, in a recent meta-analysis of IRAP studies, it was reported that even small *N* IRAP studies have sufficient statistical power" (McEnteggart, 2015)

Excluded problematic analyses

50% of effect sizes (7 of 15) were excluded

Mere presence of IRAP effects

- Widely misinterpreted due to Generic Pattern
- Not an external criterion (Flake et al., 2017)

IRAP as dependent variable

- No clinical assessment utility (Fried & Kievit, 2016)
- Incompatible with their meta analysis modelling approach
  - Correct multivariate meta: $(Y_1, Y_2) \sim IRAP$
  - Vahey et al. method: $IRAP \sim (X_1, X_2)$

IRAP's Predictive Validity

Excluded problematic analyses

Meta analysis via Hunter & Schmidt method

		95% CI	
	r	Lower	Upper
Original	.45	.40	.54
Updated	.39	.27	.51

Sample size recommendations

80% power for a bivariate correlation

	Required *N*
Original	37
Updated	105

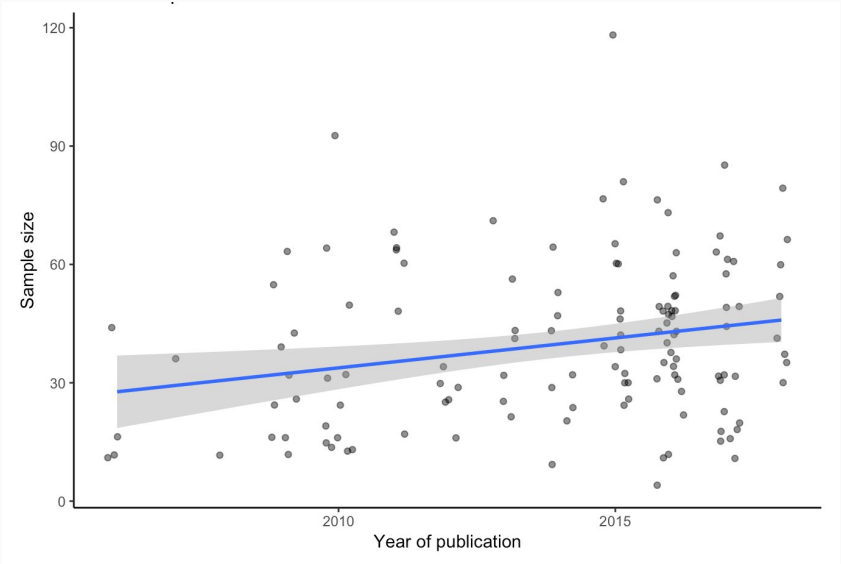
Most IRAP research is under-powered

	% of under-powered published studies
Original	50%
Updated	93%

Most IRAP research is under-powered

If current rate increase in sample sizes continues,

The average study won't be well-powered until the year **2051**



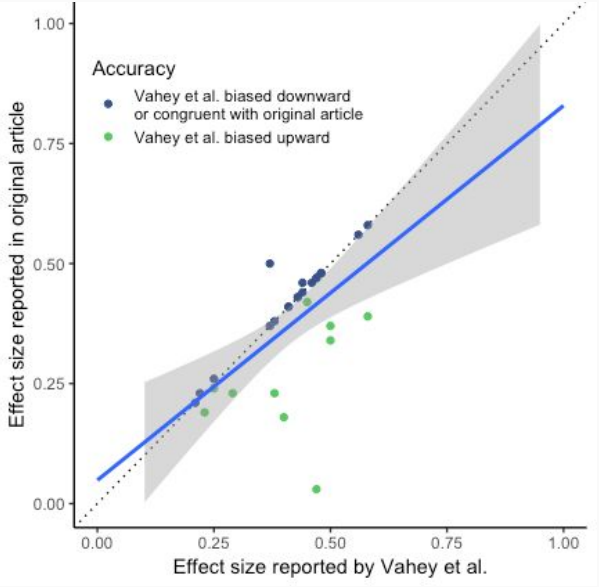
A new meta analysis

Following best practices

Issues with original meta analysis

Effect size extraction errors

- Incongruities in 33% of cases
- Biased upwards



Issues with original meta analysis

Hypothesizing After Results are Known (HARKing)

- Inclusions based on what the meta authors thought **could have been predicted** ahead of time, not what the original authors **actually predicted**

No blinding

- Researchers knew the effect size when choosing them

Issues with original meta analysis

Only relevant to deductive research

- Meta analysis of *predictable effects* can only inform future research that is making *predictions*
- But current this *deductive* meta is now inappropriately cited in *inductive* research to justify sample sizes

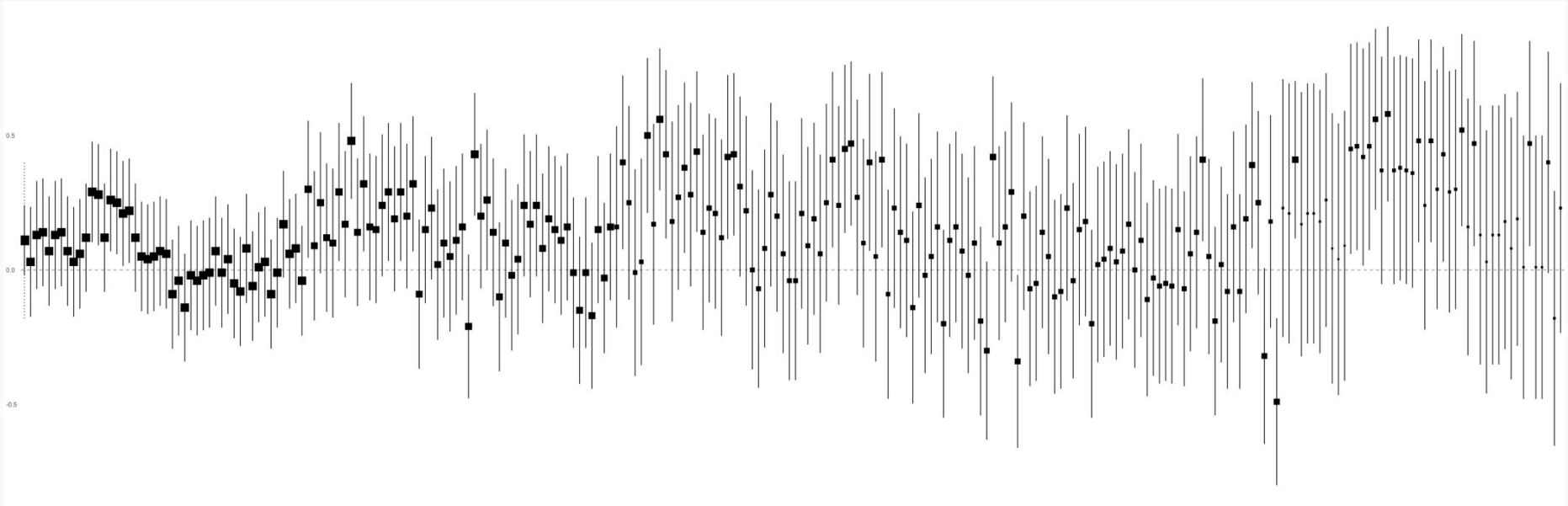
IRAP's Predictive Validity

A meta-analysis for inductive research

Modern meta-analytic best practices

- Multilevel meta analysis
- Restricted Maximum Likelihood estimation & *N* weighting
- Considered same articles as original meta
- Included all 249 effect sizes
    - Other than previously specified problematic analysis types

Sample size recommendations



Meta-effect size: r = .11, 95% CI [-.02, .24], p = .10.

Sample size recommendations

80% power for a bivariate correlation:

	Required *N*
Original	37
Updated	105
New	19,620

Conclusion

There is a problem with ~~the IRAP~~ our research practices

The way forward

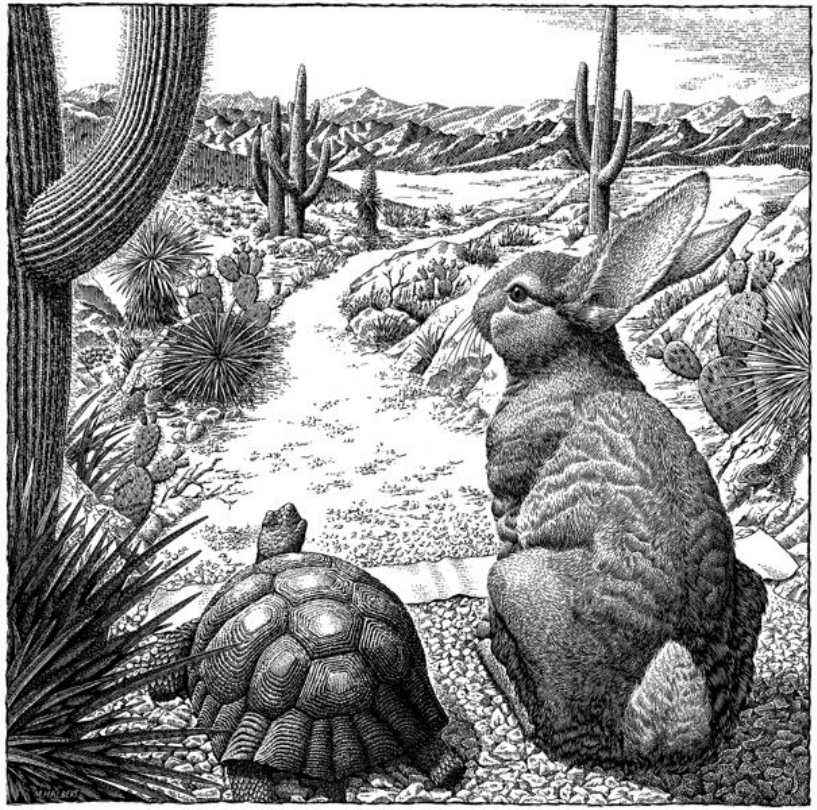
Our research practices are not exceptional

Real issue: we're resistant to change

- Other fields have had their crisis, are now 8 years into recovery

There's still time to fix this!

- More power, better use of statistics, pre-registration, direct replication
  - See Munafò et al. (2017) *A manifesto for reproducible science*



Tortoise vs. Hare approaches to science

Data & code:

osf.io/ke7zx

Ian Hussey

Postdoctoral research fellow
Ghent University
Belgium

ian.hussey@ugent.be
@ianhussey