Evaluating Evidence for Low Data Veracity
From Reported Summary Measures
A Case Study

Carel F.W. Peeters
Dept. of Epidemiology & Biostatistics
VU University medical center, Amsterdam
cf.peeters@vumc.nl

Chris A.J. Klaassen
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
&

Mark A. van de Wiel
Dept. of Epidemiology & Biostatistics
VU University medical center, Amsterdam

Amsterdam, the Netherlands
May 31, 2017
C) Experiment 8: Haptic Induction

- Global: 4.79
- Local: 2.87
- Control: 3.83
C) Experiment 8: Haptic Induction

- Local: 2.87
- Control: 3.83
- Global: 4.79
Sense Creative!
Jens Förster

Appointments

- Universiteit van Amsterdam
- Kurt Lewin Instituut
- Ruhr Universität Bochum

Research

- Experimental social psychology
- (Implicit) processing and social priming

Superstar in Social Psychology

- Highly cited
- Awarded Alexander von Humboldt Professorship in 2013
Sequence of the Förster Case

07/2013

Universiteit van Amsterdam

2013

Landelijk Orgaan LOWI
Wetenschappelijke Integriteit

03/2014

"(...) the results in Foerster and Denzler (2012) must have been manipulated (...)"

10/2014

Assignment
Shed light on the status of all JF papers published under UvA affiliation in terms of their scientific veracity
Focus
Judgments regarding the veracity of individual manuscripts

Evaluation
Evaluation of a statistical nature only

Design
Methods geared towards one-way ANOVA-type designs with 3 factor levels

Data
Only summary measures (from JF publications) at disposal

Qualitative judgements on the evidence for low scientific veracity

- Strong evidence
- Inconclusive evidence
- No evidence
Methods

Intuition

![Graphs showing population and sample data with A, B, and C categories.](image)
$\Delta F$ testing paired with Fisher’s method
ΔF testing paired with Fisher’s method
ΔF testing paired with Fisher’s method

Nested test

Does the more complex model significantly contribute to model fit?

Null hypothesis

\(H_0: \text{perfect linearity of the cell means}\)

Test statistic

\[\Delta F = \frac{n(\bar{x}_1 - 2\bar{x}_2 + \bar{x}_3)^2}{2(s_1^2 + s_2^2 + s_3^2)}\]

Distribution

\[\Delta F \sim F_{1,3(n-1)}\]
\( \Delta F \) testing paired with Fisher’s method

Behavior \( p(\Delta F) \) under \( H_0 \)

The \( p \)-values for the \( \Delta F \) test (\( p(\Delta F) \)) must be uniformly distributed

Fisher method

Quantifies how strongly an accumulation of tests favors the shared null
Intuition

Quantifies excessive closeness of the cell averages to perfect linearity given the reported cell variances
Evidential Value $\mathbb{V}$

Values

$\mathbb{V} \in [1, \infty)$

- $\mathbb{V} \approx 1$: Honest experiment
- $\mathbb{V} \geq 6$: Substantial, indicative of dependence structure
Control Studies: Trend Lines
### Table: Results control studies

<table>
<thead>
<tr>
<th>sample</th>
<th>$\Delta F$</th>
<th>$p(\Delta F)$</th>
<th>$V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hagtvedt1</td>
<td>0.2852</td>
<td>0.5951</td>
<td>1.3955</td>
</tr>
<tr>
<td>Hagtvedt2</td>
<td>0.3483</td>
<td>0.5570</td>
<td>1.1741</td>
</tr>
<tr>
<td>Hunt</td>
<td>1.5152</td>
<td>0.2224</td>
<td>1.0000</td>
</tr>
<tr>
<td>Jia</td>
<td>1.0437</td>
<td>0.3089</td>
<td>1.0000</td>
</tr>
<tr>
<td>Kanten1</td>
<td>0.9318</td>
<td>0.3362</td>
<td>1.0014</td>
</tr>
<tr>
<td>Kanten2</td>
<td>0.1478</td>
<td>0.7013</td>
<td>1.7535</td>
</tr>
<tr>
<td>Lerouge1</td>
<td>1.8439</td>
<td>0.1796</td>
<td>1.0000</td>
</tr>
<tr>
<td>Lerouge2</td>
<td>0.0018</td>
<td>0.9660</td>
<td>12.2265–13.0148</td>
</tr>
<tr>
<td>Lerouge3</td>
<td>0.8550</td>
<td>0.3595</td>
<td>1.0994</td>
</tr>
<tr>
<td>Lerouge4</td>
<td>0.3874</td>
<td>0.5364</td>
<td>1.2055</td>
</tr>
<tr>
<td>Malkoc</td>
<td>0.0143</td>
<td>0.9048</td>
<td>5.2558–5.2663</td>
</tr>
<tr>
<td>Polman</td>
<td>0.2462</td>
<td>0.6215</td>
<td>1.3369</td>
</tr>
<tr>
<td>Rook1</td>
<td>1.3421</td>
<td>0.2501</td>
<td>1.0000</td>
</tr>
<tr>
<td>Rook2</td>
<td>0.1649</td>
<td>0.6857</td>
<td>1.6933</td>
</tr>
<tr>
<td>Smith1</td>
<td>0.7938</td>
<td>0.3760</td>
<td>1.0146</td>
</tr>
<tr>
<td>Smith2</td>
<td>0.3275</td>
<td>0.5689</td>
<td>1.2640</td>
</tr>
<tr>
<td>Smith3</td>
<td>1.0743</td>
<td>0.3023</td>
<td>1.0000</td>
</tr>
<tr>
<td>Smith4</td>
<td>0.0190</td>
<td>0.8905</td>
<td>4.0388</td>
</tr>
<tr>
<td>Smith5</td>
<td>0.1588</td>
<td>0.6909</td>
<td>1.6268</td>
</tr>
<tr>
<td>Smith6</td>
<td>1.0289</td>
<td>0.3130</td>
<td>1.0000</td>
</tr>
<tr>
<td>Smith7</td>
<td>0.8435</td>
<td>0.3600</td>
<td>1.0200</td>
</tr>
</tbody>
</table>

### Table: Results independent samples regarding participant scores JF11.JEPG.

<table>
<thead>
<tr>
<th>Study</th>
<th>$\Delta F$</th>
<th>$p(\Delta F)$</th>
<th>$V$</th>
</tr>
</thead>
<tbody>
<tr>
<td>study1A</td>
<td>0.1846</td>
<td>0.6695</td>
<td>1.6403</td>
</tr>
<tr>
<td>study1B</td>
<td>0.0760</td>
<td>0.7839</td>
<td>2.3664</td>
</tr>
<tr>
<td>study1C</td>
<td>0.2342</td>
<td>0.6303</td>
<td>1.4074</td>
</tr>
<tr>
<td>study2A</td>
<td>0.0172</td>
<td>0.8960</td>
<td>4.8226</td>
</tr>
<tr>
<td>study2B</td>
<td>0.1663</td>
<td>0.6855</td>
<td>1.4756</td>
</tr>
<tr>
<td>study3A</td>
<td>0.0523</td>
<td>0.8203</td>
<td>2.6599</td>
</tr>
<tr>
<td>study3B</td>
<td>0.0000</td>
<td>1.0000</td>
<td>NaN</td>
</tr>
<tr>
<td>study4A</td>
<td>0.0142</td>
<td>0.9056</td>
<td>4.7322</td>
</tr>
<tr>
<td>study4B</td>
<td>0.0000</td>
<td>1.0000</td>
<td>15.2314–NaN</td>
</tr>
<tr>
<td>study5A</td>
<td>0.0083</td>
<td>0.9277</td>
<td>6.1719–7.0386</td>
</tr>
<tr>
<td>study5B</td>
<td>0.0063</td>
<td>0.9370</td>
<td>6.4644–7.2233</td>
</tr>
<tr>
<td>study5C</td>
<td>0.0036</td>
<td>0.9524</td>
<td>9.4973</td>
</tr>
<tr>
<td>study5D</td>
<td>0.0044</td>
<td>0.9475</td>
<td>8.7926</td>
</tr>
</tbody>
</table>
Methods in Action

JF11.JEPG: Visualization Results

Control Studies

JF11.JEPG
File Drawering Alternative Explanation?
Recreating Publication by Simulation

Absolute distance from linearity

Frequency

0 1 2 3 4

0 20000 40000 60000 80000

Carel F.W. Peeters
JF Studies: Overall Picture

24 empirical publications
UvA period

9 publications could *not* be scrutinized

15 publications could be scrutinized

Strong evidence
8
- 3 sole-authorships
- 4 first-authorships
- 1 co-authorship

Inconclusive evidence
3
- 1 first-authorship
- 2 co-authorships

No evidence
4
- 4 co-authorships
At Current

Alexander von Humboldt Professorship

04/2015: JF relinquishes von Humboldt Professorship

Retractions
- 3 retractions
- 3 expressions of concern
- 2 articles under consideration for retraction

Attention
- Quite some media attention (with exception of German media)
- Co-authors respond in public letters
- JF denies any wrongdoings
References


