Good and Bad Replications in Political Science: How Replicators and Original Authors (Should) Talk to Each Other

Nicole Janz
University of Nottingham
nicole.janz@nottingham.ac.uk
@polscireplicate

Jeremy Freese
Stanford University
jfreese@stanford.edu
@jeremyfreese
Psychology’s Replication Crisis Is Running Out of Excuses

Another big project has found that only half of studies can be repeated. And this time, the usual explanations fall flat.

- 200 Psychologists
- 28 studies
The Case of the Amazing Gay-Marriage Data: How a Graduate Student Reluctantly Uncovered a Huge Scientific Fraud
## Replication vs. Duplication

<table>
<thead>
<tr>
<th>Duplication</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verification of research results</strong></td>
<td><strong>Test robustness of research results</strong></td>
</tr>
<tr>
<td>• did errors occur that would reverse findings?</td>
<td>• due to chance?</td>
</tr>
<tr>
<td>• student learning context</td>
<td>• suboptimal methods or data?</td>
</tr>
<tr>
<td><em>same data</em></td>
<td><em>new data</em></td>
</tr>
<tr>
<td><em>same methods</em></td>
<td><em>new methods</em></td>
</tr>
</tbody>
</table>

#ossc19 @polscireplicate @jeremyfreese
## Handling different results

<table>
<thead>
<tr>
<th>Duplication</th>
<th>Replication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasons are <strong>knowable</strong></td>
<td>Reasons are <strong>mysterious</strong></td>
</tr>
<tr>
<td>- honest mistake by author or duplicator</td>
<td>- different or suboptimal methods or data</td>
</tr>
<tr>
<td>- original study opaque</td>
<td>- narrow context</td>
</tr>
<tr>
<td>- by chance</td>
<td>- by chance</td>
</tr>
</tbody>
</table>

*Duplicate your study before journal submission!*

*Describe changes exactly & avoid “failed” label for diverging results*
“Replication Chains”

- Original Study
- Duplication Replication
- Original author comment

#ossc19 @polscireplicate @jeremyfreese
"We ... find that coding errors, selective exclusion of available data, and unconventional weighting of summary statistics lead to serious errors" (Herndon et al. 2013)

“If we cannot even reproduce the original results using the same publicly available data, there is no need for further commentary.” (Miller et al, 2001)
How original authors respond

“less realistic”, “inconsistent with the substantive literature,” and “of limited utility” (Mansfield, Milner, and Rosendorff 2002)

“fundamentally flawed” (Peffley, Knigge, and Hurwitz 2001)

“statistical, computational, and reporting errors that invalidate its conclusions” (Gerber and Green 2005:301).
1
Transparent planning
Clear aim:
Are you conducting a replication or duplication?

Be transparent & reproducible:
• Why have you chosen the original study for replication?
• Is your methods and data section clear?
• Pre-register your replication to avoid accusations of p-hacking or deliberate error hunting
• Crosscheck your duplication/replication before submission

Be an expert:
• Engage with substantive literature to ensure that interpretation of differences is suitable to the field
• Extension motivated by theory or methods critique, not playing with data
Rhetorical sensitivity
Avoid binary judgments:
present diverging results step by step and interpret why this might be the case
avoid saying that the du-/replication of a study has “failed”

Don’t make it personal:
professional, courteous and collaborative (as opposed to confrontational) language
make clear what the positive contribution of the original article is--after all, you would not have chosen it if it the study to be re-examined was not crucial to the field.
honest mistakes are human

Look forward, not backwards:
your judgment on the original study is not ‘final
discuss how the literature can move forward
What **constructive** replicators write

“this is **not a critique** of existing papers, which **faithfully report careful** studies ... Rather, replication with a different event, sample, and time is a **way to move the literature forward** to assess robustness“

“**not** be taken as definitive evidence that the extant literature over-states the extent of irrelevant events; yet, it serves as a (cautionary) **prompt to the next generation** of work.”

(Busby and Druckman 2018)
Replicate others as you would like to be replicated yourself!
Thank you!

Nicole Janz
University of Nottingham
nicole.janz@nottingham.ac.uk
@polscireplicate

Jeremy Freese
Stanford University
jfreese@stanford.edu
@jeremyfreese