Training needs for reproducible research in clinical psychology

Breda Cullen
Glasgow
Training needs for reproducible research in clinical psychology

Breda Cullen

13 November 2018
Estimating the reproducibility of psychological science

Open Science Collaboration

Reproducibility is a defining feature of science, but the extent to which it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when available. Replication effects were half the magnitude of original effects, representing a substantial decline. Ninety-seven percent of original studies had statistically significant results. Thirty-six percent of replications had statistically significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of effects were subjectively rated to have replicated the original result; and if no bias in original results is assumed, combining original and replication results left 68% with statistically significant effects. Correlational tests suggest that replication success was better predicted by the strength of original evidence than by characteristics of the original and replication teams.

What psychology’s crisis means for the future of science

The field is currently undergoing a painful period of introspection. It will emerge stronger than before.

By Brian Resnick | @B_resnick | brian@vox.com | Updated Mar 25, 2016, 9:54am EDT

Hannah Devlin Science correspondent
@hannahdev
Mon 27 Aug 2018 16.00 BST

Attempt to replicate major social scientific findings of past decade fails

Scientists and the design of experiments under scrutiny after a major project fails to reproduce results of high profile studies
It’s Time to Broaden the Replicability Conversation: Thoughts for and From Clinical Psychological Science

Jennifer L. Tackett1, Scott O. Lilienfeld2, Christopher J. Patrick3, Sheri L. Johnson4, Robert F. Krueger5, Joshua D. Miller6, Thomas F. Oltmanns7, and Patrick E. Shorut8

1Northwestern University; 2Emory University; 3Florida State University; 4University of California, Berkeley; 5University of Minnesota; 6University of Georgia; 7Washington University in St. Louis; 8New York University

Abstract

Despite increasing interest in issues of replicability, open science, research transparency, and improving methods and practices in psychological science, the clinical psychology community has been slow to engage. This has been slowly shifting, and the authors of the present article hope to facilitate this emerging dialogue. We begin by examining some potential areas of weakness in clinical psychology in terms of methods, practices, and evidential base. We then discuss a select overview of solutions, tools, and current concerns of the reform movement from a clinical psychological science perspective. We examine areas of clinical science expertise (i.e., implementation science) that should be leveraged to inform open science and reform efforts. Finally, we reiterate the call to clinical psychologists to increase their engagement with these efforts toward reform that should accelerate a more credible clinical psychological science.

• Replication, Reproducibility and Open Science
• Training needs
• Barriers and opportunities
Replication
Repeating research with new data to determine whether results generalise across time and across situations.

Reproducibility
A different researcher can repeat the same analysis using the same data and obtain the same results.

Open Science
The process, content and outcomes of research are openly accessible.

Patil et al. (2016) bioRXiv; doi:10.1101/066803
Center for Open Science: https://cos.io/about/mission/
Replication
Repeating research with new data to determine whether results generalise across time and across situations

Reproducibility
A different researcher can repeat the same analysis using the same data and obtain the same results

Open Science
The process, content and outcomes of research are openly accessible
Training needs

• Attitudes and culture
• Habits
• Integrity
• Transparency
• Insight into own allegiances and biases
• Comfortable with uncertainty
  – Methods
  – Results
• Value feasibility and pilot research
The first principle is that you must not fool yourself and you are the easiest person to fool.

Richard P. Feynman
• Questionable research practices (QRP)
  – *p*-hacking
  – Hypothesising after the results are known (HARKing)
  – Undisclosed flexibility in analyses
• ‘Forking paths’ and ‘researcher degrees of freedom’

Simmons et al. (2012) Psychological Science; doi:10.1177/0956797611417632

https://xkcd.com/1478/

**Writing the Empirical Journal Article**

**Daryl J. Bem**

**Which Article Should You Write?**

There are two possible articles you can write: (1) the article you planned to write when you designed your study or (2) the article that makes the most sense now that you have seen the results. They are rarely the same, and the correct answer is (2).

[...]

To compensate for this remoteness from our participants, let us at least become intimately familiar with the record of their behavior: the data. Examine them from every angle. Analyze the sexes separately. Make up new composite indices. If a datum suggests a new hypothesis, try to find further evidence for it elsewhere in the data. If you see dim traces of interesting patterns, try to reorganize the data to bring them into bolder relief. If there are participants you don’t like, or trials, observers, or interviewers who gave you anomalous results, drop them (temporarily). Go on a fishing expedition for something—anything—interesting.
Habits

- Quality control
- Analysis methods
- Reporting
- Sharing
• **Standardised data capture and cleaning**
  – Case Record Forms (CRF)
  – Database software e.g. MS Access
  – Double data entry and comparison
  – Data validation
  – Detailed dictionary
• **Automated scoring/calculations**
• **Master versus working copy**
• **Data Management Plan (DMP)**

Bellary et al. (2014) Perspectives in Clinical Research; doi:10.4103/2229-3485.140555
DMP guidance and templates: https://dmponline.dcc.ac.uk/
Spreadsheets are widely used software tools for data entry, storage, analysis, and visualization. Focusing on the data entry and storage aspects, this paper offers practical recommendations for organizing spreadsheet data to reduce errors and ease later analyses. The basic principles are: be consistent, write dates like YYYY-MM-DD, don’t leave any cells empty, put just one thing in a cell, organize the data as a single rectangle (with subjects as rows and variables as columns, and with a single header row), create a data dictionary, don’t include calculations in the raw data files, don’t use font color or highlighting as data, choose good names for things, make backups, use data validation to avoid data entry errors, and save the data in plain text files.
Analysis methods

- Primary and secondary outcomes/endpoints
- Pre-registration
- Statistical Analysis Plan (SAP)

Pre-registration example: https://osf.io/sgrk6/
• **Planned sensitivity analyses and subgroup analyses**
  – Do the results change if I make different assumptions?
  – Do the results change if I handle the data differently?
  – What would I have done if the data had come out differently?

Researchers degrees of freedom can lead to a multiple comparisons problem, even in settings where researchers perform only a single analysis on their data. The problem is there can be a large number of potential comparisons when the details of data analysis are highly contingent on data, without the researcher having to perform any conscious procedure of fishing or examining multiple p-values. We discuss in the context of several examples of published papers where data-analysis decisions were theoretically-motivated based on previous literature, but where the details of data selection and analysis were not pre-specified and, as a result, were contingent on data.

• **Transparent and comprehensive**
  – Avoid cherry-picking and reverse-engineering
• **Appendices and online supplements**
Sharing

- Protocol, materials, data, analytic methods, results
  - Transparency and Openness Promotion (TOP) guidelines

<table>
<thead>
<tr>
<th></th>
<th>Not Implemented</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citation Standards</td>
<td>Journal encourages citation of data, code, and materials, or says nothing.</td>
<td>Journal describes citation of data in guidelines to authors with clear rules and examples.</td>
<td>Article provides appropriate citation for data and materials used consistent with journal's author guidelines.</td>
<td>Article is not published until providing appropriate citation for data and materials following journal's author guidelines.</td>
</tr>
<tr>
<td>Data Transparency</td>
<td>Journal encourages data sharing, or says nothing.</td>
<td>Article states whether data are available, and, if so, where to access them.</td>
<td>Data must be posted to a trusted repository. Exceptions must be identified at article submission.</td>
<td>Data must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.</td>
</tr>
<tr>
<td>Analytic Methods (Code) Transparency</td>
<td>Journal encourages code sharing, or says nothing.</td>
<td>Article states whether code is available, and, if so, where to access it.</td>
<td>Code must be posted to a trusted repository, Exceptions must be identified at article submission.</td>
<td>Code must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.</td>
</tr>
<tr>
<td>Research Materials Transparency</td>
<td>Journal encourages materials sharing, or says nothing.</td>
<td>Article states whether materials are available, and, if so, where to access them.</td>
<td>Materials must be posted to a trusted repository. Exceptions must be identified at article submission.</td>
<td>Materials must be posted to a trusted repository, and reported analyses will be reproduced independently prior to publication.</td>
</tr>
<tr>
<td>Design and Analysis Transparency</td>
<td>Journal encourages design and analysis transparency, or says nothing.</td>
<td>Journal articulates design transparency standards.</td>
<td>Journal requires adherence to design transparency standards for review and publication.</td>
<td>Journal requires and enforces adherence to design transparency standards for review and publication.</td>
</tr>
<tr>
<td>Study Preregistration</td>
<td>Journal says nothing.</td>
<td>Article states whether preregistration of study exists, and, if so, where to access it.</td>
<td>Article states whether preregistration of study exists, and, if so, allows journal access during peer review for verification.</td>
<td>Journal requires preregistration of studies and provides link and badge in article to meeting requirements.</td>
</tr>
<tr>
<td>Analysis Plan Preregistration</td>
<td>Journal says nothing.</td>
<td>Article states whether preregistration of analysis plan exists, and, if so, where to access it.</td>
<td>Article states whether preregistration with analysis plan exists, and, if so, allows journal access during peer review for verification.</td>
<td>Journal requires preregistration of studies with analysis plans and provides link and badge in article to meeting requirements.</td>
</tr>
<tr>
<td>Replication</td>
<td>Journal discourages submission of replication studies, or says nothing.</td>
<td>Journal encourages submission of replication studies.</td>
<td>Journal encourages submission of replication studies and conducts results blind review.</td>
<td>Journal uses Registered Reports as a submission option for replication studies with peer review prior to observing the study outcomes.</td>
</tr>
</tbody>
</table>
• **Analytic methods (code)**
  – **Why?**
    * Efficiency and accuracy
    * Flexibility
    * Commenting and logging
  – **How?**
    * Indirectly via point-and-click menus
    * Directly in syntax editor

  – SPSS example…
<table>
<thead>
<tr>
<th></th>
<th>id</th>
<th>depvar</th>
<th>indepvar1</th>
<th>indepvar2</th>
<th>indepvar3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1001</td>
<td>27</td>
<td>1</td>
<td>79</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1002</td>
<td>25</td>
<td>1</td>
<td>63</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1003</td>
<td>17</td>
<td>0</td>
<td>111</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1004</td>
<td>15</td>
<td>1</td>
<td>115</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1005</td>
<td>8</td>
<td>1</td>
<td>96</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>1006</td>
<td>9</td>
<td>0</td>
<td>97</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1007</td>
<td>20</td>
<td>0</td>
<td>106</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1008</td>
<td>10</td>
<td>0</td>
<td>101</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>1009</td>
<td>12</td>
<td>1</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1010</td>
<td>13</td>
<td>0</td>
<td>103</td>
<td>3</td>
</tr>
</tbody>
</table>

**Linear Regression**

- **Dependent:** depvar
- **Independent(s):** indepvar1, indepvar2
- **Method:** Enter
- **Selection Variable:**
- **Case Labels:**
- **WLS Weight:**
DATASET ACTIVATE
BOOTSTRAP
REGRESSION
* Chart Builder.
GGGRAPH
BEGIN GPL
END GPL

BOOTSTRAP
/SAMPLING METHOD=SIMPLE
/VARIABLES TARGET=depvar INPUT= indepvar1 indepvar2
/Criteria CLEVEL=95 CTYPE=PERCENTILE NSAMPLES=1000
/MISSING USERMISSING=EXCLUDE.

REGRESSION
/DESCRIPTIVES MEAN STDERR CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL
/Criteria=PINF(.05) POUT(.10)
/NOORIGIN
/DEPENDENT depvar
/METHOD=ENTER indepvar1 indepvar2

SAVE ZRESID.

GGGRAPH
/GGRAPHDATASET NAME="graphdataset" VARIABLES=indepvar1 depvar MISSING=LISTWISE REPORTMISSING=NO
/GGRAPHSPEC SOURCE=INLINE.
BEGIN GPL
SOURCE: s=userSource(id("graphdataset"))
DATA: indepvar1=col(source(s), name("indepvar1"), unit.category())
DATA: depvar=col(source(s), name("depvar"))
DATA: id=col(source(s), name("SCASENUM"), unit.category())
GUIDE: axis(dim(1), label("indepvar1"))
GUIDE: axis(dim(2), label("depvar"))
SCALE: linear(dim(1), include(0))
ELEMENT: schema/position(bin.quantile.letter(indepvar1*depvar), label(id))
END GPL.
Repeat the regression but add indepvar3 as well.
- R users write and edit code directly with minimal use of menus*
- RStudio is a popular user-friendly interface for working in R

* There are packages that provide a more comprehensive point-and-click interface for R, e.g. Jamovi (which mimics SPSS): https://www.r-bloggers.com/jamovi-for-r-easy-but-controversial/

Sharing

https://osf.io/
Image above is an example of a publicly viewable OSF project
https://psyarxiv.com/
Barriers

- Old habits die hard
- Governance and sensitive clinical data
- Risk of being ‘scooped’ or losing publication opportunities
- Fear of exposure (e.g. errors)
- Resources (set-up and maintenance)
- (Mis)alignment with reward structures in our field
• New generations with different attitudes and skills
• Scope to work with participants on governance issues
• Re-think nature of collaboration and credit
• ‘Push’ factor from funders and journals → justify more resources
Questions

• Other barriers and pitfalls?
• Examples of good practice?
Group of Trainers in Clinical Psychology 2018 Annual Conference

Please tweet about the Conference

#GTiCP2018