Relevance of Population Neuroscience for Understanding Human Development Conference
University of Michigan
May 16-May 17, 2014

Concluding Remarks

Victoria Stodden and Moriah Thomason
Increased access to multidimensional data
Black box
So how can we do better?

- Improved modeling
- Replication (nearly unheard of in NI)
- Move past the p value in neuroimaging – communicate effect size, perform ROC analyses
- More clinically significant meaningful variable may not be unearthed by comparing extremes
• Validation and resampling techniques (e.g., optimize on one set, train on another)
Convergent validity
Data sharing
Data sharing

- If you did not collect the data you do not know it as well
- Multiplying potential for creating false information
- Should you quarantine data
- Processed or raw data or both
- Data papers: Academic crediting, h-index
- Sticky problem: Assumption that measures or analyses selected are the right ones
Big data, neuroimaging

- Demography relays that there may be new methods we have not utilized well in neuroimaging
Big data, neuroimaging

- Paring down variables; feature selection
Representative

- Western, Educated, Industrialized, Rich, Democratic
- Psych-95% from 12% (Arnett 2008)
- Neuroimaging-90% from Western (Chiao 2009)
Representative

- CHEW: Convenience, Hospital/University Educated White
- What are the policy implications?
Causality

- Independent explanatory variables
- Genetic variability; Mendelian randomization
- Run experiments (e.g., longitudinal, controlled)
Prediction

- Call for more predictive models of population processes from neuroscientists (Colter)
- Bobby expresses that as an economist it is unsettling that WRT brain volume her collaborators could not give her a direction for prediction
- Hugh highlighted problem with retrospective prediction in longitudinal studies
  - Baseline imaging is DV instead of IV
  - Inflated prediction
  - Overfit model
  - Multidimensional problems
Need for standardization

- We do not have consensus on what the best metrics are (e.g., fMRI, behavior, phenotypic, psychiatric)
- Furthermore – one size may not fit all (e.g., Rich aggression in sibs)

Victoria Stodden, Ph.D.
Assistant Professor in the Department of Statistics at Columbia University
Code and Data Sharing

Journal Policy setting study design:

Select all journals from ISI classifications “Statistics & Probability,” “Mathematical & Computational Biology,” and “Multidisciplinary Sciences” (this includes Science and Nature).

N = 170, after deleting journals that have ceased publication.

Create dataset with ISI information (impact factor, citations, publisher) and supplement with publication policies as listed on journal websites, in June 2011 and June 2012.
### Journal Data Sharing Policy

<table>
<thead>
<tr>
<th>Requirement</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required as condition of publication, barring exceptions</td>
<td>10.6%</td>
<td>11.2%</td>
</tr>
<tr>
<td>Required but may not affect editorial decisions</td>
<td>1.7%</td>
<td>5.9%</td>
</tr>
<tr>
<td>Encouraged/addressed, may be reviewed and/or hosted</td>
<td>20.6%</td>
<td>17.6%</td>
</tr>
<tr>
<td>Implied</td>
<td>0%</td>
<td>2.9%</td>
</tr>
<tr>
<td>No mention</td>
<td>67.1%</td>
<td>62.4%</td>
</tr>
</tbody>
</table>

Source: Stodden, Guo, Ma (2013) PLoS ONE, 8(6)
# Code Sharing Policy

<table>
<thead>
<tr>
<th>Category</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required as condition of publication, barring exceptions</td>
<td>3.5%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Required but may not affect editorial decisions</td>
<td>3.5%</td>
<td>3.5%</td>
</tr>
<tr>
<td>Encouraged/addressed, may be reviewed and/or hosted</td>
<td>10%</td>
<td>12.4%</td>
</tr>
<tr>
<td>Implied</td>
<td>0%</td>
<td>1.8%</td>
</tr>
<tr>
<td>No mention</td>
<td>82.9%</td>
<td>78.8%</td>
</tr>
</tbody>
</table>

Source: Stodden, Guo, Ma (2013) PLoS ONE, 8(6)
Findings

- Changemakers are journals with high impact factors.
- Progressive policies are not widespread, but being adopted rapidly.
- Close relationship between the existence of a supplemental materials policy and a data policy.
- No statistically significant relationship between data and code policies and open access policy.
- Data and supplemental material policies appear to lead software policy.
Publishing a “Triple”

- Gentleman and Temple Lang in 2003 proposed the “Research Compendium”
- Roughly: publication + data + code => reproducibility
Reproducibility

• In the words of Buckheit and Donoho:

  “An article about a computational result is advertising, not scholarship. The actual scholarship is the full software environment code and data, that produced the result.”

• Examples..

  • http://github.com/arokem
  • http://purl.stanford.edu/ng782rw8378
  • https://github.com/nipy/nitime
nipy/nitime

Timeseries analysis for neuroscience data

master - Merge pull request #122 from arokem/mpl_units_patch

Mpl units patch

Ariel Rokem authored and committed

Build Matrix

<table>
<thead>
<tr>
<th>Job</th>
<th>Duration</th>
<th>Finished</th>
<th>ENV</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.1</td>
<td>2 min 30 sec</td>
<td>4 months ago</td>
<td>PYTHON=python PYSUF=&quot;&quot;</td>
</tr>
<tr>
<td>31.2</td>
<td>3 min 1 sec</td>
<td>4 months ago</td>
<td>PYTHON=python3 PYSUF='3'</td>
</tr>
</tbody>
</table>
Open Science Framework

• Brian Nosek & collaborators

• Testbed for studying scientific norms and behaviors

• Reproducibility Project: 60 articles from *Journal of Personality and Social Psychology*, *Psychological Science*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Do they replicate?
Reproducibility Project: Psychology

Contributors: Alexander A. Aarts | Anita Alexander | Anna Dreber Almenberg | Leslie Cramblet Alvarez | Michael Barnett-Cowan | Elizabeth Bartness | Frank Bosco | Mark Brandt

Date Created: 4/1/2012 11:49 AM | Last Updated: 5/13/2014 5:28 AM

Do normative scientific practices and incentive structures produce a biased body of research evidence? The Reproducibility Project is a crowdsourced empirical effort to estimate the reproducibility of a sample of studies from scientific literature. The project is a large-scale, open collaboration currently involving more than 150 scientists from around the world.

The investigation is currently sampling from the 2008 issues of three prominent psychology journals - *Journal of Personality and Social Psychology*, *Psychological Science*, and *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Individuals or teams of scientists follow a structured protocol for designing and conducting a close, high-powered replication of a key effect from the selected articles.

We expect to learn about:
- The overall rate of reproducibility in a sample of the published psychology literature
- Obstacles that arise in conducting effective replications of original study procedures
- Predictors of replication success, such as whether the journal in which the original finding was published, the citation impact of the original report, and the number of direct or conceptual replications that have been published elsewhere
- Aspects of a procedure that are or are not critical to a successful direct replication, such as the setting, specific characteristics of the sample, or details of the materials.

Interested? Read more or join the project! For more information, see our executive summary or visit the New Contributors page.
Learning from the brain

- Why the brain?
  - If effects are in other variables or in interactions, do you even need the brain?
  - What if brain effects are not predicted – can we still accept these as meaningful or do we increase odds of false discovery and low reproducibility by lessening standards for construct validity?
  - “Don’t have to have a neuroscientifically meaningful construct – if your goal is biomarkers – if the goal is to have a black box – measure comes out useful” (Milham)
    - When we see these findings we should not become overly married to them.
    - However, does capture our attention – gives us an idea of where things are going

- Face validity, (common sense) concept validity (accurately reflect manipulation of construct), criteria validity (consistent with an independent gold standard)
Next steps

- Interdisciplinary research
  - Can demography help neuroscience *post hoc*? (e.g., can we establish *post hoc* how representative our sample was?)
  - Sensitivity analysis; would population weighting wash out these effects?

- Exposure, communication, and engagement – Neuroscience + Population studies

- “We don’t know what we don’t know” – Rumsfeld (Dan)