Manual of Best Practices in Transparent Social Science Research

Garret Christensen[1] Courtney Soderberg[2]

[1]Berkeley Initiative for Transparency in the Social Sciences
UC Berkeley

[2]Center for Open Science

CEGA Staff Meeting, April 2015

Introduction

Our goal:

- Detailed hands-on how-to manual for transparent social science research.
- Focus on implementing the solutions in research, less on convincing of the problems.
- Cover all aspects of a transparent research project, from beginning (study design, hypothesis registration) to end (publication, data sharing).
- Keep the manual updated.
- Encourage participation from the community. `http://github.com/garretchristensen`

Introduction

Our goal:

- Detailed hands-on how-to manual for transparent social science research.
- Focus on implementing the solutions in research, less on convincing of the problems.
- Cover all aspects of a transparent research project, from beginning (study design, hypothesis registration) to end (publication, data sharing).
- Keep the manual updated.
- Encourage participation from the community. `http://github.com/garretchristensen`

Introduction

Our goal:

- Detailed hands-on how-to manual for transparent social science research.
- Focus on implementing the solutions in research, less on convincing of the problems.
- Cover all aspects of a transparent research project, from beginning (study design, hypothesis registration) to end (publication, data sharing).
- Keep the manual updated.
- Encourage participation from the community.

`http://github.com/garretchristensen`

Introduction

Our goal:

- Detailed hands-on how-to manual for transparent social science research.
- Focus on implementing the solutions in research, less on convincing of the problems.
- Cover all aspects of a transparent research project, from beginning (study design, hypothesis registration) to end (publication, data sharing).
- Keep the manual updated.
- Encourage participation from the community.

`http://github.com/garretchristensen`

Introduction

Our goal:

- Detailed hands-on how-to manual for transparent social science research.
- Focus on implementing the solutions in research, less on convincing of the problems.
- Cover all aspects of a transparent research project, from beginning (study design, hypothesis registration) to end (publication, data sharing).
- Keep the manual updated.
- Encourage participation from the community. `http://github.com/garretchristensen`

Outline

Ethical Research

- Transparency is part of being an ethical researcher.
- Admit that we're human, subject to bias and motivated reasoning, transparency can help with this (Nosek, Spies, Motyl 2012).
- Since a lot of us run experiments, we should take IRBs seriously as part of transparency (Ch. 11–13 Morton & Williams 2010, Desposato 2014).

Ethical Research

- Transparency is part of being an ethical researcher.
- Admit that we're human, subject to bias and motivated reasoning, transparency can help with this (Nosek, Spies, Motyl 2012).
- Since a lot of us run experiments, we should take IRBs seriously as part of transparency (Ch. 11–13 Morton & Williams 2010, Desposato 2014).

Ethical Research

- Transparency is part of being an ethical researcher.
- Admit that we're human, subject to bias and motivated reasoning, transparency can help with this (Nosek, Spies, Motyl 2012).
- Since a lot of us run experiments, we should take IRBs seriously as part of transparency (Ch. 11–13 Morton & Williams 2010, Desposato 2014).

Study Design and Power–BETA

- Adequately power trials to help prevent spurious significant results.
- Practical suggestions like collaboration with other labs to mutually run each others' experiments (Open Science Collaboration 2014).

Study Design and Power–BETA

- Adequately power trials to help prevent spurious significant results.
- Practical suggestions like collaboration with other labs to mutually run each others' experiments (Open Science Collaboration 2014).

Publication Bias

Existence of the problem:

- Effect sizes diminish with sample size (Gerber, Green, Nickerson 2001)
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- Data on the complete set of experiments run shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is large. (Franco, Malhotra, Simonovits 2014)

Publication Bias

If we only write up/publish significant results, and we have no record of all the insignificant results, we have no way to tell if our 'significant' results are real, or if they're the 5% we should expect due to noise.

Registrations

Explain the solution:

- Publicly stating all research you will do, what hypotheses you will test, prospectively.
- Near universal adoption in medical RCTs. Top journals (ICMJE) won't publish if it's not registered. `http://clinicaltrials.gov`
- Even better if registry requires outcomes from after study. Currently limited, but NIH is moving on this.

Registrations

Explain the solution:

- Publicly stating all research you will do, what hypotheses you will test, prospectively.
- Near universal adoption in medical RCTs. Top journals (ICMJE) won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires outcomes from after study. Currently limited, but NIH is moving on this.

Registrations

Explain the solution:

- Publicly stating all research you will do, what hypotheses you will test, prospectively.
- Near universal adoption in medical RCTs. Top journals (ICMJE) won't publish if it's not registered. `http://clinicaltrials.gov`
- Even better if registry requires outcomes from after study. Currently limited, but NIH is moving on this.

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs. `http://socialscienceregistry.org`
  - EGAP registry `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations. `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Design-Based Publication

AKA Registered Reports, moves peer review before data gathering, results, and analysis.

1. Design a project
2. Submit
3. Reviewed based on importance of question and quality of design
4. Get in-principle acceptance
5. Follow through, and nulls get published

14 Journals, 4 more with Special Issues ▸ Link

P-Hacking

Define the problem:

- Also called fishing, researcher degrees of freedom, or data-mining.
- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshhold; significance loses its meaning.
- Not something only evil people do. It's subconcious, or simply built into statistics (Gelman, Loken 2013).

Pre-Analysis Plan

Explain the solution:

- From 3ie: “A pre-analysis plan is a detailed description of the analysis to be conducted that is written in advance of seeing the data on impacts of the program being evaluated. It may specify hypotheses to be tested, variable construction, equations to be estimated, controls to be used, and other aspects of the analysis. A key function of the pre-analysis plan is to increase transparency in the research. By setting out the details in advance of what will be done and before knowing the results, the plan guards against data mining and specification searching. Researchers are encouraged to develop and upload such a plan with their study registration, but it is not required for registration.”

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Origin: FDA's Guidance for Industry

"E9 Statistical Principles for Clinical Trials" (1998) ▸ Link

§V Data Analysis Considerations

1. Prespecification of the Analysis
2. Analysis Sets
3. Missing Values and Outliers
4. Data Transformation
5. Estimation, Confidence Intervals, and Hypothesis Testing
6. Adjustment of Significance and Confidence Levels
7. Subgroups, Interactions, and Covariates
8. Integrity of Data and Computer Software Validity

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
   - Explain mean effects, FWER, FDR using Anderson (JASA 2008).
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

Examples

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`, 6 published papers:
  - Sierra Leone CDD, Oregon Medicare, Turkey Job Training, El Salvador TOMS, two in Indonesia (Olken et al.)
- OSF: Hawkins, Fitzgerald, Nosek—Conception Risk and Prejudice

Wide range of when exactly to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

Replication

1. The Problem (JMCB Project)
2. Project Protocol, Reporting Standards
3. Organizing Workflow
4. Code & Data Sharing

Project Protocol, Reporting Standards

Make sure you report everything another researcher would need to replicate your research.

- Find the appropriate reporting standard for your field and follow it: `http://www.equator-network.org/`
- Report the nuts and bolts of the project implementation in a detailed protocol: `http://www.spirit-statement.org`

Workflow

"Reproducibility is just collaboration with people you don't know, including yourself next week"
—Philip Stark, UC Berkeley Statistics

Workflow

Practical coding and organizational suggestions

- Long (2008) *The Workflow of Data Analysis Using Stata*
  - Making any changes to a file that has been posted/shared means it gets a new name.
  - Use version commands to ensure others get same results.
- Literate programming (extensive commenting, making the aim of code reading by a human)
- R Markdown, integration of analysis and output.

Workflow

Practical coding and organizational suggestions

- Long (2008) *The Workflow of Data Analysis Using Stata*
  - Making any changes to a file that has been posted/shared means it gets a new name.
  - Use version commands to ensure others get same results.
- Literate programming (extensive commenting, making the aim of code reading by a human)
- R Markdown, integration of analysis and output.

Workflow

Practical coding and organizational suggestions

- Long (2008) *The Workflow of Data Analysis Using Stata*
  - Making any changes to a file that has been posted/shared means it gets a new name.
  - Use version commands to ensure others get same results.
- Literate programming (extensive commenting, making the aim of code reading by a human)
- R Markdown, integration of analysis and output.

Workflow

Practical coding and organizational suggestions

- Long (2008) *The Workflow of Data Analysis Using Stata*
  - Making any changes to a file that has been posted/shared means it gets a new name.
  - Use version commands to ensure others get same results.
- Literate programming (extensive commenting, making the aim of code reading by a human)
- R Markdown, integration of analysis and output.

Workflow

Practical coding and organizational suggestions

- Long (2008) *The Workflow of Data Analysis Using Stata*
  - Making any changes to a file that has been posted/shared means it gets a new name.
  - Use version commands to ensure others get same results.
- Literate programming (extensive commenting, making the aim of code reading by a human)
- R Markdown, integration of analysis and output.

Data Sharing

Post your code and your data in a trusted public repository.

- Find the appropriate repository: `http://www.re3data.org/`
- Repositories will last longer than your own website.
- Repositories are more easily searchable by other researchers.
- Repositories will store your data in a non-proprietary format that won't become obsolete.

Conclusion

OK, how do I implement this in my own research?
Read the manual.

To do:

- P-curve
- Dynamic documents with R Markdown
- If you have suggestions, it's on GitHub for a reason.
  `https://github.com/garretchristensen/BestPracticesManual` ▸ Link