SIOP REPRODUCIBLE RESEARCH GUIDELINES

Considerations for Submissions that Include Reproducible Research

- Read this document completely for descriptions of reproducible research (RR) at SIOP
- All session types may include partial or full participation in RR.
- Check the box for “Reproducible Research” in the online form for your Submission.
- Consider including “Reproducible” in the title and/or description of specific RR presentation(s) to increase visibility with other attendees.
- Briefly describe the type of data/mock data (if any) and/or code that will be reproducible.
- Please do not include any reproducible data or code with your submission. Instead, simply describe the type(s) of reproducible files (including their coding language) you intend to make available by the time of your presentation in the required 500 character maximum description.
- Upon acceptance of your submission, SIOP will provide additional instruction for uploading/sharing your code files to ensure they are searchable and reproducible.
- Any additional questions about SIOP Reproducible Research not answered below should be directed to program@siop.org.

Why Reproducibility at SIOP?

Within the past few years, the field of data science has made it easier than ever to share analytical code and techniques with others. During this time, researchers at SIOP and other scientific research outlets (e.g., *Nature* and *Psychological Science*) have called for increased replication and reproducibility of research methods and findings. As a result, researchers across a wide range of disciplines are now presenting “reproducible research” (RR) with great success. Reproducible research approaches can help advance our field by quickly sharing data (or mock data), process, and code for analysis.

Beginning with SIOP 2017 in Orlando, the Program Committee encourages and supports research presentations that strive for “reproducibility.” Authors of all accepted RR presentations will receive RR Presentation Guidelines from the Program Committee in December with further details and best practices for your RR presentation(s) at SIOP.

What is “Reproducible Research” and Why is It Important?

Reproducible research refers to the process of fully documenting the steps that were taken to conduct a research study so that anyone can efficiently replicate the study’s findings, understand exactly what analyses were conducted, or update an analysis (Fomel & Claerbout, 2009). One method of creating a reproducible research process is to embed both the data and the analysis code (or, syntax) within a single research document using **Markdown** programming language (for which many programs are available).

Reproducible research provides a number of potential benefits:

- Transparency and replication, which are core values of science
- It is much easier to share details about data and quantitative methods shortly after conducting analysis versus waiting to provide this information “upon request” after presentation or publication
SIOP REPRODUCIBLE RESEARCH GUIDELINES

- As studies become more complex and datasets get larger, reproducible research allows both academics and practitioners to replicate research findings and leverage the latest analytical techniques
- Sharing data and analysis details in Markdown files expedites the data analysis process and advancement of our science
- Sharing code or data can more formally document your analysis process
- Sharing structured analyses allows researchers in our field to more efficiently conduct similar investigations or update results with new data
- Sharing data and code as a community reinforces diligence and integrity in our methods while making it easier to conduct data cleaning, transformation, and statistical analyses.

Various Degrees of Reproducibility

Organizing your code or syntax in a way that flows logically and surrounds code with natural language explanations is most useful (see literate programming; Knuth, 1992). Embedding code in this way is useful for documents that are updated over time with new data, and allows all audiences to understand and begin using code quickly. Certain programs like Markdown make it easy to embed code within the natural language flow (see examples here and here).

Sharing with SIOP. While it is most helpful to the I-O psychology community to provide all of the necessary materials (e.g., code, data, decision points, etc.) to fully replicate a study, any materials that can be shared and that authors feel comfortable providing will be most welcome!

- **Data**: Nothing beats the dataset used for your research; however, if this data is restricted, then changing variable names and producing “mock” datasets can be just as helpful.
- **Code/Syntax**: RR presentations became popular with the R stats language (see tutorial), but the process would also work for syntax in SAS, SPSS, Notepad, or other file formats.

**Industry Standards.** For a more detailed treatment of the varying degrees of sharing available to researchers, the Center for Open Science (COS) provides the Transparency and Openness Promotion (TOP) Guidelines which discuss three levels of transparency across eight categories.

**References**