Reproducible Research

Grieven Otieno

February 26, 2016

Objectives

In this session we shall cover the following:

1. Introduce the concept of reproducible research
2. Compare reproducibility and replicability
3. Identify ways to make work reproducible
4. Different platforms available to make work reproducible

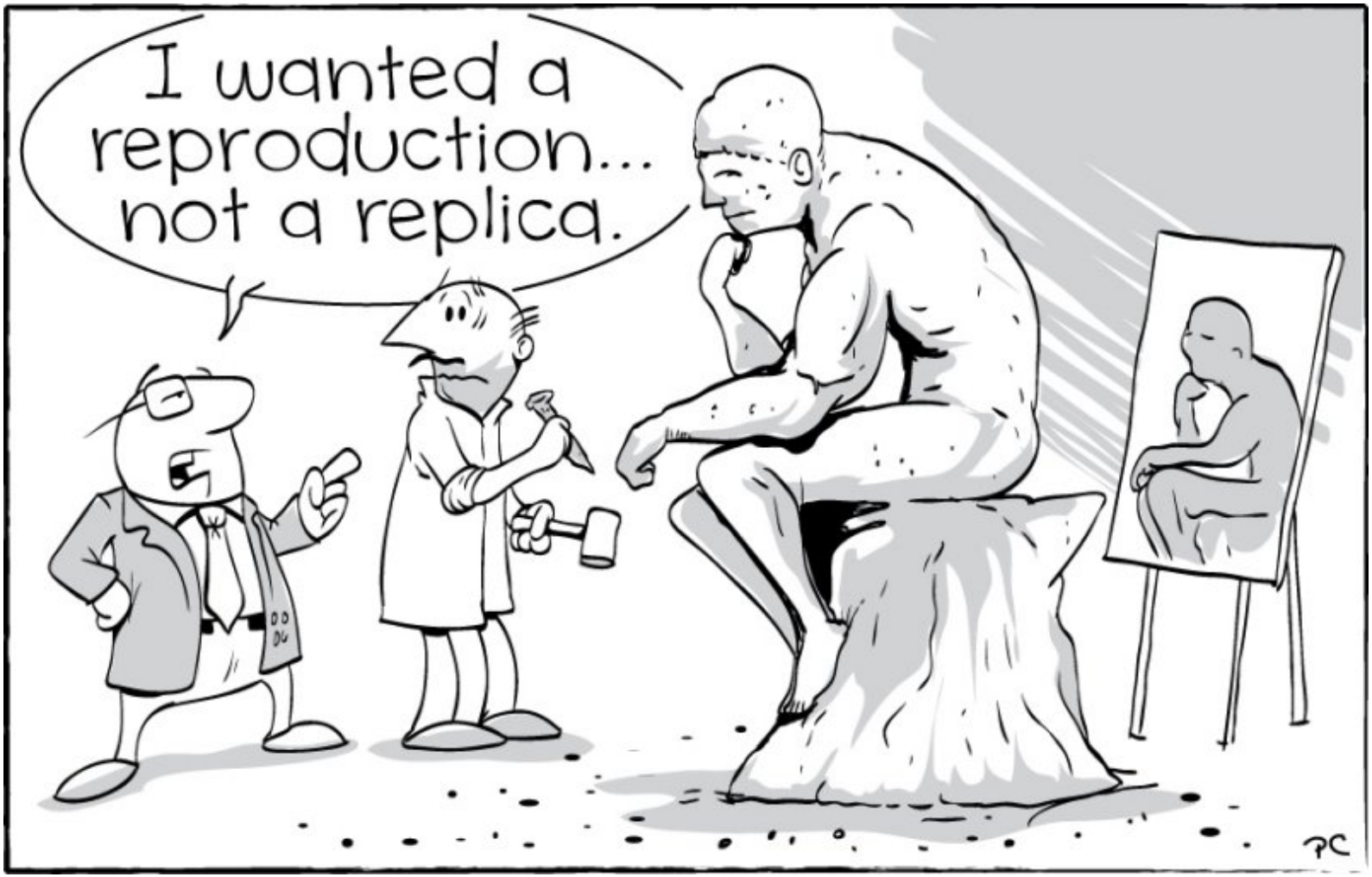
Replication

- Replication involves the process of repeating a study using the same methods, different subjects and different experimenters.
- Some studies cannot be replicated due to:
  - Time constraints
  - Financial constraints
  - Uniqueness of the original study

Reproducible Research

- Reproducible Research involves methods and approaches which enhance the integrity of research and promote consistent expectations within the scientific community.
- Generally, reproducibility ensures that independent scientists can reproduce published results by using the same procedures and data as the original investigator(s). (Laine et al.,2007)
- A good practice in reproducible research is to document as much as possible all the methods, processes and assumptions made when conducting your analysis

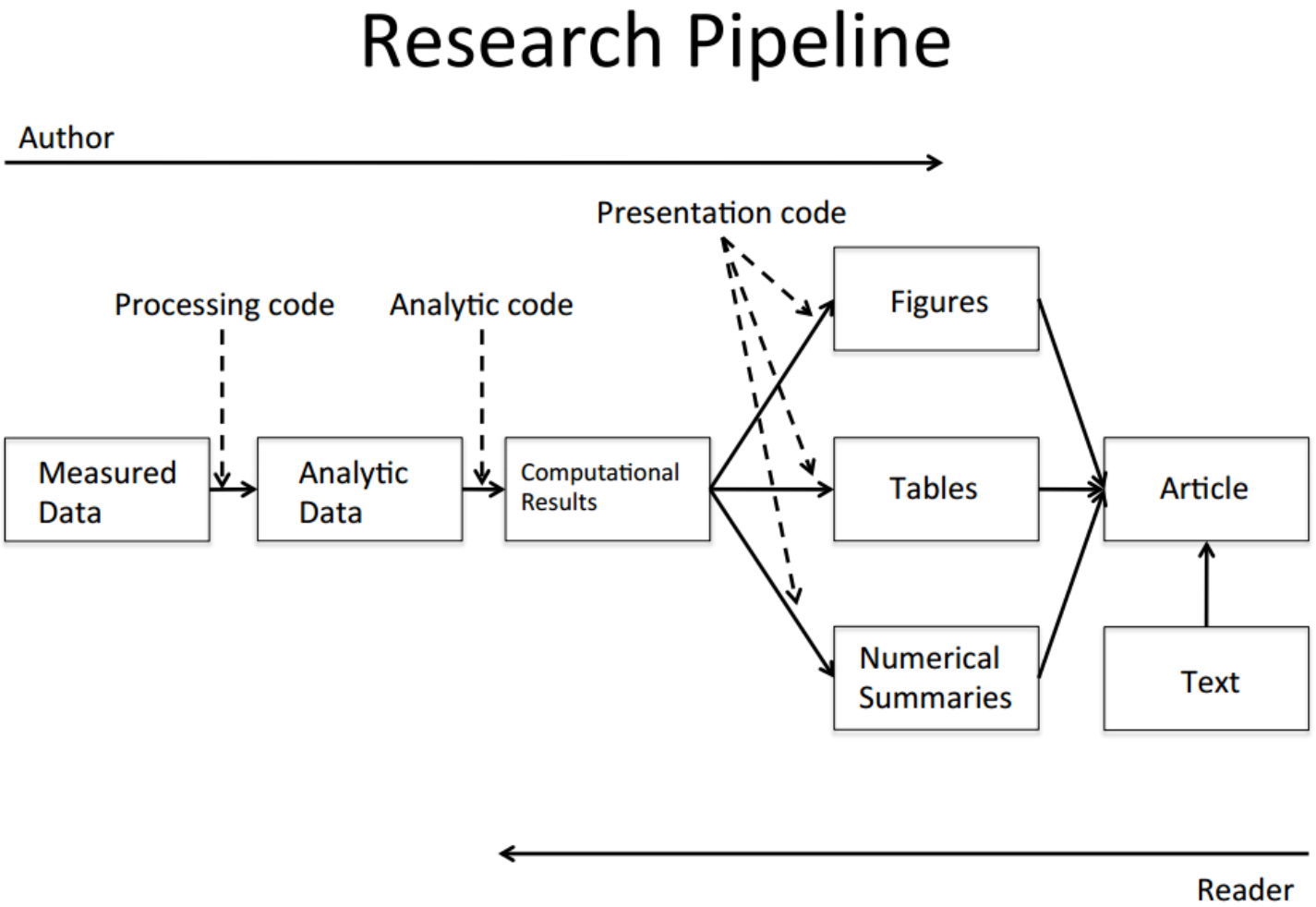
Reproducibility v/s Replicability



How do I make my work reproducible?

- Ensure that the following are available:
  - Analytic data
  - Analytic code
  - Documentation of data and code
- These steps help during review of work and enable continuity of your work.

General Process of Reproducibility



R Markdown

- In R mark down, the analysis code is divided into text and code **"chunks"**
- Each code chunk loads data and executes to give the results
- Article text explains what is going on
- For Mac users install Mactex while windos users use Miktex

Sample R code in Markdown

```
```{r, echo=FALSE, message=F}
setwd("/Users/gpotieno/Documents/Presentations/Internship Program/Data")

# 1.0 Load packages
library(foreign)
#library(epicalc)
library(epitools)
library(epiR)
library(fmsb)
library(knitr)

# 2.0 Import the child deaths data
childdeaths <- read.table("child_deaths.csv", header=TRUE, sep=",")
```
```

Summary

- Reproducible research is important as a **minimum standard**, particularly for studies that are difficult to replicates
- There is a growing number of tools that can be used to create reproducible documents, so we should be able to account for every process in our research

Questions