Egocentric network analysis with R

[Raffaele Vacca](http://www.raffaelevacca.com/)

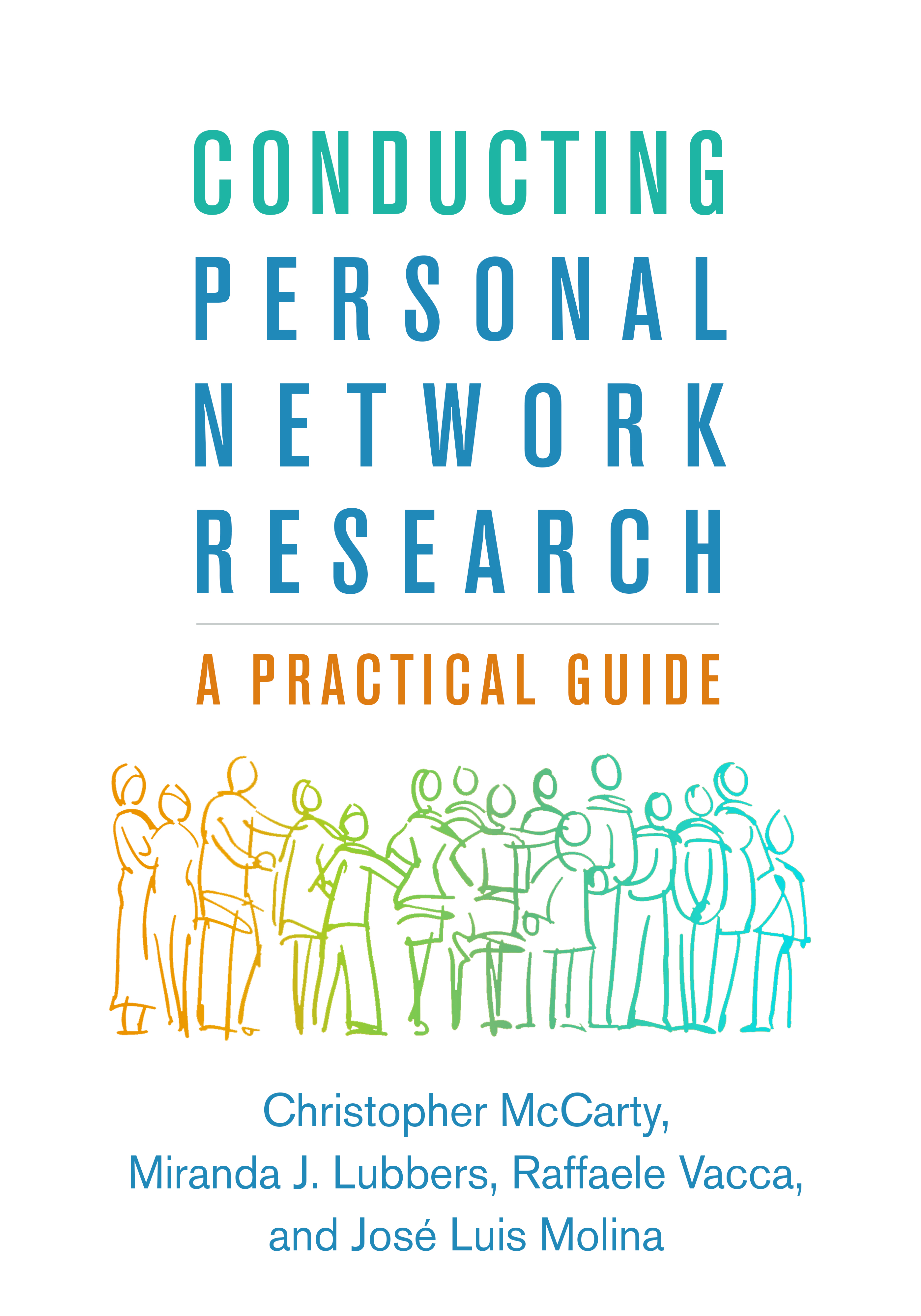
r Sys.Date()

# Introduction

This book is a companion to [my workshop](http://www.raffaelevacca.com/teaching/egocentric-r/) on egocentric network analysis with R. Over the past several years I have taught this workshop at different conferences and summer schools on social network analysis, personal networks, and network science – such as [INSNA Sunbelt](https://www.insna.org/) conferences, [NetSci](https://netscisociety.net/home) conferences, and the [UAB Barcelona Course on Personal Network Analysis](https://sway.office.com/1TpXMhGqKa7fTsAC). The book is a work in progress and I’ll keep updating it as I continue to teach the workshop and related courses.

A Spanish translation of part of this book has appeared in the journal [REDES - Revista hispana para el análisis de redes sociales](https://revistes.uab.cat/redes/article/view/v31-n2-vacca). My colleagues and I provide a more in-depth discussion of personal network analysis and the methods covered in this workshop in our textbook on personal network research [@mccarty\_conducting\_2019]. See the bibliography at the end of the book for a list of other useful references on egocentric or personal network analysis.

Feel free to [contact me](http://www.raffaelevacca.com/) to know more about this workshop and how to take it, or to give me feedback or report any issue about this book.

[](https://www.guilford.com/books/Conducting-Personal-Network-Research/McCarty-Lubbers-Vacca-Molina/9781462538386/authors)

## Workshop materials

The materials for this workshop consist of this book and the **workshop folder**. The workshop folder can be downloaded [here](http://www.raffaelevacca.com/teaching/egocentric-r/folder/) and includes:

* Scripts subfolder: all code shown in this book.
* Data subfolder: all data we’re going to use.
* Slides subfolder: the workshop slides.
* Egocentric\_R.Rproj: the workshop R project file.

In the Scripts subfolder you’ll see seven script (.R) files. You can access and run the R code in each script by opening the .R file in RStudio. Each script corresponds to one of the following chapters (see the table of contents):

1. [Basics of the R language](#basics).
2. [Representing and visualizing ego-networks in R](#represent).
3. [Analyzing ego-network composition](#composition).
4. [Analyzing ego-network structure](#structure).
5. [Modeling tie- or alter-level variables with multilevel models](#multilevel).
6. [Introduction to the egor package](#egor).
7. [Supplementary topics](#supplementary).

## Workshop setup

To take this workshop you need to:

1. Download the last version of **R** [here](https://cran.r-project.org/mirrors.html) (select a location near you)
   * Follow instructions to install R in your computer
2. Download **RStudio** (free version) [here](https://www.rstudio.com/products/rstudio/download/)
   * Follow instructions to install RStudio in your computer
3. Install the **R packages** listed [below](#packages)
   * Open RStudio and go to Top menu > Tools > Install packages...
   * Install each package in the list
4. Bring your **laptop** to the workshop
5. Download the **workshop folder** [here](http://www.raffaelevacca.com/teaching/egocentric-r/folder/) (I recommend that you do this in class at the beginning of the workshop so as to download the most updated version of the folder).
   * Click on the link > Click on Download in the top-right > Click on Direct download
   * Then save the folder to your computer

Once in class, go to the workshop folder (point 1 above) and double-click on the workshop *R project* file (Egocentric\_R.Rproj) in it. That will open RStudio.

**NOTE:** It’s very important that you save the workshop folder *as downloaded* to a location in your computer, and open Egocentric\_R.Rproj *within that folder*. By doing so, you will be opening RStudio *and* setting the workshop folder as your [R working directory](#starting-R-and-loading-packages). All our R scripts assume that working directory. In particular, they assume that the Data subfolder is in your R working directory. You can type getwd() in your R console to see the path to your R working directory and make sure that it’s correctly pointed to the location of the workshop folder in your computer.

### Required R packages

* For descriptive statistics
  + [janitor](https://cran.r-project.org/web/packages/janitor/vignettes/janitor.html)
  + [skimr](https://cran.r-project.org/web/packages/skimr/vignettes/skimr.html)
* For network data, measures, visualization
  + [egor](https://github.com/tilltnet/egor)
  + [ggraph](https://github.com/thomasp85/ggraph)
  + [igraph](https://igraph.org/)
  + [intergraph](https://mbojan.github.io/intergraph/)
  + [statnet](http://statnet.org/)
  + [tidygraph](https://github.com/thomasp85/tidygraph)
* To fit multilevel statistical models and view their results
  + [broom.mixed](https://cran.r-project.org/web/packages/broom.mixed/vignettes/broom_mixed_intro.html)
  + [car](https://cran.r-project.org/web/packages/car/index.html)
  + [ggeffects](https://strengejacke.github.io/ggeffects/)
  + [lme4](https://github.com/lme4/lme4)
* General
  + [magrittr](https://magrittr.tidyverse.org/)
  + [tidyverse](https://www.tidyverse.org/). This is a collection of different packages that share a common language and set of principles, including dplyr, ggplot2, and purrr. See [Wickham and Grolemund (2017)](http://r4ds.had.co.nz/) for more information.

### RStudio settings

RStudio gives you the ability to select and change various settings and features of its interface: see the Preferences... menu option. These are some of the settings you should pay attention to:

* Preferences... > Code > Editing > Soft-wrap R source file. Here you can decide whether or not to wrap long code lines in the editor. When code lines in a script are *not* wrapped, be aware that some code will be hidden if script lines are longer than your editor window’s width (you’ll have to scroll right to see the rest of the code). With a script (.R) file open in the editor, try both options (checked and unchecked) to see what you’re more comfortable with.
* Preferences... > Code > Display > Highlight R function calls. This allows you to highlight all pieces of code that call an R function (“command”). I find function highlights very helpful to navigate a script and suggest that you check this option.

## Data

This workshop uses real-world data collected in 2012 with a personal network survey among 107 Sri Lankan immigrants in Milan, Italy. Out of the 107 respondents, 102 reported their personal network. All data are in the Data subfolder.

The data files include ego-level data (gender, age, educational level, etc. for each Sri Lankan respondent), alter attributes (alter’s nationality, country of residence, closeness to ego etc.), and information on alter-alter ties. Each personal network has a fixed size of 45 alters. Information about data variables and categories is available in ./Data/codebook.xlsx.

All data objects are saved as R objects in the R data file data.rda. Data objects are the following:

* ego.df: A data frame with ego-level attributes for all respondents (egos).
* alter.attr.all: A data frame with alter-level attributes for all alters from all respondents.
* gr.list: A list. Each list element is one ego-network stored as an igraph object.
* alter.attr.28: A data frame with alter-level attributes only for alters nominated by ego ID 28.
* gr.28: The ego-network of ego ID 28 stored as an igraph object.
* gr.ego.28: The same as gr.28, but with the node of the ego included in the igraph object.

The R data objects above were imported from raw csv data files. All csv files are in the ./Data/raw\_data/ subfolder:

* ego\_data.csv: A csv file with ego-level data for all the egos.
* alter\_attributes.csv: A single csv file including attributes of all alters from all egos.
* alter\_ties\_028.csv: The edge list for ego ID 28’s egocentric network.
* alter\_attributes\_028.csv: The alter attributes in ego ID 28’s egocentric network.
* adj\_028.csv: The adjacency matrix for ego ID 28’s egocentric network.
* alter\_ties.csv: A single csv file with the edge list for all alters from all egos.

For most of the workshop we will directly use R data objects in data.rda. We will *not* focus on importing ego-network data from outside sources (for example, csv files). However, Section @ref(egor-import) shows you how the data objects in data.rda were created by importing csv files with the egor package. Section @ref(import-igraph) covers importing ego-network data using just tidyverse and igraph.

## Author and contacts

I am an assistant professor of sociology at the [University of Florida](https://soccrim.clas.ufl.edu/). My main research and teaching interests are social networks, migration, health inequalities, and studies of science. I also teach and do research on data science, statistics, and computational methods for the social sciences. More information about me, my work and my contact details is [here](http://www.raffaelevacca.com/about-me/).

{r include=FALSE} # automatically create a bib database for R packages knitr::write\_bib(c( .packages(), 'bookdown', 'knitr', 'rmarkdown' ), 'packages.bib')