



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

SBA SISTEMA BIBLIOTECARIO
DI ATENEO

Repli**cabilità e riproducibilità.**

Buone pratiche per la scienza aperta

Conversazioni sull'Open Science

24 Aprile 2025

Ufficio Biblioteca Digitale CAB

Costantina Bruno, Raffaella Cattarinussi

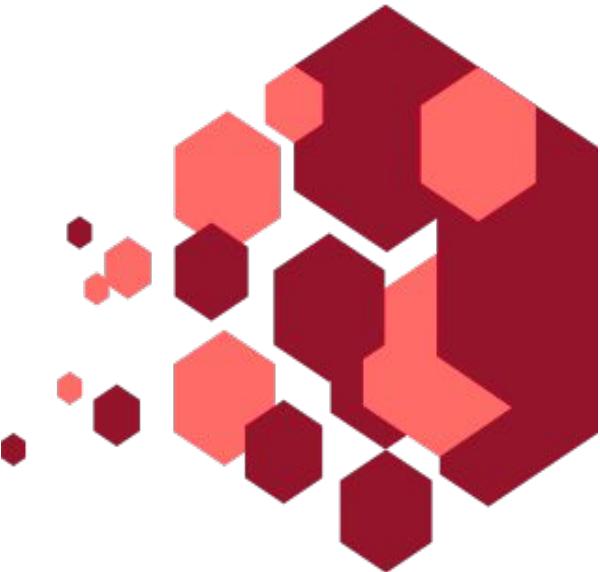
Gruppo Supporto alla ricerca SBA

Dott. Filippo Gambarota

Dipartimento di Psicologia dello Sviluppo e della Socializzazione -
DPSS



Le Conversazioni sull'Open Science



Serie di incontri mensili

Ogni terzo o quarto giovedì del mese alle ore 14.30

Online

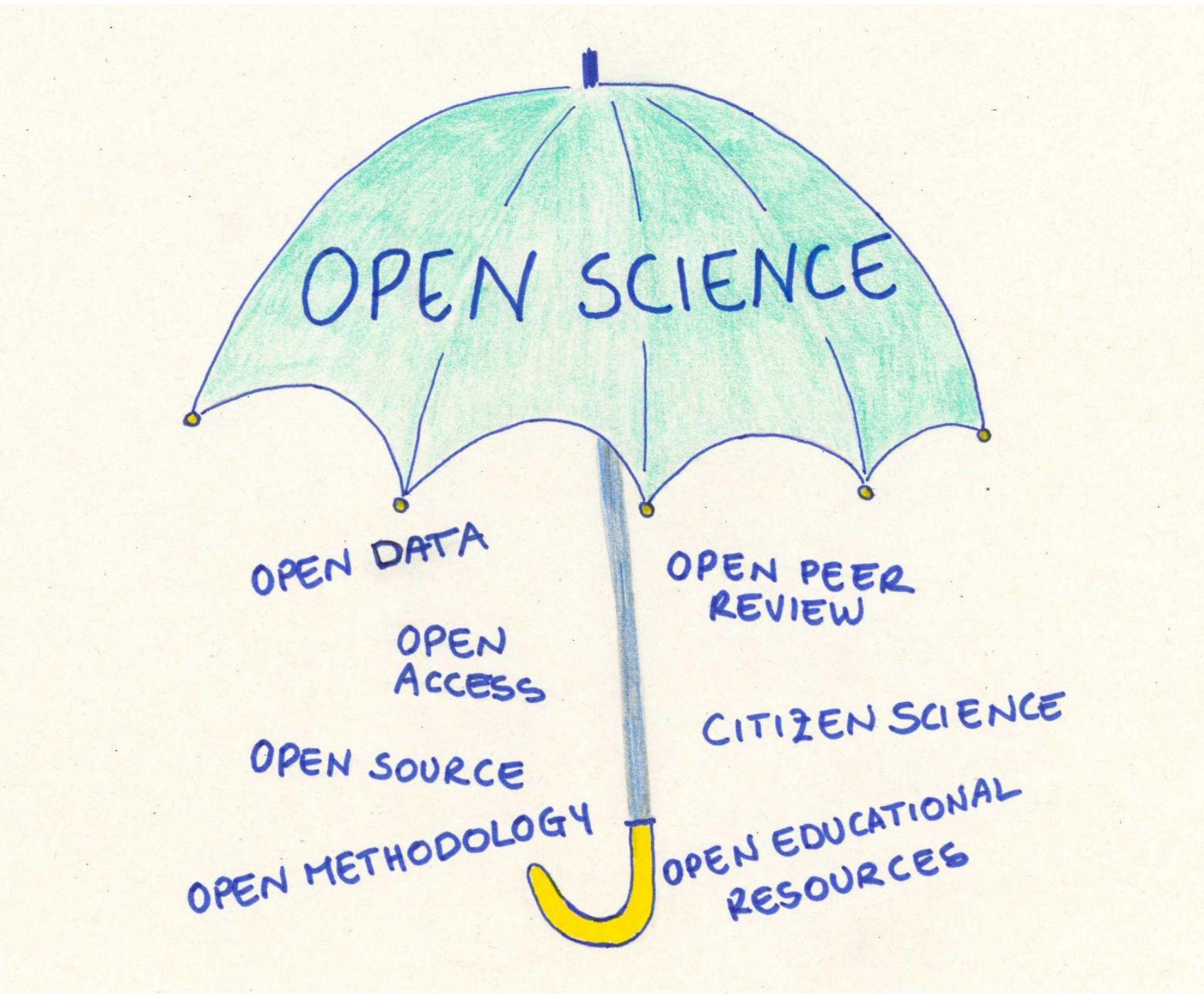
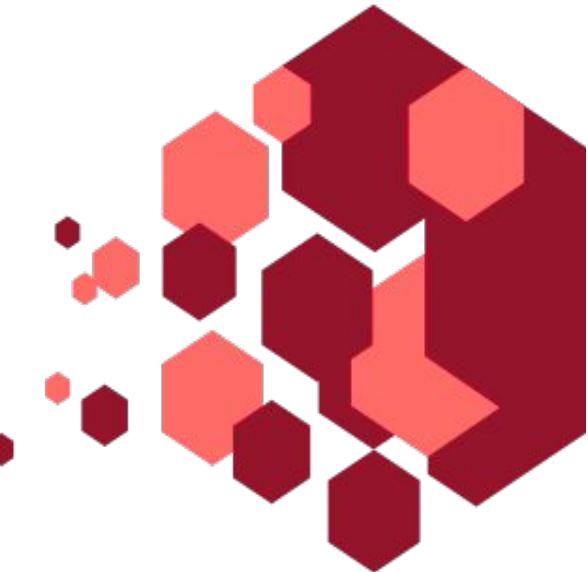
Partecipazione gratuita, su prenotazione

Momenti informativi su temi e novità dal mondo della Scienza Aperta e occasioni per conoscere e approfondire il funzionamento dei servizi di Ateneo, allo scopo di aumentare l'efficienza della ricerca e rendere la conoscenza scientifica rintracciabile, accessibile, interoperabile e riutilizzabile.

Per rimanere aggiornati sulle prossime Conversazioni:

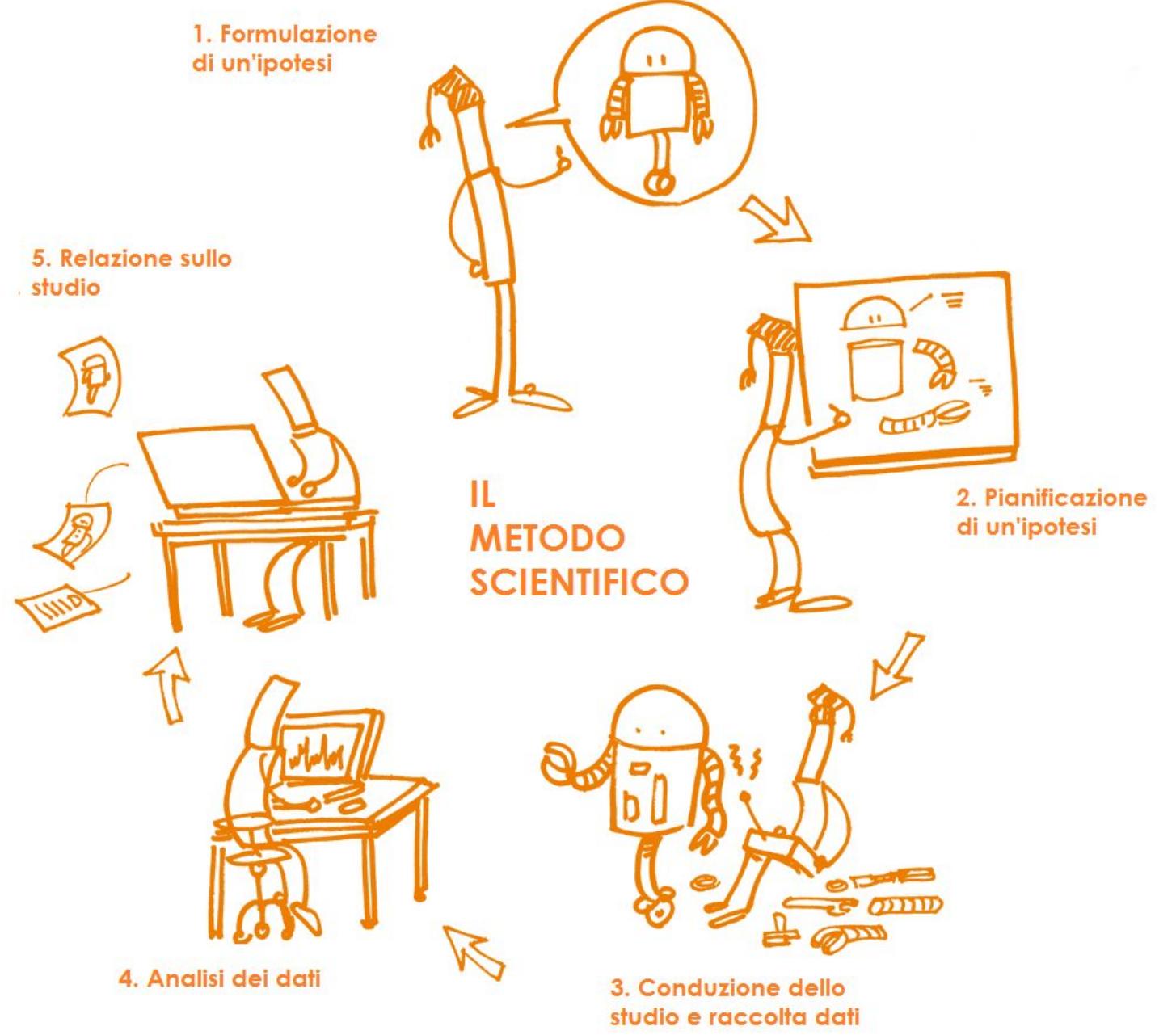
- Sito del Sistema bibliotecario di Ateneo
- Newsletter SBA
- Eventbrite

OPEN SCIENCE



DI COSA PARLIAMO

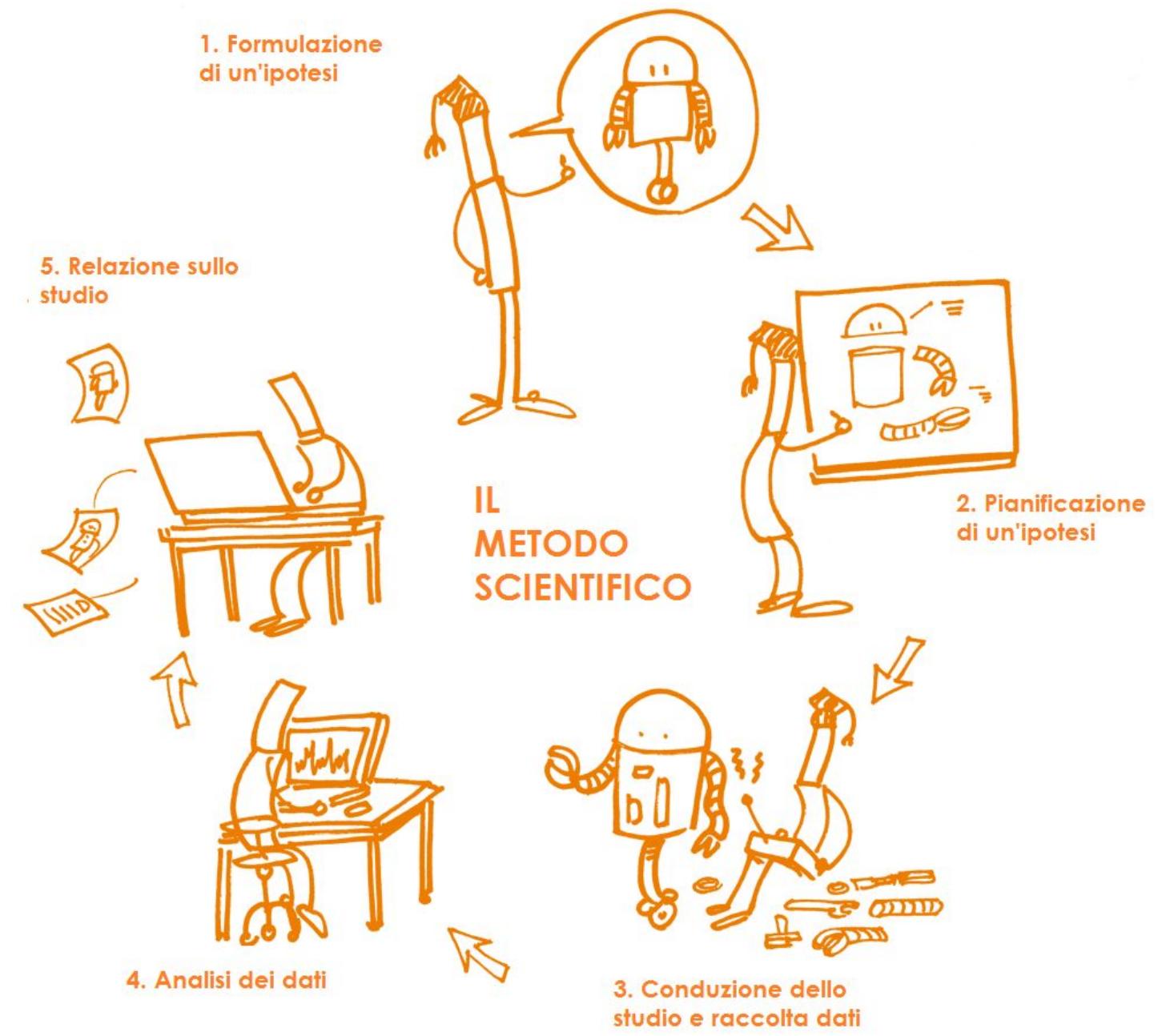
Riproducibilità
Replicabilità
Crisi della riproducibilità e della replicabilità



RIPRODUCIBILITÀ - REPLICABILITÀ

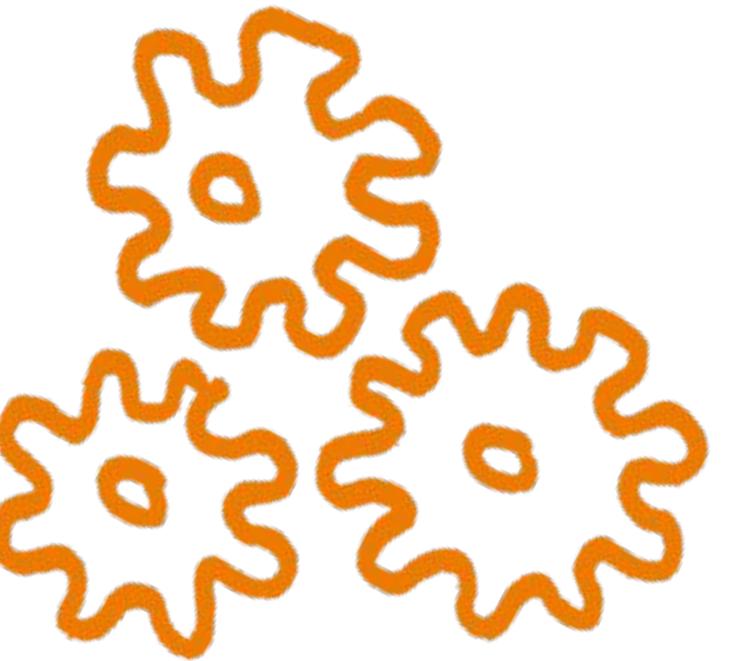
Riproducibilità:
applicabilità dello stesso metodo
agli stessi dati, verificando il raggiungimento
degli stessi risultati

Replicabilità:
applicabilità dello stesso metodo a diversi dati di
partenza, ottenendo diversi risultati



BUONE PRATICHE

1. Formulazione del quesito di ricerca (preregistrazione)
2. Individuazione di una procedura (protocollo)
3. Descrizione dettagliata (notebook)
4. Identificazione univoca dei materiali
5. Utilizzo di software open source
6. Deposito in archivio aperto
7. Verifica dei risultati (rispetto alla preregistrazione)
8. Pubblicazione in OA





Un esempio: PSICOSTAT



Gruppo interdipartimentale tra psicologia e statistica, impegnato nei settori della psicologia quantitativa, psicometria, testing psicologico e statistica.

<https://psicostat.dpss.psy.unipd.it/index.html>

Filippo Gambarota

Dipartimento di Psicologia dello Sviluppo e della Socializzazione - DPSS



Psicostat



Siamo un gruppo di ricerca fondato presso il Dipartimento di Psicologia dello Sviluppo e della Socializzazione (Università di Padova).

Siamo formati da Statistici e Psicologi di varie discipline (clinici, sperimentali, del lavoro)

Siamo interessati ad applicare e sviluppare metodi statistici nella ricerca psicologica

Siamo particolarmente interessati alla riproducibilità e replicabilità della ricerca e alla meta-science.



<https://psicostat.dpss.psy.unipd.it/>
<https://github.com/psicostat>

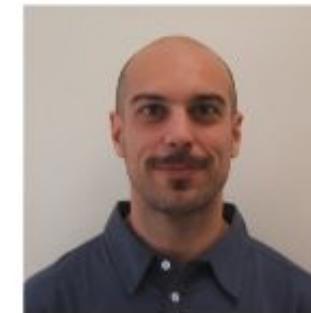
Psicostat, core team



Gianmarco Altoè



Massimiliano Pastore



Livio Finos



Giulia Calignano



Ottavia Epifania



Filippo Gambarota



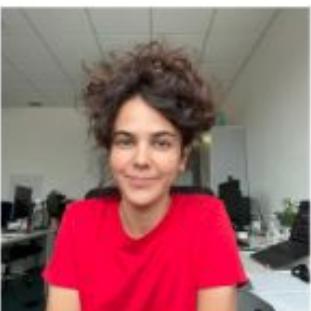
Enrico Toffalini



Luca Menghini



Ambra Perugini



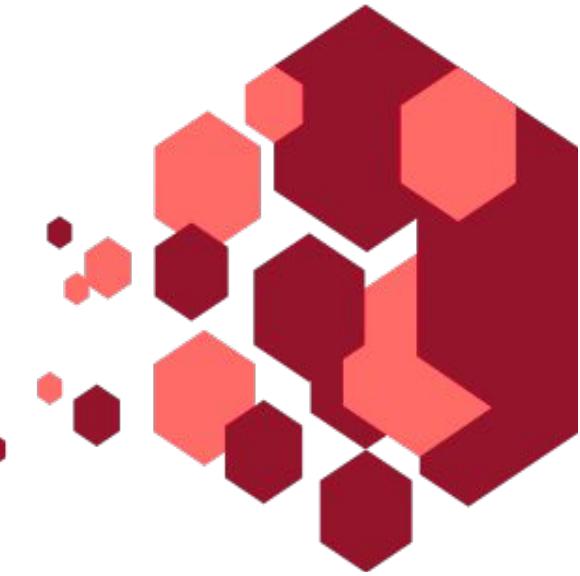
Margherita Calderan



Marco Tullio Liuzza



Tommaso Feraco



Riproducibilità e replicabilità



Riproducibilità

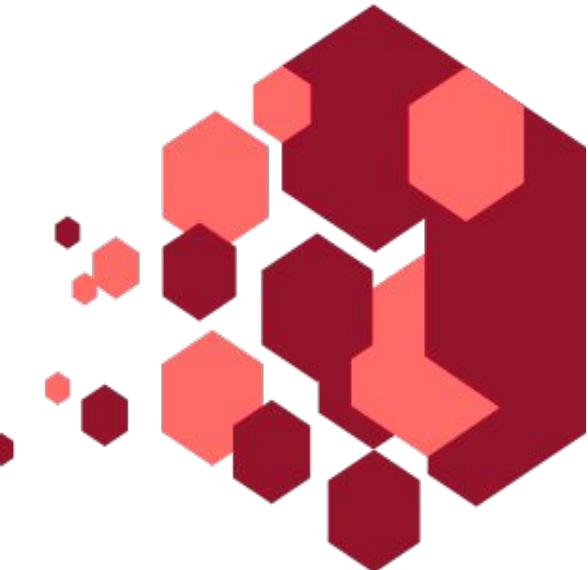
“Reproducibility includes the act of a second researcher recomputing the original results, and it can be satisfied with the availability of data, code, and methods that makes that recomputation possible.”

Più comunemente questo tipo di riproducibilità viene definita **computazionale** essendo relativa alla possibilità di riprodurre risultati, figure, tabelle, etc.



Come agevolare la riproducibilità?

- Utilizzo di **strumenti open-source**, gratuiti ed ampiamente utilizzati
- Utilizzo di **scripting** (scrivere codice) o di strumenti che forniscono il codice
- Utilizzo di **formati di file aperti** e facilmente utilizzabili (e.g., csv vs xlsx)
- Utilizzo di sistemi di **literate programming** per produrre documenti
- **Condivisione in spazi appositi** di dati (dove possibile), codici, procedure, etc. in modo comprensibile ed utilizzabile



Step di una ricerca (riproducibile)

- Raccolta dati
- Gestione dati
- Analisi dati
- Reporting
- Condivisione



Raccolta dati

In Psicologia sperimentale e neuroscienze, solitamente si svolgono esperimenti programmati in laboratorio.

- Utilizzo di software **open-source** e possibilmente **gratuiti**
- Utilizzo di software diffusi, aggiornati e supportati
- Limitare passaggi *a mano*, non riproducibili a favore di script per eseguire ogni step (creare e/o modificare immagini, etc.)



Raccolta dati, qualche esempio

Psychopy (gratuito e open-source)

stackoverflow About Products OverflowAI

Home Questions Tags Discussions Labs Chat Users Companies COLLECTIVES + TEAMS

Search Results

Results for psychopy
Search options not deleted

500 results Relevance Newest More ▾

6 votes **Exiting PsychoPy Window**
I was stepping through some **PsychoPy** code in the coder view, and ran the Window function:
<http://www.psychopy.org/api/visual/window.html> This had the helpful side-effect of opening a grey win...
4 answers 9k views python macos psychopy alexplanation 1,498 asked Nov 12, 2013 at 20:44

1 vote **Python psychopy imageStim**
I'm using **psychopy** to get an image on a window like this: from **psychopy** import visual window = visual.Window(size= (1000,600), units = 'pix', pos = (10,10)) stimulusimage =...
1 answer 1k views python psychopy Caeline 119 asked Nov 19, 2015 at 11:41

Eprime (pagamento e ~open-source)

stackoverflow About Products OverflowAI

Home Questions Tags Discussions Labs Chat Users Companies COLLECTIVES + TEAMS

Search Results

Results for eprime
Search options not deleted

60 results Relevance Newest More ▾

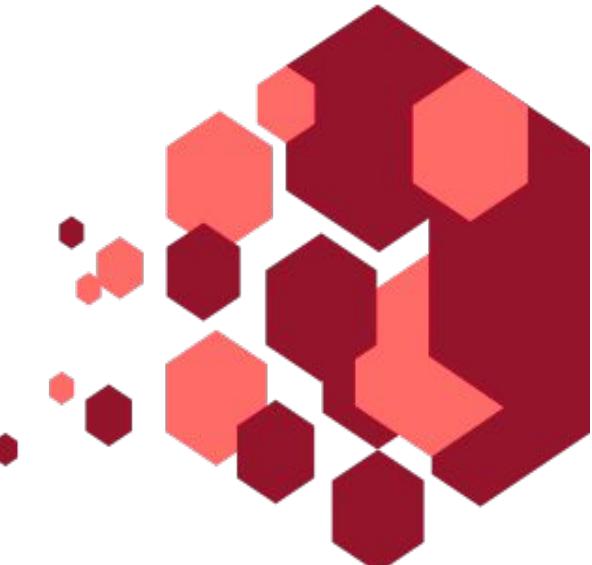
0 votes **eprime/vba highest value of 4 attributes**
I'm using **eprime** (psychology software tools) to design an experiment. ... Thus, I actually need a smart if-statement or loop or something like that telling **eprime** to look for the highest of these values - if the...
1 answer 102 views vba loops if-statement Jan 61 asked Jul 16, 2021 at 12:37

1 vote **Thread in C++ DLL called by EPrime causes exception**
The function is thus declared in **EPrime**: Declare Function moop stdcall Lib "meep.dll" () As Long Dim oPtr As Long oPtr = moop() and the DLL code looks like this: void* event_loop_func(void* eye_track...
0 answers 81 views c++ windows multithreading dll dk 561 asked Apr 18, 2016 at 8:06



Step di una ricerca (riproducibile)

- Raccolta dati
- Gestione dati
- Analisi dati
- Reporting
- Condivisione



Gestione dati

- Preferire formati non proprietari (.txt, .csv vs .xlsx)
- Avere sempre versioni originali e grezze dei dati
- Agire sempre in modo riproducibile e reversibile sui dati
- Non solo questione di condividere il codice MA di essere in grado di eseguirlo.
Senza una licenza di un software a pagamento non è possibile eseguire dei codici in teoria perfettamente riproducibili.



Step di una ricerca (riproducibile)

- Raccolta dati
- Gestione dati
- Analisi dati
- Reporting
- Condivisione

Analisi dati



Preferire un linguaggio di programmazione (R, Python, etc.) piuttosto che un software specifico (spesso closed-source ed a pagamento)

- Script sono riproducibili e il software è molto utilizzato
- Eventuali errori sono facilmente risolvibili (community)
- Nuovi metodi, modelli statistici, etc. sono rapidamente introdotti e sviluppati



Un esempio con R

Community attiva e costantemente in aggiornamento. Maggiore probabilità di parlare un “linguaggio” comune.

Contributed Packages

Available Packages

Currently, the CRAN package repository features 22254 available packages.

[Table of available packages, sorted by date of publication](#)

[Table of available packages, sorted by name](#)

[CRAN Task Views](#) aim to provide some guidance which packages on CRAN are relevant for tasks related to a certain topic. They provide tools to automatically install all packages from each view. Currently, 48 views are available.

Installation of Packages

Please type `help("INSTALL")` or `help("install.packages")` in R for information on how to install packages from this repository. The manual [R Installation and Administration](#) (also contained in the R base sources) explains the process in detail.

Package Check Results

<https://cran.r-project.org/>



Step di una ricerca (riproducibile)

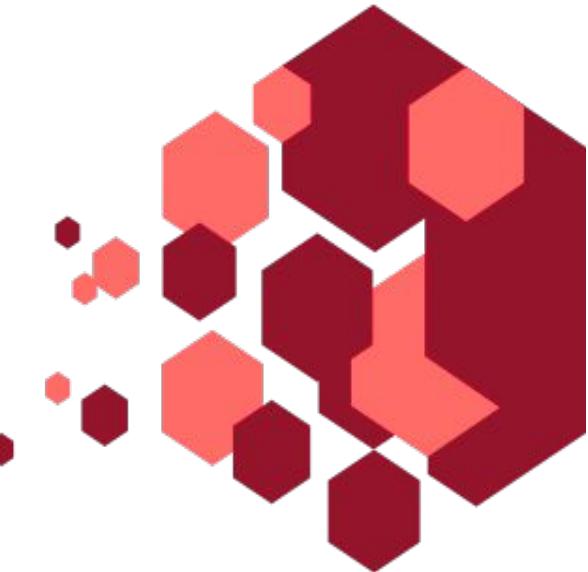
- Raccolta dati
- Gestione dati
- Analisi dati
- Reporting
- Condivisione



Reporting

Anche la condivisione dei risultati (pubblicazioni scientifiche, conferenze, etc.) può agevolare o limitare la riproducibilità.

- Figure e/o tavelle non riproducibili
- Errori nelle analisi o nel riportare i risultati
- Poca trasparenza nel passaggio dagli script di analisi al risultato finale



Literate programming

La soluzione migliore è usare un framework di literate programming.
Ovvero un documento che contenga **codice** (per analisi, figure, tabelle) e
testo per la parte narrativa del documento.

Codice

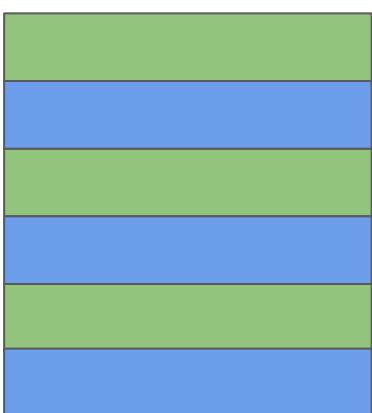


Testo



Copia e incolla manuale tra codice
e documento per integrare risultati

Codice + Testo



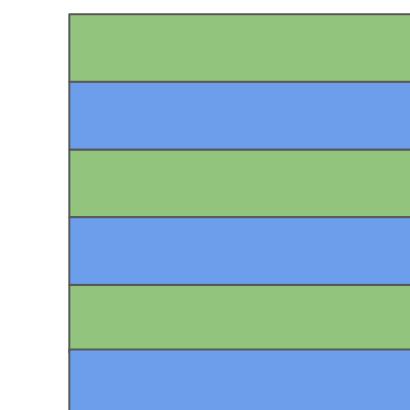
Nessuno step intermedio, stesso
documento

- Non intuitivo
- + Errori meno probabili
- Completamente riproducibile

Paper



Paper





Strumenti

Quarto → documenti (slide, paper, etc.) integrando Markdown e Python, R, etc.

R Markdown (predecessore) → documenti integrando Markdown ed R

Jupyter Notebook → abbastanza usato prevalentemente in Python

...

<https://github.com/shared-research/simulating-meta-analysis/blob/main/documents/paper/paper.Rmd>
<https://github.com/shared-research/simulating-meta-analysis/blob/main/documents/output/paper.pdf>



Tool per rilevare possibili errori



statcheck on the web

To check a PDF, DOCX or HTML file for errors in statistical reporting, upload it below. See the FAQ page for more information about what statcheck can and cannot do.

Upload files (pdf, html, or docx):

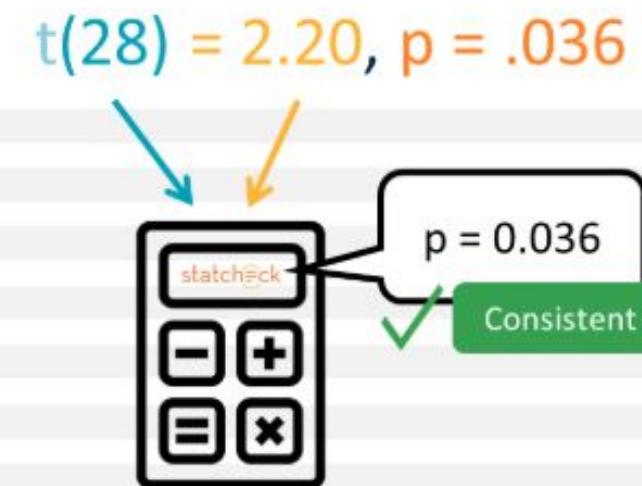
No file selected

Settings:

Try to identify and correct for one-tailed tests

statcheck is a "spellchecker" for statistics. It checks whether your *p*-values match their accompanying test statistic and degrees of freedom.

statcheck searches for null-hypothesis significance test (NHST) in APA style (e.g., $t(28) = 2.2$, $p < .05$). It recalculates the *p*-value using the reported test statistic and degrees of freedom. If the reported and computed *p*-values don't match, statcheck will flag the result as an error.



<https://michelenuijten.shinyapps.io/statcheck-web/>



Step di una ricerca (riproducibile)

- Raccolta dati
- Gestione dati
- Analisi dati
- Reporting
- Condivisione

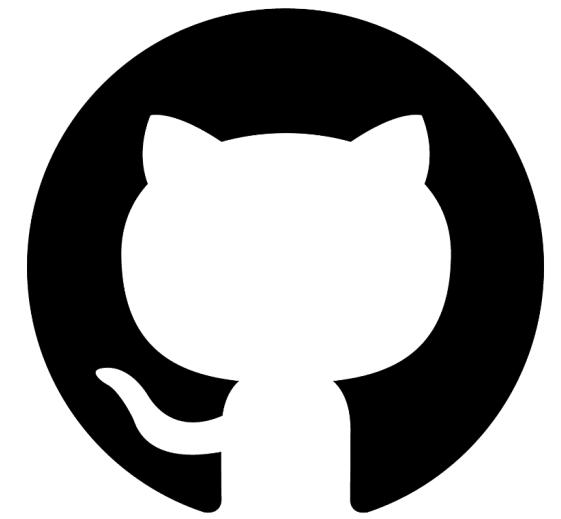


Condivisione materiale (dati, codice, etc.)

A prescindere dallo specifico strumento, ci sono alcune caratteristiche utili:

- Facilmente accessibile (idealmente un link univoco) e utilizzabile
- Future-proof (difficilmente modificabile, corruttibile)
- Con identificativo, magari anche citabile (DOI)
- Integrabile con altri strumenti utilizzati
- Accettato dalle riviste scientifiche

Condivisione codici, Github



Version control locale

- Versioning trasparente delle modifiche
- Ogni versione dei file è tracciata

Version control online

- collaborazione
- condivisione del codice

<https://github.com/shared-research/simulating-meta-analysis>

Condivisione dati, Open Science Framework



OSF HOME ▾

Understanding meta-analysis through d... Metadata Files Wiki Components Analytics Registrations Contributors Add-ons Settings

0.0B Make Private Public ⚡ 0 ...

Understanding meta-analysis through data simulation with applications to power analysis

Contributors: Filippo Gambarota, Gianmarco Altoè

Date created: 2023-02-16 05:46 PM | Last Updated: 2023-09-27 12:12 PM

Identifier: DOI 10.17605/OSF.IO/54DJN

Category: Project

Description:

Meta-analysis is a very important tool to combine evidence from multiple studies. When comes to understanding a new statistical method, data simulation is a powerful educational tool. In this tutorial, we present how to simulate the most common meta-analysis models and meta-regressions.

License: CC-By Attribution 4.0 International

Has supplemental materials for Understanding meta-analysis through data simulation with applications to power analysis on PsyArXiv ▾

Wiki

This repository contains the code to reproduce the analysis of the paper **Understanding meta-analysis through data simulation with applications to power analysis** by Gambarota Filippo and Gianmarco Altoè. All the materials are also available on [Github](#).

Files

Files can now be accessed and managed under the Files tab.

Citation

Components

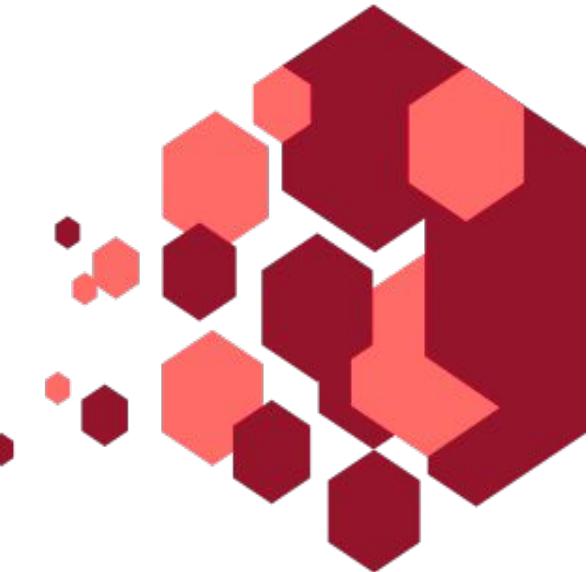
Add Component Link Projects

Add components to organize your project.

Tags

data simulation x meta-analysis x monte carlo simulation x R x Add a tag

<https://osf.io/54djin/>



Integrazione, Open Science Framework

Oltre a condividere file, è possibile collegare OSF a vari altri servizi (tra cui Github) per ottenere un **hub centrale di condivisione con DOI associato**.

Add-ons

All Add-ons Connected Accounts

Sync your projects with external services to help stay connected and organized. Select a category and browse the options. To manage all add-ons connected to your account, visit your profile settings.

Filter add-ons

Additional Storage

Citation Manager

 Bitbucket Connect	 Box Connect	 Dataverse Connect
 Dropbox Connect	 figshare A Digital Science Solution Connect	 Github Configure
 GitLab Connect	 Google Drive Connect	 Onedrive Connect



Crisi di replicabilità

Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015

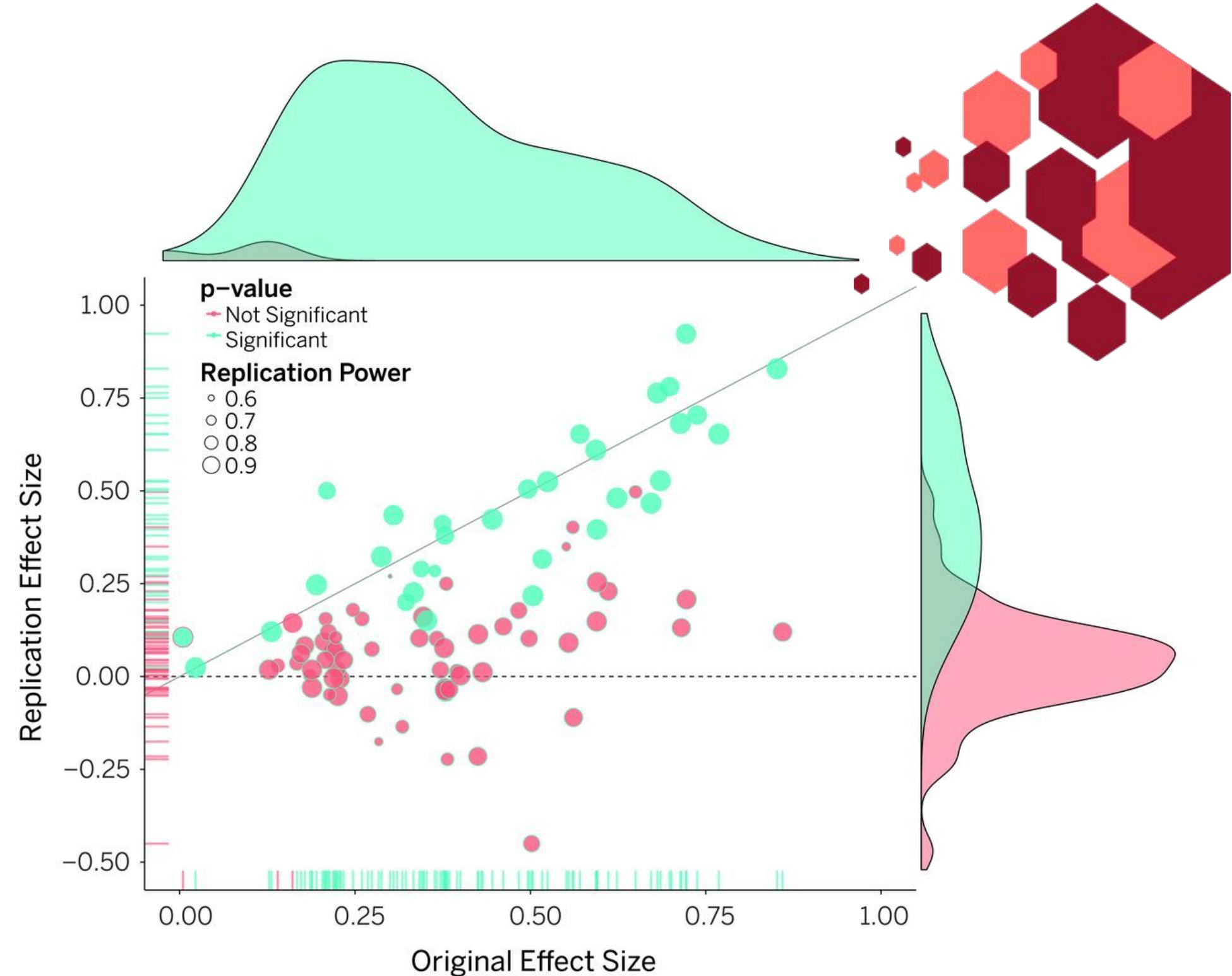
Colin F. Camerer^{1,16}, Anna Dreber^{2,16}, Felix Holzmeister^{10,16}, Teck-Hua Ho^{4,16}, Jürgen Huber^{3,16}, Magnus Johannesson^{11,16}, Michael Kirchler^{3,5,16}, Gideon Nave^{6,16}, Brian A. Nosek^{12,16*}, Thomas Pfeiffer^{13,16}, Adam Altmejd^{14,16}, Nick Buttrick^{7,8}, Taizan Chan¹⁰, Yiling Chen¹¹, Eskil Forsell¹², Anup Gampa^{7,8}, Emma Heikensten², Lily Hummer⁸, Taisuke Imai^{10,13}, Siri Isaksson², Dylan Manfredi⁶, Julia Rose³, Eric-Jan Wagenmakers¹⁴ and Hang Wu¹⁵

RESEARCH ARTICLE SUMMARY

PSYCHOLOGY

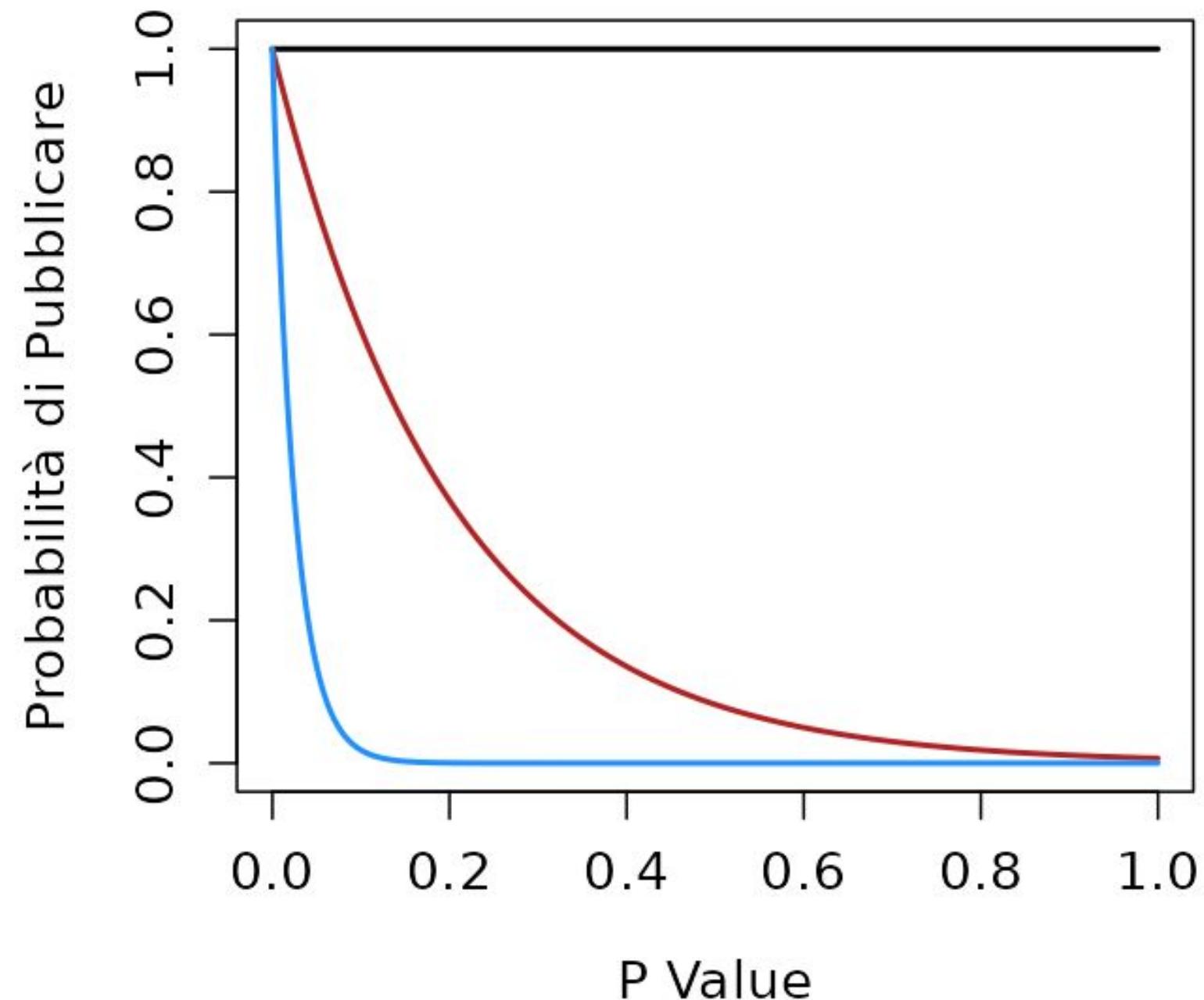
Estimating the reproducibility of psychological science

Open Science Collaboration*



Effetti non replicati o sistematicamente inferiori rispetto a quelli originali

Publication Bias



Se sistematicamente solo una parte dei risultati viene pubblicata, la letteratura non è un campione rappresentativo con effetti, solitamente, inflazionati o dei falsi positivi.

Essay

**Why Most Published Research Findings
Are False**

John P. A. Ioannidis



Questionable Research Practices (QPRs)

Oltre al publication bias, la letteratura può essere *biased* e non replicabile (e riproducibile) per condotte considerate problematiche se non fraudolente:

- Eliminazione e/o manipolazione di dati
- Selezione delle ipotesi di ricerca a posteriori
- Omissione di ipotesi testate o variabili non dichiarate a priori
- ...



Preregistrazione e registered report

La preregistrazione ma in particolare il registered report sono delle possibili soluzioni al problema del publication bias e alle QPRs.

La necessità di pubblicare qualcosa di innovativo, groundbreaking etc. forza a perseguire queste pratiche.

Una soluzione è quindi rendere il risultato e la pubblicazione più indipendenti possibili e valutare altri step della ricerca.



Preregistrazione

L'idea è quella di dichiarare le ipotesi, metodo, raccolta dati, etc. prima di fare l'esperimento o la ricerca.

Un documento viene registrato in modo irreversibile su un portale (ad esempio OSF).

Questo può essere visionato, citato e riportato quando poi viene proposto il lavoro ad una rivista.



Preregistrazione, pro e contro

- + Non è un formato rigido, è abbastanza flessibile sia nel modo di formalizzare che nella quantità di dettagli
- + Permette di avere un riferimento (permalink) indipendente e non modificabile
- + Non è vincolante
- Non da garanzie di pubblicazione
- Seppur non modificabile, non vincola a deviare o completamente ignorare ciò che è stato prestabilito
- Troppa flessibilità



Esempio di preregistrazione

OSF REGISTRIES ▾ Add New My Registrations Help Donate Join Login

Auditory selective attention under working memory load

Public registration ▾ Updates ▾

Overview Metadata Files Resources Wiki

Study Information

Title Auditory selective attention under working memory load

Authors

Description There is currently evidence for both increased and reduced distractor interference.

Contributors Rena Bayramova, Enrico Toffalini, Mario Bonato, and Massimo Grassi

Registration type OSF Preregistration

Date registered October 16, 2019

<https://osf.io/pnq6m>



Esempio di preregistrazione

- Come è formato e reclutato il campione
- Come viene **creato l'esperimento, condizioni, randomizzazione** etc.
- La *sample size justification* quindi quanti partecipanti vengono reclutati
- Il campione viene determinato tramite una procedura di **simulazione di dati in vari scenari**
- Quali analisi, variabili considerate, **esclusione** di partecipanti, etc.

In altri termini tutti gli elementi che se non specificati o scelti a posteriori possono rendere i risultati meno replicabili.

Registered Report (RR)



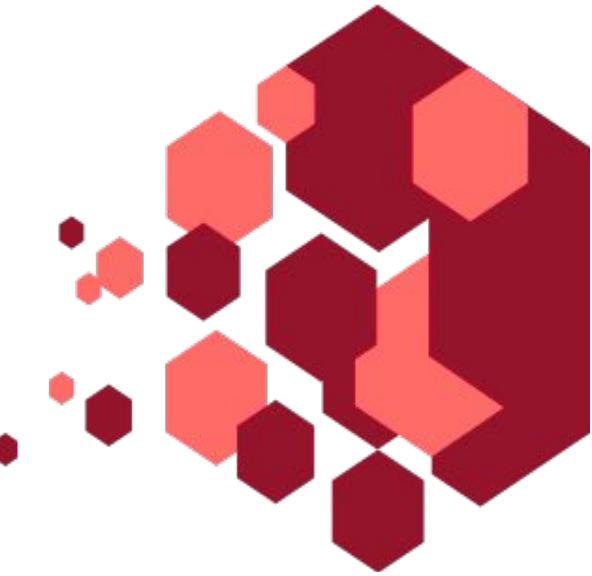


Registered Report (RR)

Ci sono dei template (e.g., <https://osf.io/93zhn>) e ogni rivista può avere il suo formato ma l'idea è simile alla preregistrazione MA il protocollo di pre registrazione viene sottoposto a *peer-review*.

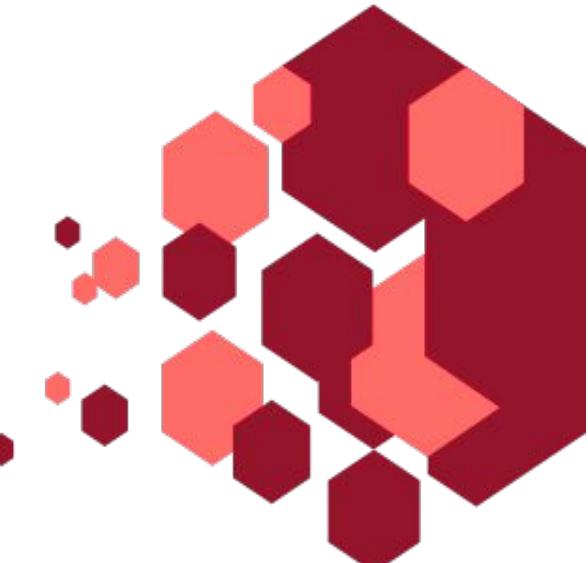
In caso di accettazione (generalmente) il lavoro sarà pubblicato a prescindere dai risultati.

L'incentivo è di produrre la metodologia più adeguata e dettagliata (con i feedback della peer-review) invece che nel produrre un risultato significativo.



Registered Report (RR) Pro e contro

- + Idealmente elimina il publication bias
- + Permette di pensare ed aggiustare il metodo prima di raccogliere i dati
- + Sposta lo sforzo dal trovare/confezionare un risultato a pensare all'esperimento ed alle ipotesi
- È un formato tendenzialmente molto rigido e le deviazioni dal protocollo devono essere esplicitate e concordate
- Spesso non è semplice considerare tutti gli scenari o problematiche in anticipo
- Non tutte le riviste lo accettano



Open Science Badges

E' comune vedere dei paper pubblicati che riportano uno o più badges riguardo buone pratiche di open science

Allo stesso modo è sempre chiaramente riportato se un paper è un registered report o meno (spesso anche nel titolo)





Multilab Replication Study

REPLICATING THE UNCONSCIOUS WM EFFECT

1

Registered Report (Stage 1 Manuscript)

**Replicating the unconscious working memory effect:
A multisite preregistered study**

Alicia Franco-Martínez¹ , Ricardo Rey-Sáez¹ , David R. Shanks³ , David Soto² , and
Miguel A. Vadillo¹ 

¹Autonomous University of Madrid

²Basque Center on Cognition, Brain and Language

³University College London

I multilab replication studies sono un tipo particolare di lavoro (nella maggior parte dei casi con registered report) che direttamente testano la replicabilità di un certo effetto, esperimento etc.

<https://osf.io/tkyfh>

Methodological Review Board (MRB)



A partire da un'iniziativa dedicata all'Open Science condotta da Psicostat per il Dipartimento di Psicologia dello sviluppo e della Socializzazione e finanziata dall'Ateneo, stiamo organizzando il **primo MRB in psicologia in Italia**

Prendiamo ispirazione dal modello introdotto dal Prof. Daniel Lakens (Eindhoven University of Technology)

L'idea è di:

- valutare i progetti di ricerca prima della raccolta dati
- diffusione di linee guida riguardo l'uso di open data e open code



Methodological Review Board (MRB)



- Partecipazione volontaria
- 3 reviewers per progetto (statistica, psicometria e psicologia)
- Eventuale indirizzamento ad altri esperti per situazioni specifiche

Portata a termine la pre-registrazione, si riceve un **badge MRB**, garante del fatto che lo studio soddisfa standard di chiarezza e trasparenza.





Replicability crisis in Science 9-11 Giugno 2025

Methodological School riguardo la crisi di replicabilità nella scienza organizzata dal Dipartimento di Scienze Statistiche e dal corso di Dottorato in Psychological Sciences.

Con speaker internazionali (e nazionali)

Prof. Giovanni Parmigiani (Harvard University)

Prof. Branden Fitelson (Northeastern University)

📍 Scuola di Psicologia, Università di Padova

⌚ Scadenza iscrizioni 5 Maggio 2025

Per iscrizioni <https://psicostat.github.io/replicability-psy-unipd-phd/>

Per qualsiasi informazione filippo.gambarota@unipd.it

ARCHIVI UNIPD



Per i contributi della ricerca:

Padua Research Archive (IRIS)

<https://www.research.unipd.it/>



Per i dati sottesi alla ricerca:

Research Data UNIPD

<https://researchdata.cab.unipd.it/>





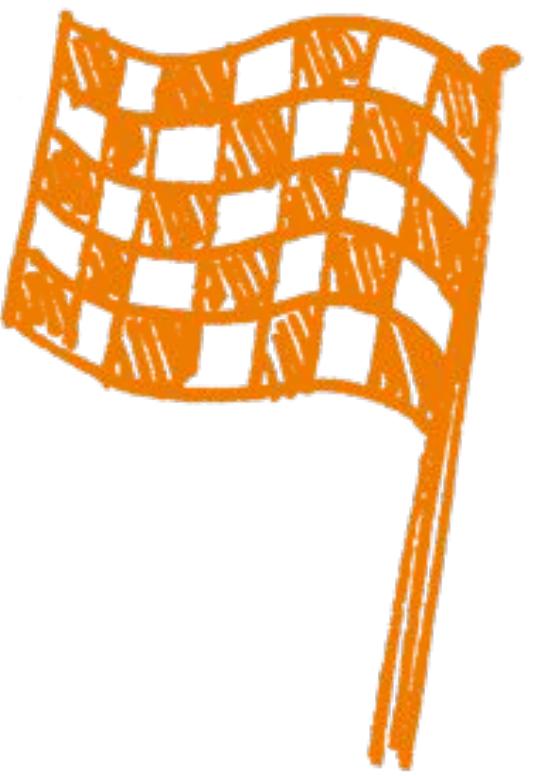
Servizio Aiuto

Il Servizio Aiuto fornito dal Sistema Bibliotecario di Ateneo dell'Università di Padova è un servizio di supporto online rivolto a tutta l'utenza, sia istituzionale che esterna.

Code specifiche sui temi dell'Open Science:

- Supporto alla pubblicazione accademica
- Consulenza tematiche diritto d'autore
- Supporto Open Science (Open Access, Open Data)

<https://biblio.unipd.it/aiuto>





UNIVERSITÀ
DEGLI STUDI
DI PADOVA

SBA

SISTEMA BIBLIOTECARIO
DI ATENEO



Grazie per l'attenzione!

Università degli Studi di Padova
Ufficio Biblioteca Digitale - Centro di Ateneo per le Biblioteche

Sistema Bibliotecario di Ateneo



This work is licensed under the Creative Commons license
[Attribution-ShareAlike 4.0 International](#)



Le immagini sono tratte dal sito
<https://open-science-training-handbook.gitbook.io/book>

