Practices to improve and guard research integrity

Presenter: Zoltan Kekecs, PhD

ELTE, Institute of Psychology

kekecs.zoltan@ppk.elte.hu

What is trustworthy research?

- Trust that the researchers did everything in their power to address the research question with high quality and rigor
- The research methods used are adequate
- Limitations of the research project that limit the generalizability of the findings are transparent

- Is the effect reliably reproducible

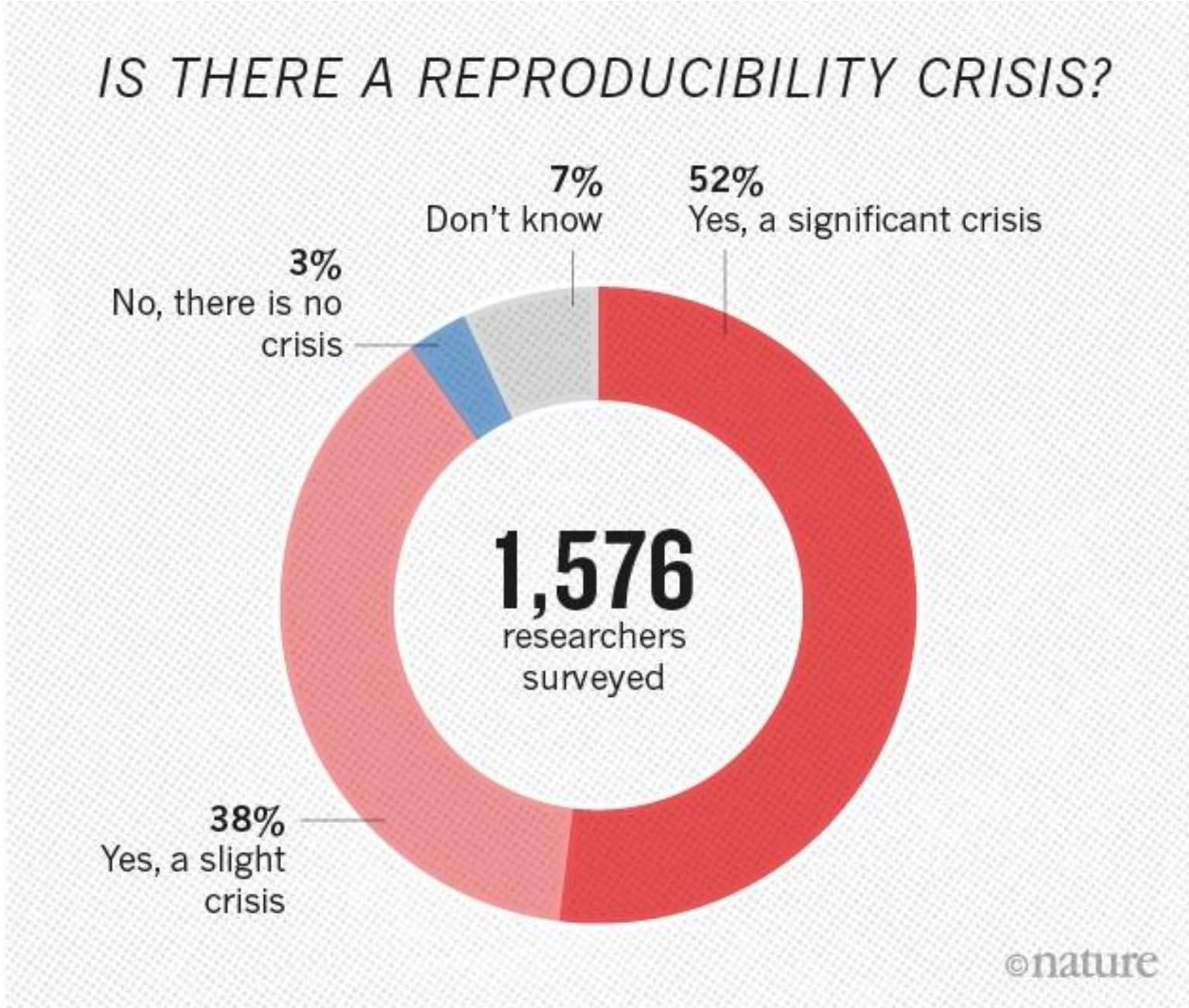


Image from: https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

Reproducibility of scientific findings

- Reproducibility of findings published in top journals in
  - Psychological Science - **47%** (Owens, 2019; Open Science Collaboration, 2015; Camerer et al., 2018)
  - Economics – **61%** (Camerer et al., 2016)
  - Preclinical cancer biology - **40%** (Errington et al., 2021)

Journal	% Findings Replicated
Journal of Personality and Social Psychology: Social	23
Journal of Experimental Psychology: Learning, Memory, and Cognition	48
Psychological Science, social articles	29
Psychological Science, cognitive articles	53
Overall	36

Table from: Diener & Biswas-Diener (2016)

61-86% accuracy in predicting replicability of an effect. (Cramer et al 2016, 2018)

Replication to the rescue

- Every finding should be replicated.

Issues with replication

- Time consuming
- Costly
- Cannot replicate everything ourselves
- If done by others replications have the same trust-issues as original studies
- Post-hoc criticism

- We need tools to be able to **create trustworthy original studies** that are acceptable for the stakeholders instead of relying heavily on replications

Producing robust and trustworthy research

How to produce trustworthy replicable research

- Make sure that the **methodological approach is appropriate**
- Being able to **demonstrate the integrity** of research steps on demand

Research process

Planning


Plans
Experiment
Data management
Data analysis
Interpretation

Get feedback about research plans

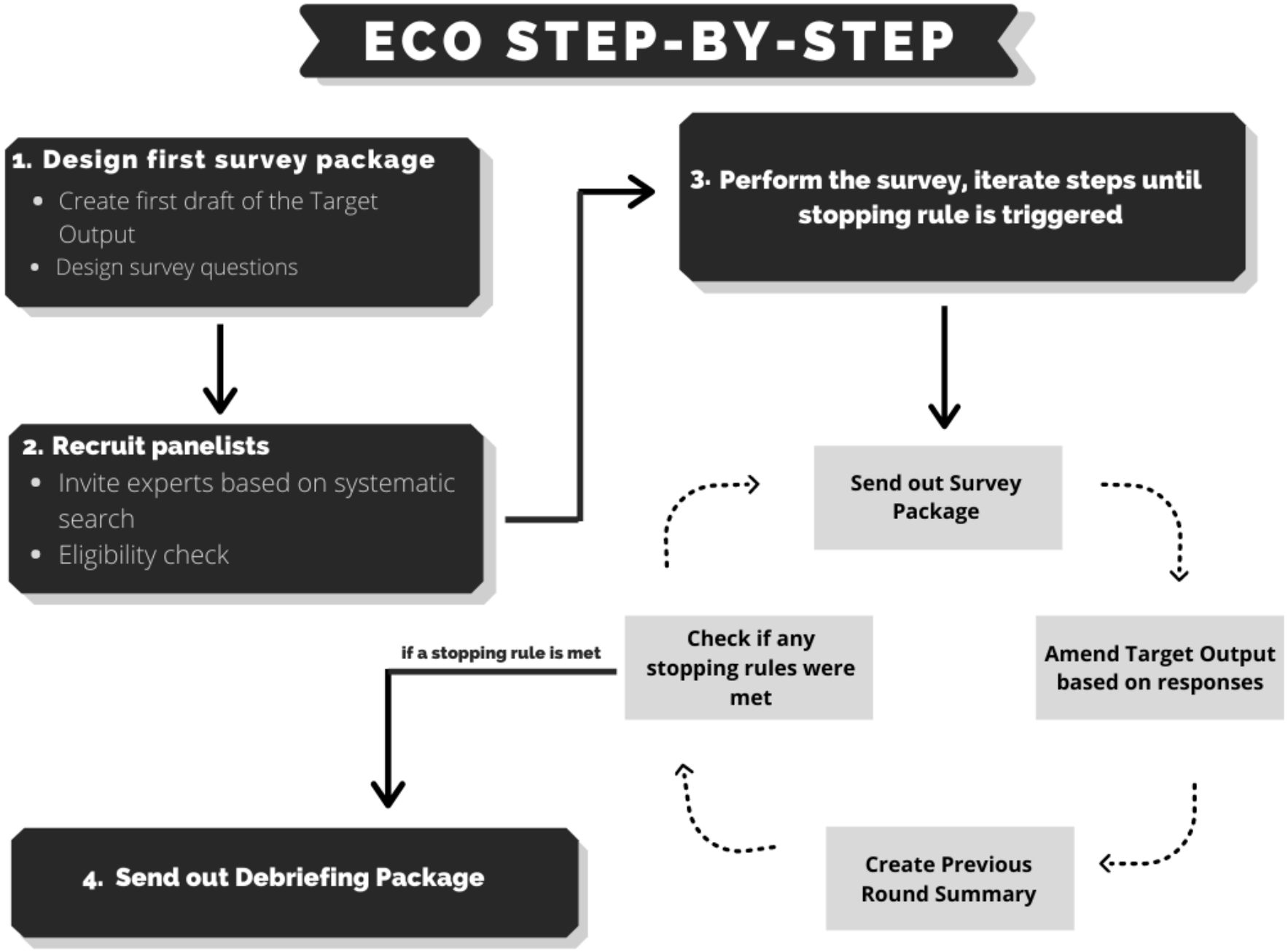
- Present research plans to colleagues
- Publish research plans
  - Journals like BMC Trials
- Registered reports

Adversarial or collaborative study design

- Adversarial collaboration
- Expert Consensus design (ECO)

The purpose of ECO

- ECO aims to facilitate consensus among a panel of experts about a target scientific output
  - to decrease the chance that the target output will be subject to **conceptual or methodological mistakes**, and
  - to increase the chance that the target output would be **acceptable by the stakeholders** on the field



Kekecs, Z., Szaszi, B., & Aczel, B. (2020, September 25). Expert Consensus Procedure (ECO): Facilitating Robust Scientific Outputs. https://doi.org/10.31234/osf.io/9gqru

Adversarial or collaborative study design

- Adversarial collaboration
- Expert Consensus design (ECO)

- Co-design mutually acceptable **conclusions**

Test the procedure and be transparent

- Pilot study
  - To detect unforeseen events
  - To fine tune sampling and experimental procedures

- Preregistration
  - To make the final research **plan transparent**, including aims, hypotheses, data collection procedures, data management plans, data analysis plan, and planned conclusions.
  - Use dedicated registry, like **euclinicaltrials.eu, clinicaltrials.gov, OSF**
  - Including **analysis code**
  - Run **power analysis** using the preregistered analysis code
  - Be transparent about the **operational characteristics** of the study

Protocol execution

Ensure and demonstrate protocol fidelity

- Manual for experimenters
- Checklists
- Training
  - This training needs to be verifiable
- Laboratory logs
  - Automated logs and manual notes about the research sessions
- Tamper-evident software
  - Experiment software is run from a version controlled repository (e.g. GitLab)
  - Can demonstrate integrity of the experimental software through the study lifecycle
- External research audit
  - Dedicated people checking research integrity

Data management

Demonstrating integrity of the data management process

- Data management plan
- Direct Data Deposition
  - Data is directly saved in real time to a trusted third party **data repository** with version control
  - This way the **integrity** of research data can always be demonstrated
  - Guide to DDD
- Born-open data
  - Data is made accessible in real time as it is being collected
  - Radical trasnparency about data
  - Allows immediate reusability
- Real-time Research Report
  - Data is analyzed in real-time and offered for interpretation via graphs or tables (e.g. via shiny apps)
  - Allows easy interpretability on the fly

Data analysis and interpretation

Trustworthy data analysis

- Reproducible data analysis
  - Open code (e.g. GitHub)
  - Open data (e.g. OSF, Zenodo)

- Trust in findings
  - Preregistration
    - Analysis code
    - Hypotheses
  - Real-time Research Report
  - Give access to raw data
  - Direct data deposition

Trustworthy data analysis

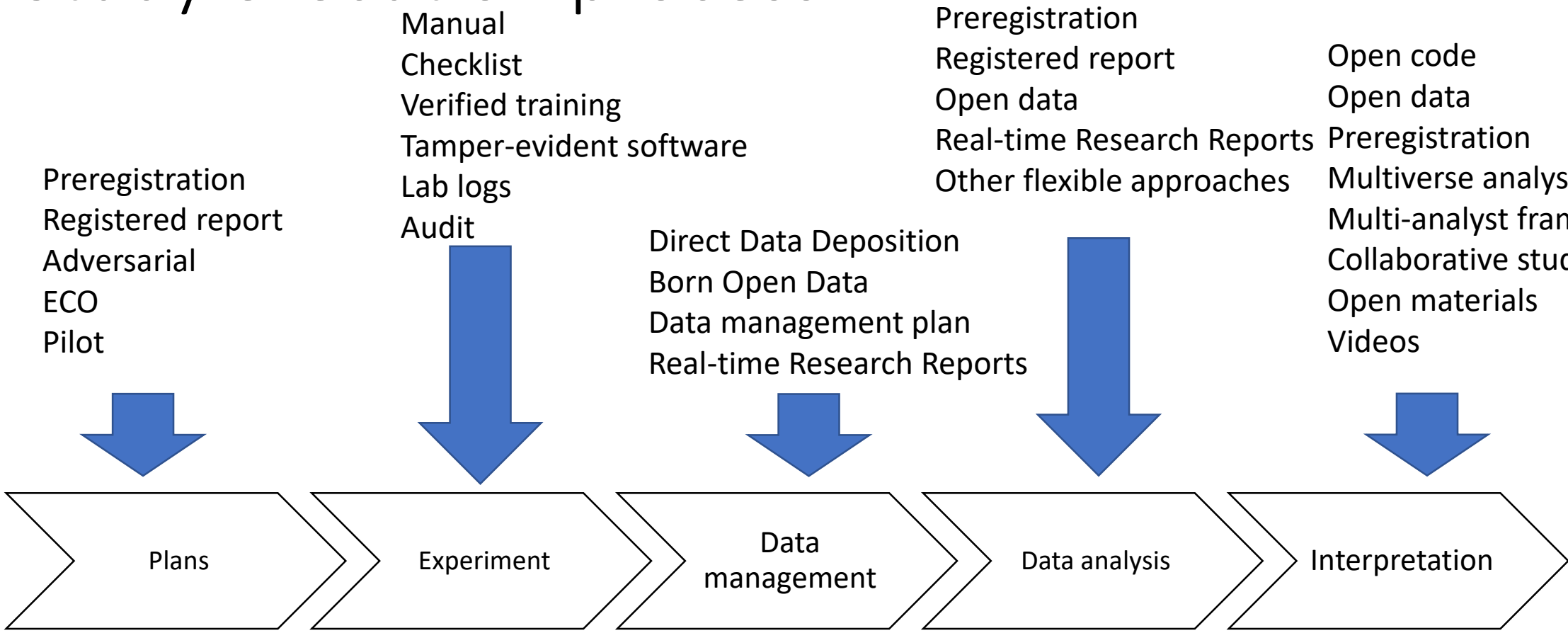
- When flexibility is needed
  - Training-set – Test-set approach
  - Blinded data analysis
  - Multi-analyst framework

- When data is unsharable
  - Remote data analysis on deposited data

Transparency and reproducilibity of the interpretation

- Reproducible interpretation
  - **Preregistered** hypotheses, analysis code, **stopping rules, and conclusions**
  - All of these **unchanged**
- Making the procedures reproducible
  - Open materials (including manual, and data collection software)
  - Videos documenting research sessions could help to capture contextual details that may not seem important to researchers but may influence results
- Robustness of the interpretations
  - Multiverse analysis
  - Multi-analyst framework
  - Collaborative study design with agreed conclusion

Study execution process



Percieved usefulness of credibility enhancing tools

- Laboratory logs
- Manual for experimenters
- Checklist for experimenters
- Preregistration
- Open materials
- Consensus Design (ECO)
- Direct Data Deposition
- Born open data
- Real-time research report
- Verifiable training
- External research audit
- Tamper-evident software

Useful tools for big-team-science

- Consensus Design – helps to agree in things
- Manual, Checklists, and Video-verified training, Centralized tamper-evident software – standardize procedures, high fidelity across sites
- Laboratory logs – Keep track of progress and problems
- Direct Data Deposition – No data-transfer issues, seamless open data
- Real-time research report – Easy and real-time check on progress

Acknowledgement

FUNDAÇÃO Bial

Institution of public utility

Acknowledgement

- Kekecs, Z., Palfi, B., Szaszi, B., Szecsi, P., Zrubka, M., Kovacs, M., ... & Aczel, B. (2023). Raising the value of research studies in psychological science by increasing the credibility of research reports: the transparent psi project. *Royal Society Open Science*, *10*(2), 191375.
- Kekecs, Z., Szaszi, B., & Aczel, B. (2020, September 25). Expert Consensus Procedure (ECO): Facilitating Robust Scientific Outputs. https://doi.org/10.31234/osf.io/9gqru

Contact

kekecs.zoltan@ppk.elte.hu

Twitter: @kekecs_zoltan