Open Science and Reproducibility

Venue: Mogens Zielerstuen, Fredrik Nielsens Vej 4, 8000 Aarhus C

Tuesday 12th March

09:00	**Christine Parsons**	Introduction
Chair:	**Joshua Skewes**	**Context session**
09:15	**Ivan Flis**	A radical historical reading of psychology's replication crisis
10:15	**Berna Devezer**	Toward a Theory of Scientific Discovery and Reproducibility
11:15	Coffee	
11:30	**Danielle Navarro**	Science, statistics and the problem of "pretty good inference"
12:30	Panel discussion	Joshua, Ivan, Berna, Danielle
13:00	Lunch	
Chair:	**Andreas Roepstorff**	**Tools 1 session - software**
14:00	**Joachim Vandekerckhove**	Metastudies for robust tests of theory
15:00	**Britta Westner**	Open software in open science
16:00	Coffee	
16:15	Panel discussion	Andreas, Joachim, Britta

Wednesday 13th March

Chair:	**Micah Allen**	**Tools 2 session - practices**
09:30	**Lisa DeBruine**	Everything is cool when you're part of a team
10:30	**Zoltan Dienes**	The inner workings of Registered Reports
11:30	Panel discussion	Micah, Zoltan, Lisa
12:00	Lunch	
Chair:	**Riccardo Fusaroli**	**Populations session**
13:00	**Eiko Fried**	Theory and measurement crises as obstacles to replicability in Clinical Psychology
14:00	**Christina Bergmann**	Can developmental science provide some practical solutions to improve transparency?
15:00	Coffee	
15:15	**Bret Beheim**	Implementing Open Science Principles in Longitudinal Field Data Collection
16:15	Panel discussion	Riccardo, Eiko, Christina, Brett
16:45	Summing up	

The workshop is organized by Christine Parsons, Joshua Skewes & Andreas Roepstorff, IMC, AU

Ivan Flis: A radical historical reading of psychology's replication crisis

In this paper I will discuss some historical antecedents to the replication/replicability crisis in psychology. Mainstream scientific psychology as we know it today went through its disciplinary formation and standardization of research practices during the late nineteenth and twentieth century. Especially important was the middle of the twentieth century, when the discipline was thoroughly Americanized, professionalized, and exhibited an accelerated expansion in the number of scientists/professionals and their research outputs. Here, I will highlight three historical topics of epistemological relevance: operationism, literature expansion, and the move toward thorough quantification of the subject matter and concurrent spread of inferential statistics (the so-called 'inferential revolution'). These three topics will be used as a means for opening up a critical debate about the current crisis of the dominant research programs in psychology; and the role Open Science practices and advocacy play in it. The 'radical' in the paper's title is meant in two ways: (1) radical for academic historians of science and historians of psychology, because I wish to engage in normative evaluations of psychological knowledge production; and consequently (2) radical for research psychologists, because I am arguing that psychologists should start working on more fundamental epistemological questions within their areas of research, not just methodological innovation. Although I will focus on psychology, I will try to draw conclusions about Open Science as a transformative movement in general.

Berna Devezer: Toward a Theory of Scientific Discovery and Reproducibility

Scientists have discussed the role of replication and reproducibility in scientific progress for two millennia. However, these discussions have not yielded a theoretical understanding of the role of reproducibility in scientific process. Recent literature on the reproducibility crisis identifies several putative causes for the proliferation of irreproducible results, most of which are methodological. Without a theory of scientific discovery that delineates the role of reproducibility, whether these putative causes can solely account for the proliferation of irreproducible results is not clear. Drawing from an historically informed conception of science that is open and collaborative, we identify the components of an idealized experiment and analyze these components as a precursor to develop such a theory. We show that there are some impediments to obtaining reproducible results that precede many of the causes of irreproducibility often cited in literature on the reproducibility crisis. Even in the absence of methodological misdeeds at the individual level, reproducibility may still not be guaranteed due to behavioral aspects of science or structural characteristics of scientific phenomena.

Danielle Navarro: Science, statistics and the problem of "pretty good inference"

A central problem facing scientists is choosing the most appropriate explanation of some observed phenomenon. In statistics, we face an analogous problem of selecting the model that provides the "best" account of a data set. The two problems have much in common, but in this talk I'll argue that in practice they are not the same. Part of the scientific problem we face - particularly in psychology - is that all of our theories (both formal and informal) are wrong, and usually quite badly wrong, yet we still need to make decisions about which of these (bad) theories is the most useful one to guide our future work. In statistics, the analogous problem is one of model misspecification - we cannot select the "true" model because in all likelihood no such thing exists, and even if it did it most certainly does not exist among the models under consideration. This leaves us facing the problem of "pretty good inference" - of trying to make inferences that will guide us toward sensible actions despite our ignorance of the world. In this talk I do not propose any strong "solutions" to this problem, but will aim to highlight how this perspective creates a certain tension between what we hope to achieve (learning about the world) and the tools we usually rely on - as open scientists - to do so.

Joachim Vandekerckhove: <u>Metastudies for robust tests of theory</u>

We describe and demonstrate an empirical strategy useful for discovering and replicating empirical effects in psychological science. The method involves the design of a meta-study, in which many independent experimental variables—that may be moderators of an empirical effect—are indiscriminately randomized. Radical randomization yields rich data sets that can be used to test the robustness of an empirical claim to some of the vagaries and idiosyncrasies of experimental protocols and enhances the generalizability of these claims. The strategy is made feasible by advances in hierarchical Bayesian modeling which allow for the pooling of information across unlike experiments and designs, and is proposed here as a gold standard for replication research and exploratory research. The practical feasibility of the strategy is demonstrated with a replication of a study on subliminal priming. All materials and data are freely available online via https://osf.io/u2vwa/.

Britta Westner: Open software in open science

Open software, spanning from openly shared code to open source toolboxes, is one of the pillars of open science and essential for the reproducibility of data analyses. In the first part of this talk, I will argue for the sharing of data analysis code and further look into the details of how this can be done in an organized and easily accessible way. The second part of the talk is focused on open source toolboxes. I will emphasize their influence on (open) science and illustrate some of their struggles. Lastly, I will encourage contributing to open source toolboxes by showing some common steps in that process and highlighting the positive effects for the science community.

Lisa Debruine: Everything is cool when you're part of a team:The role of large-scale collaboration in improving replicability and generalisability

The "replication crisis" has led to a call for initiatives to increase the replicability of psychological science, such as data and code sharing, pre-registration, registered reports, and reproducible workflows. Similarly, researchers have questioned the extent to which studies of WEIRD populations (Western, Educated, Industrialised, Rich, and Democratic) generalise to the majority of people in the rest of the world. Here, I will discuss how large-scale collaborations can improve both replicability and generalisability, with a focus on the Psychological Science Accelerator, a globally distributed network of more than 360 laboratories from 45 countries across all six populated continents.

Materials available: https://osf.io/be3gy/

Zoltan Dienes: The inner workings of Registered Reports

I discuss what features a Registered Report has to have in order to solve (some) existing problems in our scientific procedures. These features illustrate the most common reasons a submission to RRs at Cortex are desk rejected. Writing a Registered Report does not come naturally for many people it seems, but I hope to guide people in this process. The guidelines will also make pre-registration (i.e. as done by authors when not submitting to a Registered Reports format) more scientifically valuable as well.

Eiko Fried: Theory and measurement crises as obstacles to replicability in Clinical Psychology

Open science advocates in Clinical Psychology would likely agree that the adoption of open science practices and implementation of replicability efforts in the field have been scarce, with some notable exceptions. In this talk, I discuss how a crisis of theory and a crisis of measurement provide fundamental obstacles to replicability, and sketch some ways forward. Regarding the theory crisis, Clinical Psychology is full of great ideas, but lacks formalized theories that can be falsified. I provide an example of a formalized computational model for panic disorder to exemplify advantages of such theories. In the second part, I discuss the lack of attention to measurement practices in many areas of Clinical Psychology, which leads to problems with scientific inference. Together, theory and measurement crises impede progress towards meaningful replications, and adopting open science practices will help overall these challenges.

Christina Bergmann: Can developmental science provide some practical solutions to improve transparency?

Developmental scientists face unique hurdles to reproducibility - in addition to those highlighted for hypothesis-testing experimental research more generally: We work with an (often) uncooperative population which is difficult to test and recruit, which in turn leads to noisy measures and small sample sizes. To make matters more intransparent, much of the raw data (video and audio recordings in particular, but also many questionnaire responses) are sensitive and thus cannot easily be shared.

At the same time, and possibly because of those challenges, developmental scientists have (partially) embraced transparency (see e.g. childes.talkbank.org for decades worth of open data, or more recently wordbank.stanford.edu, lookit.mit.edu, and manybabies.stanford.edu).

A consequence are unique solutions to our hurdles, which may also be of interest to other experimental researchers. One example among those I will discuss are "walkthrough videos" to provide detailed documentation of the procedure in the absence of sharing actual recordings of experiments. Such videos can be highly valuable, for example for teaching and to uncover systematic methodological variation.

Bret Beheim: Implementing Open Science Principles in Longitudinal Field Data Collection

Field data collection is a critical weak point in the task of making scientific analysis more transparent and reproducible. In addition to the standard problems facing an experimental dataset in controlled conditions, the open science field practitioner must contend with apocyphal data origins, unaccounted-for protocol variation, mysterious revisions, missing metadata, and structural inconsistencies. Field data collection also presents unique ethical and logistical challenges for implementing new methods among large, diverse teams. Here I outline some of the specific solutions we are implementing in our department, focusing on several longitidunal field projects currently running around the world.