(*Last revised on October 1, 2018*)

Understanding Replication Across the Sciences

Rotman Institute of Philosophy Annual Conference
October 12 - 14, 2018
London, Ontario
#URAS2018

Friday, October 12

9:00 a.m. - 9:30 a.m. Coffee

9:30 a.m. - 9:45 a.m. Welcoming - Introduction

9:45 a.m. - 11:00 a.m. HANS RADDER (Vrije Universiteit Amsterdam)
"Reproduction and reproducibility in experimental and observational research"

The paper will address three issues. First, it will briefly situate the topic of reproducibility in the (history of the) philosophy of science. Next, it will explain and develop my own account of experimental and observational reproducibility, starting from the questions of *what* should be reproduced and *who* should be capable to perform the reproduction. Finally, I will apply this account to analyze and assess central aspects of the recent debate on the failures of reproducibility in some specific disciplines.

11:00 a.m. – 11:15 a.m. Coffee Break

11:15 a.m. – 12:30 p.m. ALLAN FRANKLIN (University of Colorado)
"Is It the Same Experimental Result? Replication in Physics"

One of the interesting issues in the philosophy of experiment is that of the replicability of experimental results. The scientific community enthusiastically endorses the idea that "Replication – the confirmation of results and conclusions from one study obtained independently in another is considered the scientific gold standard." The underlying argument for this is that if an experiment has succeeded in revealing a real phenomenon or accurately measuring a quantity then that success should reappear when the experiment is repeated under the same circumstances or when it is reproduced in a different experiment. There are, however, questions about whether this standard is universally, or even typically, applied. There are also questions concerning what constitutes a successful or failed replication.

In this paper I will discuss two clear examples of successful replications: The discovery of the Higgs boson and the detection of gravitational radiation. Two failed replications will also be presented: early experiments on the Fifth Force, a proposed modification of Newton's Law of Gravity; and attempts to measure G, the universal gravitational constant in Newton's law. More complex episodes in which the success or failure of replication was not clear will also be discussed. These include measurements of physical constants; claims of low-mass electron-positron states; and experiments on the pentaquark, the case of the disappearing particle. The methods used to resolve the issues in these more complex cases will also be discussed.

12:30 p.m. – 1:30 p.m. Lunch

1:45 p.m. – 3:00 p.m. JUTTA SCHICKORE (Indiana University Bloomington)
"Repetition and replication in the life sciences: historical perspectives"

I plan to discuss the ways in which past scientists discussed methodological issues pertaining to experimentation, what methodological strategies they considered most important, and how they justified these strategies. Focusing on controversies in 19th-century agricultural and bacteriological research, I examine what significance the researchers attached to replications, what they meant by "replication," and whether non-reproducibility was perceived as a problem.

3:00 p.m. – 3:15 p.m. Coffee Break

3:15 p.m. – 4:30 p.m. YVES GINGRAS (Université du Québec à Montréal)
"The Social and epistemological constraints on replication"

Recalling the basic work of Harry Collins on replication in the 1970s and 1980s, and on Pierre Bourdieu on the struggle for authority in the scientific field, we will survey the general social constraints that contribute to explain the lack of interest in replicating previous science, as a direct effect of Robert K. Merton's model of the scientific community. Despite the recent upsurge of discourse on the topic of replication within the scientific community, it does not seem that they are ready to take the steps necessary to solve that problem which is an effect of the structural constraints that affect the present dynamic of scientific research. I will also link these discourses to those concerning the «Slow science movement».

4:30 p.m. – 6:00 p.m. Poster Session

Saturday, October 13

9:00 a.m. – 9:30 a.m. Coffee

9:30 a.m. – 10:45 a.m. JACQUELINE SULLIVAN (Western University)
"Coordinated Pluralism and Cumulative Neuroscience"

One positive outcome of the "replication crisis" in science is the recognition that we need an appropriate conceptual framework for (a) evaluating scientific research, (b) determining when and why experiments succeed or fail and (c) increasing the successes and minimizing the failures. In this talk, I put forward one such conceptual framework and use it to critically evaluate several case studies from past and recent neuroscience. I show that in designing and implementing experiments and interpreting experimental results, scientists face a number of competing epistemic desiderata that cannot be satisfied within the context of a single experiment or research study. Rather, satisfying these desiderata requires an unprecedented amount of coordination within individual laboratories and across research groups. I put forward recent work in translational cognitive neuroscience as an example of how such "coordinated pluralism" may work in practice. I end by considering whether coordinated pluralism may offer the right recipe for a cumulative neuroscience.

10:45 a.m. – 12:00 p.m. MICHAEL ANDERSON (Western University)
"The in-principle limits of reproducibility in the cognitive neurosciences"

Most individual regions of the brain are multi-modal and active across many different cognitive tasks. Moreover, two kinds of developmental plasticity ensure that this architecture is continually being remodeled: Hebbian learning that changes the strength of synaptic connections to tune local function, and a neural "search" or "reuse" process that acts to establish the functional partnerships between regions that will support newly acquired abilities. These facts call for a reconsideration of both the mathematical and conceptual tools we bring to bear in our understanding of the brain. This talk discusses some of these tools and how they can be used to capture the brain's complexity. Moreover, I will offer an analysis of what kind and degree of reproducibility we should expect these tools to reveal. The brain is constantly changing at multiple spatial and temporal scales. The tools I present can be used to capture and quantify these changes.

12:00 p.m. – 1:00 p.m. Lunch

1:15 p.m. – 2:30 p.m. MIRIAM SOLOMON (Temple University)
"After Disclosure"

Industry funding of research is the greatest known systematic threat to the objectivity of medical research. This paper clarifies the nature and scope of industry funding bias and attempts to quantify it. It reviews four kinds of remedy for industry bias suggested so far: disclosure, standards and regulation, steps towards independence for all clinical research, and case by case assessments, finding most of them helpful but not sufficient. The paper proposes two possible further interventions to reduce the effects of industry bias: qualitative and quantitative discounting of industry results.

2:30 p.m. – 2:45 p.m. Coffee Break

2:45 p.m. – 4:00 p.m. JACOB STEGENGA (University of Cambridge)
"The Perils of P-Hacking and the Promise of Pre-analysis Plans"

P-hacking involves the manipulation of data to find a statistically significant result. Many claim that p-hacking is a problem in science, especially in the medical and social sciences, while others deny this. The problem with p-hacking is usually articulated from a frequentist perspective. In this presentation we articulate the epistemic peril of p-hacking using Bayesian confirmation theory and model selection theory, which we then draw on to explain the arguments on both sides of the debate. This requires a novel understanding of Bayesianism, since a standard criticism of Bayesian confirmation theory is that it cannot accommodate the influence of biased methods. A methodological device widely used to mitigate the peril of p-hacking is a pre-analysis plan. Some say that following a pre-analysis plan is epistemically meritorious while others deny this, and in practice pre-analysis plans are often violated. We use the formal groundwork developed, to resolve this debate, offering a modest defence of the use of pre-analysis plans. In the longer run our ambition is to use this approach to make sense of scenarios in which scientists depart from pre-analysis plans.

4:00 p.m. – 5:15 p.m. AYELET SHAVIT (Tel-Hai College)
"Scientific replication and academic outreach: can it work together?"

The last five years have unfolded a deep controversy over scientific replication, with substantial and ongoing impacts on the trustworthiness of science and its routine working standards in the biological, bio-medical, bio-psychological and psychological research. As a philosopher of science, an important inquiry in this context is to study the existing meanings of 'replication' and 'reproducibility', in the hope of articulating a more coherent and useful concept of replication. It should be clear enough, broadly applicable enough, and provide specific enough tools to answer two crucial questions for any scientific endeavor: to what degree are the results accurate and how generalizable are they? The first question asks

whether the description of the observed results more accurately fit a description of a causal, a random or some other sequence of events? Further, even if the results are "statistically significant" (i.e. non-random), are they reproducible with repeated measurements? The first question in effect deals with the realism of a scientific description and analysis, i.e. what "actually" happened in a certain place and time, while the second question deals with the generality of that analysis and its predictions: to what degree is a result observed for a given population—whether random or not—representative of other populations? These two questions are not the same, yet both address the challenge of replication, and both are concurrently required yet typically conflict in answering research questions that use local knowledge of heterogenous populations. In this paper I will unfold the various dimensions of this accuracy-generalizability conflict within the concept of replication in the biological sciences, and sketch a possible resolution, since a general solution is currently not in sight.

Sunday, October 14

9:00 a.m. – 9:30 a.m. Coffee

9:30 a.m. – 10:45 a.m. RICHARD SHIFFRIN (Indiana University Bloomington)
"Scientific progress despite irreproducibility: A seeming paradox"

It appears paradoxical that science is producing outstanding new results and theories at a rapid rate at the same time that researchers are identifying serious problems in the practice of science that cause many reports to be irreproducible and invalid. Certainly, the practice of science needs to be improved and scientists are now pursuing this goal. However, in this perspective I argue that this seeming paradox is not new, has always been part of the way science works, and likely will remain so. I first introduce the paradox. I then review a wide range of challenges that appear to make scientific success difficult. Next, I describe the factors that make science work-in the past, present, and presumably also in the future. I then suggest that remedies for the present practice of science need to be applied selectively so as not to slow progress, and illustrate with a few examples.

10:45 a.m. – 12:00 p.m. STUART FIRESTEIN (Columbia University)
"Replicate That! Important Failures in Science"

Replication failures are not a crisis in science, as the popular press would have you believe. Rather in most cases they should be viewed as part of the normal process of

scientific discovery - a process that depends on failure to advance. A crucial distinction must be made between a replication (positive or negative) and a *replicable* experiment. Only the second is a demarcating factor in science. The first can often lead to new knowledge and is to be expected if not welcomed. Failures are crucial to discovery, replication failures are one specific kind of scientific failure.

12:00 p.m. – 1:00 p.m. Lunch

1:00 p.m. – 2:15 p.m. LORNE CAMPBELL (Western University)

"Teaching Good, Open Science by Conducting Close Replications in the Classroom"

The past few years has witnessed much debate regarding research practices that can potentially undermine the accuracy of reported research findings. A strong case can be made that, for example, that the Type I error rate is in fact much higher than the nominal a level of .05 because of study design and data analytic flexibility. Combined with the typically low levels of statistical power (~50%) in the published research, and the fact that over 90% of published findings are reported as statistically significant, a non-trivial number of published research findings are false-positives. But which ones? The most effective way to reliably distinguish true positives is to conduct high quality close replications of published research in order to derive more precise effect size estimates for presumed effects. In this presentation I will discuss how such a hands-on replication approach to teaching replication and open science practices is an excellent opportunity to teach how to do good science at an early stage in training new scientists. The goal is to make these teaching resources openly available to encourage widespread adoption as well as solicit feedback for improvement.