Society for Philosophy of Science in Practice (SPSP)
11th Biennial Conference
Cambridge, UK
15–18 July 2026

Book of Abstracts

Table of contents

pp. 2–3 **Abstracts of Plenary Lectures**

pp. 4–28 **Abstracts of Posters**

pp. 29–134 **Abstracts of Contributed Papers**

pp. 135-276 **Symposium descriptions (including abstracts of individual talks)**

Abstracts of Plenary Lectures

(in order of appearance)

The Fine Print of Experimental Inferences in Psychology and Beyond

Uljana Feest
Leibniz Universität Hannover
feest@philos.uni-hannover.de

When we take a medication or sign a contract, we are often presented with a list of provisos, the so-called "fine print." These provisos specify the conditions under which the medication is expected to work or the contract is legally binding. In this talk, I argue that experimental inferences likewise depend on a form of fine print: a set of conditions under which experimental data warrant the intended conclusion. Experimental fine print resembles the fine print of contracts and package leaflets in one important respect: People often do not read it carefully. It differs, however, in two crucial ways. First, whereas the conditions stated in contracts or package leaflets are typically assumed to hold, the conditions underlying experimental inferences are often uncertain. Second, this uncertainty concerns the very individuation conditions of the object under investigation. When experimental findings are moderated by unanticipated background conditions, this may either reveal a confounder or indicate that the object itself manifests differently across contexts. I will clarify this distinction and its implications for experimental inferences in psychology and in other fields that study heterogeneous objects.

How Scientific Practice Unlocks the History of Science: The Case of the Personal Equation

Matthew D. Lund
Rowan University
Lund@rowan.edu

The degree to which scientific observations can vary from individual to individual has been a significant concern in philosophy of science since the middle of the 20th century. However, most discussions of the epistemically deleterious effects of theory-laden observation have been situated in debates about theory-testing or scientific change rather than focused on scientific practice. This talk shows how the recovery of historical scientific practices can provide insight not only into how disputes about the comparative accuracy of observations emerged and were resolved, but can also correct received historical narratives.
In 1796, David Kinnebrook, assistant to Astronomer Royal Nevil Maskelyne, was dismissed for marking stellar transits too slowly. Several decades later (in 1823), F.W. Bessel conducted comparative observational studies that confirmed the existence of individual differences that were five to six times greater than was expected, eventually giving rise to the notion of the 'personal equation', an important corrective to individual observational differences that allowed for astronomical observing on an 'industrial' scale. A key part of Bessel's argument concerned his assumptions about Kinnebrook's observational practice. This study uses the Royal Observatory's raw observational data, along with Kinnebrook's private letters (recovered in 1985), to reconstruct Kinnebrook's practice. Since the consequences of practice can often be independently confirmed, practice can ameliorate

historical narratives, like the one surrounding Kinnebrook. Reappraisal of Kinnebrook's defenses of his observational practice discloses that he took important steps toward the development of the personal equation decades before its recognized discovery.

PSP, Inquiry and Science Education

Hasok Chang
University of Cambridge
hc372@cam.ac.uk

There has been a curious lack of connection between science education and the philosophy of science (or history and philosophy of science, HPS). Despite great efforts by Michael Matthews, Douglas Allchin and others HPS remains on the fringes of science education, and science education is not a major concern in mainstream philosophy, HPS, or even science studies more broadly. Yet philosophy of science has clear relevance in science education, for example in the teaching of the "nature of science" (NoS). Particularly for the philosophy of science in practice (PSP), this lack of engagement with education is tragic, since education could really connect up the two intended meanings of "science in practice": the practices of science, and science operating in practical realms of life.
To improve this situation it is not sufficient, nor helpful, for philosophers to go in with the idea that we know what science truly is, and will teach educators how their work should be done. Instead I take inspiration from Dewey's advocacy of learning through inquiry, with a focus on the experience of students. So I ask: how can HPS help students have a true experience of science? How can students be given the opportunity to truly experience natural phenomena, and to experience scientific inquiry through the practices of discovery and justification? Here I return to my idea of HPS as "complementary science". Even with the best intentions, it is difficult to promote independent inquiry on the part of the students if we are teaching scientific content that is already widely known and accepted. HPS can help by pointing to questions that today's mainstream science does not address and does not have ready answers to. I illustrate the proposed mode of teaching through examples taken from my previous and current work, including the anomalous boiling of water, the puzzle over the workings of the Voltaic battery, and the "tower argument" against heliocentrism.

Abstracts of Posters

(in alphabetical order by last name of corresponding author)

Can science education really remove superstitions?

Deepika Bansal
Mahindra University
deebans.88@gmail.com

This paper is an attempt to revisit and examine the claim that learning science can help individuals get rid of superstitions. While superstition is not a topic of much contemporary research interest, the motivation for this paper comes from my engagement with teaching a course entitled 'science and society' to undergraduate science students at an elite higher education institution in India. The course begins with a discussion on the legacy of scientific temper, and the efforts to disseminate it in Indian society. The phrase scientific temper is similar in ethos to phrases such as scientific attitude or scientific thinking. It is supposed to help the country and the country-people get rid of superstitions, prejudices, and all kinds of backwardness. Through institutional mechanisms of formal schooling to social movements which took scientific knowledge to the people, the task was arduous, and the concept not sharp enough. All of this work rested on the assumption that if people learnt science, they would cease to be superstitious. They will become rational; they will examine each of their beliefs and practices scientifically, seek and then follow the evidence. But this is not a story of India alone. For most philosophers, the scientific revolution of the seventeenth century was the basis to attack superstitions and other religious beliefs and practices across Europe. However, in light of critical scholarship on science that challenges its special nature, its privileged claims to rationality and objectivity, and the uniqueness and singularity of its method, such assertions seem tenuous. Superstitions disappear from a society through the concerted impact of many factors of which science/education is one part. Keeping other factors constant, how do we understand the role of science learning or scientific thinking in the elimination of superstitions? In this paper, I wish to challenge the notion prevalent since the dawn of scientific revolution that science is the antidote to superstition. I will do this by examining the assumptions behind such a claim. One, most superstitious thinking was often conflated with religious thinking. As science and religion increasingly became opposites of each other, all elements of religions were, including most superstitions which were of religious nature, became amenable to be obliterated by science. By showing that science and religion were not always at odds with each other, I will challenge the soundness of the proposition. I will then go on to argue that scientific knowledge or method as taught in science classrooms alone do not offer sufficient intellectual resources to combat all superstitions in the lives of ordinary people. It is my contention that learning scientific principles and laws, and learning how to conduct experiments will not automatically lead students to not engage in superstitious practices. The superstitions are to be discussed explicitly in the classroom and students should be involved in some discussion on epistemology and metaphysics (of science) to help them distinguish between rational and irrational beliefs and practices. In other words, epistemology and metaphysics of science are more useful in fighting irrational beliefs than Newton's laws and the Periodic table.

How can an integrated history and philosophy of science help in the context of the debate between new mechanism and processual ontology? The relevance of the other biological synthesis of Julian S. Huxley (1887-1975)

Stefan Bernhardt-Radu
University of Leeds
StefanB-R@outlook.com

Yihan Jiang
Chinese Academy of Sciences
jiangyihan@ucas.ac.cn

The integrated history and philosophy of science (iHPS) offers a powerful methodological toolkit for addressing live debates in the philosophy of science by grounding them in the details of scientific practice. This paper demonstrates the value of iHPS by intervening in the current debate between processual and "new mechanical" philosophies in the philosophy of biology (Dupré & Nicholson 2018). Recent work has advanced a processual interpretation of genes, viewing genes not as autonomous, self-subsisting mechanisms but as elements embedded within, and constituted by, developmental processes (Griffiths & Stotz 2018). Nonetheless, processual accounts continue to face resistance from philosophers who defend a more object-based conception of genes, treating them as stable, discrete entities. Austin (2016) has argued that processualists reject a mechanicist philosophy that was current in the 17th-century, but has since gone 'outmoded'. A revised mechanistic ontology can explain various biological phenomena just as well as and even better than process ontology. We disagree.

There are both historical and philosophical components to Austin's argument. On the historical side, processualist philosophy, he argues, packs a punch because it picks on a version of the mechanistic philosophy that is old-fashioned. Old mechanicist philosophers used to be thinking of objects in isolation – but today 'new' mechanistic philosophers do this no longer. Philosophically, mechanisms, they say, work in contexts: the activity of a part dynamically emerges in interaction with other parts, rather than in isolation. Processualists fail to see this because of their selection bias, picking on old-fashioned mechanicist accounts. In a word, they are guilty of cherry-picking.

Here we challenge Austin's 'dynamic' mechanistic account and the cherry-picking fallacy. Integrating our philosophical claims with history to challenge Austin's cherry-picking charge, we demonstrate our case for processualism through an account about the evolutionary biologist Julian S. Huxley (1887-1975). He is especially relevant in relation to Austin's argument for cherry-picking, for he accepted the importance of mechanisms, with remarkable relevance for our debate in the 21st. He was a serious subscriber to genetics, including prominent geneticists T.H. Morgan's work, and Richard Goldschmidt's view that gene actions emerge only in context. While we cannot understand Huxley retrospectively from 21st-century philosophy, we can, by grounding his thinking historically in early 20th-century Oxford, see how he challenged the mechanics of Mendelian genetics of his day, and why he thought that it was not just context that mattered, but contexts in time. One had to track the evolutionary and developmental history of the organism in question to understand how a specific mechanism worked the way it did. Mechanisms themselves emerged in time, through what he at one time called 'processes', rather than being 'transcendental', as Austin argued. This paper will use hitherto unexamined and unpublished material, in conjunction with an exploration into the debate between processualists and 'new' mechanists, to show how, to Huxley, time preceded mechanisms, and how processes were crucial to understand them. In

this way, we gain traction for a processualist account with a more recent historical case study, one highly engaged in genetics. Integrating a history and philosophy of Huxley's biology can then shed new light on a lively debate in contemporary philosophy of science.

Fostering integration across the diverse methodological landscape of human behaviour research

Ingo Brigandt
University of Alberta
brigandt@ualberta.ca

Various fields are involved in the study of human behaviour. Previous philosophical discussions have either emphasized the disparity of different methodological approaches to human behaviour that would make integrative connections across fields impossible, or pointed to integration that has already been achieved between some of them. To first obtain a more nuanced understanding of the methodological prospects as well as hindrances to integration, my discussion does not characterize human behaviour research in terms of monolithic approaches. I instead analyze the methodological landscape in terms of different individual investigative, methodological, or explanatory commitments, some of which facilitate integration, while others that discourage integration are also present within the same field. My case study on the epigenetics of neuropsychiatric conditions will document such methodological diverseness. Second, beyond prior accounts' focus on descriptive issues (the impossibility or presence of integration), my discussion centers on the desirability of more integration for normative reasons, especially for social-political reasons. To this end, my analysis of the molecular field of epigenetics will assess integration with perspectives in neuropsychiatry that address the role of the human social environment. With particular attention to gender differences in psychiatric conditions, I argue that such integration would be valuable for improving human mental health and reducing societal inequities. While there are methodological hindrances to integrative connections between molecular and social-environmental perspectives, this calls for a critical engagement with such methodological commitments, combined with the fostering of those methodological resources that are conducive to normatively desirable integration.

The teratological blindspot: challenges for a philosophical understanding of normativity in the developmental study of malformations and synthetic embryos

Leyre Celada-Marcén
leyrecel@ucm.es

This poster addresses the problem of biological normativity, specifically the challenges contemporary philosophy of biology faces in explaining abnormal development emphasizing its direct connection with scientific practice in developmental teratology and synthetic embryo research. Within naturalism, two dominant schools can be distinguished: the etiological approach, which defines function based on evolutionary history but risks epiphenomenalism by overlooking the organism's present state; and the systemic approach, which focuses on current causal contributions but struggles to distinguish genuine

malfunction. The Organizational Approach (OA) has emerged as a promising third option, grounding normativity in organizational closure, the circular causal network through which a biological system maintains itself. Under OA, a trait is functional if it contributes to this self-maintenance, and dysfunctional if it fails to do so (even if the organism continues to survive), thereby supporting an actualist view of biological teleology.
However, as is the case with other normative frameworks focused on evolved lineage, the OA, by centering on the constituted adult risks overlooking the very generative process of constitution itself, including the development of abnormalities, or teratologies. I argue that such abnormal developmental outcomes cannot be interpreted neither as mere functional failures or diseases, nor as simple evolutionary dead ends. Instead, they sometimes represent viable, albeit atypical, structural possibilities. Drawing on the work of the evolutionary developmental biologist Pere Alberch, this poster argues that these forms arise from internal physicochemical constraints rather than external selective pressures, occupying a discontinuous morphospace that challenges adaptationist and function-based accounts. Nonetheless, this framework does not deal with the functional dimension of teratological forms. Therefore, the challenge is reconciling the OA's emphasis on present-time self-maintenance with the reality of developmental deviation of normal function. If a teratological embryo maintains its organization, the OA might classify it as functional, while ignoring its trajectory toward an unviable future form or one with drastically reduced lifespan. However, this solution does not allow discrimination between normal and abnormal forms in development.
To address this blindspot, I propose extending the OA by incorporating what I call a normativity of becoming: a framework that recognizes that, in development, the normative reference point is not only the present system but the constitution of a future one. I advocate a pluralist framework that integrates organizational viability with etiological and morphogenetic understandings of development. I propose replacing binary conceptions of biological normativity with a more flexible approach that treats it as a cross-system, cross-scale phenomenon, aligning with the biological mistake-making framework recently advanced by Oderberg et al. Such an approach, I suggest, could reframe teratologies and aid in addressing practical challenges in clinical practice. In both teratological research and synthetic embryo studies, embryos are evaluated not solely on present metabolic integrity but on their developmental potential to constitute a viable organism. These fields therefore presume a forward-looking normative criterion already operative in scientific practice. Therefore, to align philosophy with scientific practice, I argue that we must expand our framework to recognize that the goal of development is not just maintenance, but constitution.

A Blind Morphism from the Iteration of Multimodal Determination and Model Design

Mario Chalita
Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba (UNC), Córdoba, Argentina
mario.chalita@mi.unc.edu.ar

This poster examines how multimodal determination (MD) generates a shared pattern between model and target during model construction. In MD, heterogeneous methods jointly constrain what can count as a robust result. The result is then incorporated into a broader network or pattern. This pattern is not a pregiven formal entity but an extract of regularities

that the model designer draws from nature and projects into the model. By treating a model as a material device, it can be argued that the pattern it embodies stands in a morphic relation to the pattern in the target system. This view reproduces certain structuralist schemas, but replaces formal (logical or mathematical) structure with a robust, empirically determined pattern.
Initially, the pattern is only partially defined. A model-design proposal leads to the design of a new experiment, which can either confirm the proposal or require its revision. Successive designs gradually give rise to a completion of the pattern, which is always assumed to be equivalent to the pattern in the target. The process of successive determinations produce an asymptotic elimination of unknowns and a reduction to the minimum of the conditions required for the result to occur. This minimal set of elements, considered to be the condition of the result, includes the pattern. Once the pattern is sufficiently delineated, the cycle of MDs reaches its end. This means the dynamics of MD produces successive reconfigurations of the pattern oriented toward a threshold.
The final step in this process of model design consists in testing the pattern under worst-case conditions, outside the parametrized environment in which it was defined. This threshold is only the epistemic condition for the decision to end the multiple-determination process—a decision that also involves social, political and economic constraints. A corollary of this is that the value of MDs is not expressed in a quantitative increase in the robustness or reliability of the result or the pattern, but in a dynamic of successive model designs and pattern redefinitions oriented toward a stability threshold.
The threshold separates a validation period, in which multiple determinations are being performed, from a stability period, in which the result and the pattern are considered reliable and their robustness is no longer questioned, even though their stability may continue to increase through use. In the validation period, even when a determination fails, the morphism itself is not thematized; instead, the proposed model pattern is considered wrong or flawed. In the stabilization period, the morphism becomes a fact and is rendered invisible: the model's pattern and the target's pattern are treated as one and the same. This omission of reflection on the morphism is a condition for the functioning of science, which always operates in a highly parametrized environment—that is, with a huge number of elements whose reliability and robustness are not questioned.

Effects of incorporating scientific uncertainty into undergraduate STEM classes

Elizabeth Coody
Morningside University
coodye@morningside.edu

Brandon Boesch
Morningside University
boeschb@morningside.edu

Shannon Claxton
Morningside University
claxtons@morningside.edu

Aaron Bunker
Morningside University

bunkera@morningside.edu

Brian McFarland
Morningside University
mcfarlandb@morningside.edu

To meet the growing demand for a qualified STEM workforce and a scientifically literate citizenry, there is a critical need to educate scientifically literate students in STEM classes. Scientific uncertainty is an important part of the advancement of science. Despite its centrality to scientific practice, many students do not understand this uncertainty well. It includes the familiar uncertainty associated with measurement and statistics, but also the inherent uncertainty associated with the ever-present possibility of theory change, explanatory uncertainty when it comes to identifying the etiology of individual instances of complex phenomena, and the uncertainty of scientific predictions. The purpose of our NSF-funded project was to improve undergraduate students' scientific literacy by introducing concepts of scientific uncertainty within the context of introductory STEM courses. We hypothesized that enhancing STEM education courses with lessons that integrate the concepts of uncertainty into the scientific content knowledge and science process skills of the course would improve students' understanding of uncertainty, their scientific literacy, and their ability to communicate scientific uncertainty to others. To determine the effects of incorporating scientific uncertainty into STEM courses faculty from a wide range of STEM disciplines who were invited to participate in a week-long summer workshop. Faculty were selected to participate in the workshop based on the quality of ideas for how they might incorporate scientific uncertainty into their classes. The workshop was broken into 3 parts: 1) A theoretical and philosophical account of scientific literacy and scientific uncertainty 2) guided examples of how one might incorporate scientific uncertainty in a STEM class and 3) workshopping towards a plan of how each participant will incorporate scientific uncertainty into their STEM class. Specifically, during the workshop time we worked with faculty participants to ensure that each proposed project was:

1. Meaningfully integrated into the course as a whole, over several weeks and units, rather than being offered as a one-off, stand-alone module.
2. Deep and substantive in its presentation and exploration of uncertainty.
3. Feasibly designed to be accomplished in the time allotted.
4. Original with a novel design and approach to the question of uncertainty in the course content.
5. Designed to ensure time for pre- and post-tests in the courses.

A pre- and post-implementation survey was conducted for both control classes (with no added material on scientific uncertainty) and intervention classes (with the added material on scientific uncertainty). Additionally, students from intervention classes were randomly selected for pre- and post-interviews to determine whether students were able to understand scientific uncertainty and better prepared to communicate about it with others. Data is still being collected, but will be completed and analyzed by June of 2025, so results will be shared at the SPSP meeting.

Is PSP bound to ontological pluralism, and what does it entail? A metaphysical discussion in the context of an inquiry with biologists.

Sélène Domino
IHPST, Université Paris 1 Panthéon-Sorbonne. ISYEB, Muséum national d'Histoire naturelle, Paris.
selene.domino@mnhn.fr

This contribution tries to address the question of the possibility or impossibility of a demonstrative value of empirical studies of scientific practice for different types of ontological claims, as it arose in our ongoing empirical philosophical inquiry, and to spell out its implications.
To set the frame, it joins the time-honored discussion on realism and the conditions of knowledge through a critical appraisal of the project of naturalistic metaphysics in its landmark formulation by Ladyman & Ross (2007). By agreeing with the requirements they state but noticing some incoherences, we claim that the tension between their attempted monism and their naturalism (there understood as fidelity to science) is symptomatic of an unavoidable pressure exerted by the latter towards some form of ontological pluralism. It constitutes an a fortiori argument for a convergence between naturalistic metaphysics and positions resembling Hasok Chang's pluralism.
Showing how an a priori choice on the primary relevance of current science (be it seen from its practical or its theoretical side) for metaphysics is tied to the stance of ontological pluralism leads us to reconsider what an empirical study of science can adequately hope to bring to the metaphysical discussion. The difficulties of our ongoing empirical inquiry on the ontologies of diverse researches across biology are tackled through a tripartition of ontological questioning. Structured interviews with biologists spread across all biological fields were able to yield descriptions and comparisons of "ontological landscapes" and "ontological statuses". On the other hand, discourse on "ontological stances", including pluralism, has to assume another form in order to escape a circular reasoning, which would take its interpretation of scientific practice as a proof of what is actually an already-implicit stance.
This observation leads us to open up the debate on the aims and expectations of inquiry in the philosophy of science, going beyond the exclusivity of a logic of normativity emerging from a descriptive enterprise of philosophical inference with science standing as proof, towards either a set of proposals to be taken up as tools by scientists according to their usefulness for their own needs, or an outlook on the sense of scientific activity in its relationship with reality (which allows a reconsideration and a broadening of what discourse is deemed useful to scientists). The latter brings a justified interest in ontological pluralism for scientists, even as it is advanced through forms of inquiry in which it is embedded rather than resulting.

A Philosophical Approach to Rare Diseases

Anne Fenoy
IRPHIL, Faculté de philosophie, Université Jean Moulin Lyon III
anne.fenoy@univ-lyon3.fr

This poster introduces an ongoing research project in the philosophy of medicine on the concept of "rare disease." The poster is structured around three main axes.

First axis — Definition, context, and usage.
Rare diseases are primarily defined through quantitative and epidemiological criteria. In the European Union, a disease is considered rare when it affects fewer than one in two thousand people. This statistical approach delineates a heterogeneous set of highly diverse diseases. Sociologist Caroline Huyard (2012) traced the recent history of this category, which emerged in the 1970s in the United States. Today, the notion of rare disease plays a structuring role in public health, organizing research and care activities around reference centers and specialized care pathways, particularly in France, through initiatives such as Orphanet and the National Plans for Rare Diseases. In May 2025, the WHO made rare diseases a global health priority.
Second axis — Literature review.
Exploratory review shows that while social sciences have widely addressed this notion (Gargiulo & Mandel, 2023), the philosophy of medicine has so far devoted little attention to it. In social sciences, research focuses on (i) the definition of rare disease and its relative nature across geographic or political contexts (Richter et al., 2015), including qualitative criteria (e.g., the predominance of genetic or pediatric diseases; Boucand, 2018); (ii) clinical and diagnostic uses; (iii) the progressive structuring of care and research, including medical classifications nationally and internationally; and (iv) specific rare diseases. Philosophical reflections primarily concern (i) resource allocation, (ii) patients' lived experiences, (iii) research challenges due to scarce data, (iv) access to treatments, and (v) connections with other philosophical concepts such as genetic diseases, chronic diseases, neglected diseases, personalized medicine, and experiential knowledge.
Third axis — Working hypothesis.
I hypothesize that rare diseases constitute a paradigmatic case for thinking about epistemic plurality in medicine, although this perspective is still largely unexplored (Rosàs Tosas, 2025). These conditions involve specialized knowledge that most of the medical community does not possess, giving patients an active role in knowledge production, circulation, and validation. By adopting a "conceptual scaffolding" approach (Friedman, 2025), I aim to analyze the unique distribution of knowledge to enrich the philosophical understanding of epistemic justice in healthcare and to refine reflections on the role of experiential knowledge in contemporary medicine.

On synthetic medical data: values hierarchies, reification, and zombies

Yael Friedman
Integreat, Department of Mathematics, University of Oslo
yaelf@uio.no

Anna Smajdor
Integreat and Department of Philosophy, Classics, History of Art and Ideas, University of Oslo
anna.smajdor@ifikk.uio.no

With rapid advances in AI technology, there is an increasing demand for data to train AI models. These models are used in medicine and health care for clinical research, diagnosis, drug development, and prevention. However, while significant amounts of health-related data are generated, large datasets may not be available due to privacy concerns and regulatory frameworks (e.g., GDPR). To obtain patients' data, informed consent is required, therefore

creating a high threshold for data sharing. In addition, for rare diseases or cases of “no diagnosis”, large data sets are not available.
Synthetic medical data (SMD) has been introduced to address data-acquisition challenges; It has been described as a solution that enables model training without compromising patients’ privacy and is therefore considered "ethically desirable" (Bellovin et al. 2019, Savage, 2023).
In this talk, we will introduce three main models of SMD: Generative models, Differential Privacy (DP) models, and the C-vine copula-based synthetic data generator. As we will show, these models are not value-neutral, and they pose both epistemic and ethical trade-offs.
We will illustrate how SMD use creates new value hierarchies:
Privacy over transparency: SMD can obscure the specific characteristics and nuances of real patient populations, which can create explanation gaps that undermine clinicians’ ability to interpret model outputs, erode trust, hinder bias mitigation (Susser et al., 2024), and knowledge development, and risk long-term de-skilling. These risks are especially relevant in the case of rare diseases, where small original datasets and oversimplified synthetic approaches can lose critical dependencies and idiosyncratic features, leading to misleading clinical insights.
Scalability over fidelity: developers of SMD aim for low individual fidelity (to protect privacy) but high distributional fidelity, yet, as we will show, noise and selective-distortion methods can break crucial dependencies or entrench biased dependencies from the original data (ibid.), perpetuating the “garbage in, garbage out” paradox (in the process of keeping what seems as important dependencies in C vein copula models), or even introduce new biases (when adding noise in DP models). The preference for scalable, quantified data over individual experience holds risks of alienating patients, erasing contextual and minority experiences (creating what we call “zombie data”), expanding diagnostic disparities, and institutionalizing these trade-offs as a new normal.
We will discuss how this prioritization fosters fungibility in medical practice, which we conceptualize as an advanced stage of the medical reification process. We will distinguish between the stage of “datafication”, which is characterize by decontextualization and translation of messy human experiences into transportable data points that creates a partial, curated technical form of objectivity (Leonelli,2016; Nguyen,2024), and SMD, which goes a step further, generating “zombie data” that deepens reification by stripping lived experience, context and agency from those experiences. Finally, we will argue that this new level of reification is a step forward in the process of commodification and, hence, in capitalism.
References
Bellovin, S. M., Dutta, P. K., & Reitinger, N. (2019). Privacy and synthetic datasets. Stan. Tech. L. Rev., 22, 1.
Leonelli, S. (2019). Data-centric biology: A philosophical study. University of Chicago Press.
Nguyen, C. T. (2024). Value capture. J. Ethics & Soc. Phil., 27, 469.
Savage, N. (2023). Synthetic data could be better than real data.
Susser, D., Schiff, D. S., Gerke, S., Cabrera, L. Y., Cohen, I. G., Doerr, M., ... & Wagner, J. K. (2024). Synthetic health data: real ethical promise and peril. Hastings Center Report, 54(5), 8-13.

Reconnecting "Science Practices": What Science Education Can Learn from Philosophy of Science in Practice

Ron Gray
Northern Arizona University
ron.gray@nau.edu

Over the past decade, science education has embraced "scientific practices" as the foundation for reform efforts such as the Next Generation Science Standards (NGSS) in the U.S. While this shift drew inspiration from philosophy of science, its conceptual grounding largely reflects mid-twentieth-century accounts of scientific reasoning. The result is a simplified depiction of practice that fails to capture the complexity of how knowledge is actually produced in contemporary science.

Recent work in philosophy of science in practice offers a richer and more dynamic picture of inquiry. Scholars such as Tarja Knuuttila, Sabina Leonelli, Adrian Currie, Angela Potochnik, Wendy Parker, Mary Morgan, and Melinda Bonnie Fagan describe science as a heterogeneous, tool mediated, and context dependent activity shaped by epistemic aims, material constraints, and methodological pluralism. Their research highlights key themes that resonate with science education: models as epistemic artifacts (Knuuttila), data as mobile and transformative (Leonelli), reasoning under uncertainty (Currie), idealization as an epistemic strategy (Potochnik), robustness and credibility in simulation modeling (Parker), models as mediating instruments that connect theory and the world (Morgan), and explanation as locally embedded in practice (Fagan). Collectively, these perspectives reveal how scientific practices are dynamic systems of coordination among people, tools, and aims, rather than static sets of procedures.

These philosophical insights can also inform how teachers structure modeling activities, data analysis, and explanation building in classrooms. Teacher education programs could integrate these perspectives by helping teachers design inquiry routines that foreground epistemic tools, uncertainty, and iterative model refinement.

To illustrate their relevance, this poster connects these ideas to examples from ongoing research on model-based inquiry (MBI) and historical epistemic literacy. In MBI classrooms, students engage in iterative model construction, testing, and revision in ways that mirror Knuuttila's concept of models as epistemic artifacts and Currie's methodological omnivory. Students construct competing explanations for the same phenomenon, exploring uncertainty in ways consistent with Parker's and Potochnik's discussions of idealization and robustness. Work on historical epistemic literacy extends these connections, drawing on Fagan's emphasis on locally embedded explanation and Currie's attention to reasoning from traces to design classroom activities where students reconstruct past events using fragmentary evidence. These examples demonstrate how philosophical accounts of practice can inform the design of inquiry experiences that better reflect the epistemic diversity of real science.

The poster concludes by inviting collaborative formats such as joint commentaries and conceptual papers, interdisciplinary workshops bringing philosophers and educators together, and empirical philosophical studies that examine classrooms as living laboratories of scientific reasoning. It invites dialogue between philosophers and science educators around shared empirical questions: How can philosophical accounts of practice inform classroom research? How might science education serve as a living laboratory for studying epistemic activity? Reconnecting the two fields could advance both, ensuring that what counts as "scientific practice" in education remains aligned with the evolving understanding of practice in philosophy of science.

Bias and Expertise: Think tank framing in Canadian news media

Jacob Hamel-Mottiez
Université Laval
jaham43@ulaval.ca

News media often rely on experts. Since experts are highly fallible and since their messages are tainted by their own values, news media most certainly have a responsibility to choose appropriately their experts, and to present their message in a suitable way. However, news media can fall short of this duty, a phenomenon which has been documented as media bias in the literature. This bias can happen at multiple levels such as the source gathering stage (e.g. which experts are chosen to contextualize the issue); at the writing process (e.g. is this expert framed as well-established or controversial) or at the editing stage (e.g. is the news making the cover or it is in the middle). Given the impact of slanted news on issues such as political polarization, reflecting in a systematic way on how news media is prone or not to these various biases is of crucial importance. We focus on a particular and contested expertise which has seen an increase in its use by journalists in recent years: think tanks. Some studies have already documented source selection bias when it comes to which think tank news media use. However, almost no studies documents to which extent news media are biased in the way they frame think tanks and their message.
In this poster, we contribute to this gap in the literature at the intersection of epistemology of testimony and journalism studies. Given that it has already been shown that journalists are biased towards the selection of think tanks (i.e. the gathering stage), we look in this study at how think tanks and their messages are framed (i.e. the writing stage). Do journalists make use of textual cues that are more charitable to think tanks with which they share ideological affinity? Conversely, are they less charitable in their word choice and labeling when framing think tanks and their messages that are ideologically unaligned with their world view? Considering the gatekeeping role of news media and the fact that an important majority expect that journalists exhibit impartiality, assessing if media fall short of this expectation is of primary importance, especially regarding their accountability to the public.
With a corpus of over 20,000 mentions of think tanks in Canadian media, we show that the propensity of a media organization to use credibility cues in relation to a think tank depends on the ideological affinity between the two. More specifically, ideological affinity is linked to media bias in the contextualization of the expert and its message measurable as:
1. a higher propensity to use positive cues;
2. a lower propensity to use negative cues;
3. a lower propensity to use polarizing cues;
Given these empirical results, we reflect on the normative consequences of such findings.

Social science, metaethics and public policy: Conceptual problems in the 'turn to values'

Keryn Hassall
University of Queensland
k.hassall@uq.edu.au

THE CONTEXT: Across multiple social science disciplines, researchers are writing about 'values' in relation to public policy issues. Research into 'values' is seen as a means to better

understand what matters to people and communities in order to inform policy development or evaluate policy effectiveness. It is presented as a way to include ethical, environmental or social values that have been excluded from conventional policy analysis centred on economic self-interest.
THE CONCEPT: This literature contains diverse colloquial conceptions of 'values' which are often poorly articulated and disconnected from philosophical work on values. Much of this work uses an older noncognitivist concept of 'values' that has been superseded in philosophy but persists in social science. The literature appears to reflect twentieth century debates on metaethics and the role of values within science and politics, as translated through social science methods textbooks and canonical texts.
THE PROBLEM IT CREATES: Unconsidered use of this older concept of 'values' does not align well with researchers' stated intention of using 'values' to draw attention to ethical and normative concerns. Instead, it reinforces the personal interest framing of policy analysis which they seek to counter and effectively positions citizens as akin to customers who expect to have their preferences served by the state.
THE UNDERLYING PROBLEM: Social scientists are rarely exposed to topics in moral philosophy, metaethics or discussion of values in science and therefore they have few conceptual resources to think about 'values' in relation to public policy. In the 1990s, a few social scientists with training in philosophy identified the persistence of emotivist and prescriptivist metaethical assumptions as a problem for social science but nothing happened and now the problem is mostly invisible as researchers are not aware that their colloquial concept of 'values' has a specific historical context.
WHY DOES THIS MATTER? This conceptual problem in social science is not simply a problem for social science. It influences the advice given to policy-makers and is a barrier to the effective use of ethical and normative arguments in policy advice. It needs philosophical attention.
MY INVITATION TO THE SPSP COMMUNITY: I've spent 20 years working in social science and seeing many conceptual problems that need philosophical attention as social science lacks the conceptual resources. I want to make this problem more visible to the SPSP community and invite you to bring your expertise to help social science and policy practitioners.
THE POSTER: This is a work-in-progress study of the conceptual problems in the 'turn to values' for policy-oriented research. My poster will give examples of the problem in practice, the metaethical backstory in social science textbooks, and the implications for policy advice. I will also explain my data collection method to elicit and unpack the values conceptions with social science researchers which provides greater depth to my literature analysis.

Causal Process Modelling for Better Evaluation and Planning

Brendan Kelters
Durham University
brendan.t.kelters@durham.ac.uk

In their new book Causal Processes and their Warrant: A Practical Guide Cartwright et al develop an approach to detailed modelling of causal processes, describing a novel taxonomy of evidence types and how evidence of these types, concerning each step in a proposed causal process, may be used to build cases that this process does or doesn't work as proposed. This approach, Cartwright et al argue, may be used to enrich both post hoc and ex ante evaluation

of practical policy, as well as research, and can draw on both qualitative and quantitative evidence.
The proposed poster will extend this approach, combining a detailed and condensed articulation of the approach itself with an account of how the researcher or evaluator can negotiate systematically between their developing causal model and emerging evidence to efficiently produce a well and transparently validated model of the causal process they're studying. It will then describe a method for systematically using such models in evaluation and planning, through supplementing Cartwright et al's 'enriched theories of change' and 'evidence-roll maps' with novel 'pathway aggregation charts'. These charts will methodically guide the user through the process of refining a plan through reducing their produced account of available causal pathways into sets of required inputs and using the results of the modelling stage of analysis to determine chances of success in a manner amenable to cost-benefit analysis.
The poster will be organised into three stages of inquiry, with a schematic account offered of how the evaluator or researcher should traverse through these stages and back to earlier stages in an epistemically sound manner. Stage 1 will be 'envisaging'; the process of describing the causal pathways in which you are interested, incorporating significant details highlighted by Cartwright et al, each of which will be defined and discussed. Stage 2 will be 'testing', the process of collecting and integrating evidence of kinds outlined by Cartwright et al, through a structured process (condensed from Cartwright et al) here described. Adverse evidence concerning 'underlying systems', at this stage, will return the user to the envisaging stage, while evidence concerning absent supporting factors and other defined model components will complicate implementation options. Stage 3 will be 'integration', the process of forming a unified account of the causal system under investigation through integrating the potentially many causal pathways identified and evidenced through navigation of Stages 1 and 2. Here, I will describe pathway aggregation charts and show how they may be used to organise, in a transparent and helpful way, the potentially complex causal possibilities that exist within the investigated setting. The output of this final stage of inquiry will be a list of intervention-sets, each associated with a set of uncertainty-adjusted causal pathways and a local estimation of chances of success. These, in turn, are apt to serve as inputs for cost-benefit analyses.
An illustrative example will be continuously developed alongside the description of the overall method, such as to help the reader see how to use the approach described.

Is Cognition Low-Dimensional?

Phillip Kieval
University of Florida
pkieval@ufl.edu

An influential idea in machine learning hypothesizes that many kinds of real-world data occupy a ``low-dimensional'' manifold embedded in very high-dimensional data.
In recent years, this perspective has begun to inform how neuroscientists analyze populations of neural activity, suggesting that biological neural networks also acquire a low-dimensional representational geometry. Such neural manifolds describe the collective dynamics of neural population activity in terms of the geometric properties of low-dimensional subspaces embedded in high-dimensional neural activation spaces. Yet, the nature and explanatory

status of manifolds remains vexed because neuroscientists cannot study them directly. Researchers investigate neural manifolds by using data-analytic techniques to reduce the dimensionality of patterns of neural activity. These techniques involve modeling choices that affect the structure of low-dimensional projections, which can turn out to be fragile to changes in the choice of technique or choices of parameters. Hence, disagreement remains over whether neural manifolds represent mere statistical summaries of neural population data or the real, emergent dynamics that realize neural computations. This paper aims to cut through this strict dichotomy. I argue that empirical evidence supports the idea that different neural systems exploit both high- and low-dimensional representational geometries. I claim that plausible a explanation for such variation lies in the idea that different systems must solve tasks that constrain the optimal solution to the bias-variance tradeoff. The bias-variance tradeoff is a core methodological assumption in statistics and machine learning that implies that models should aim to strike a balance between two competing sources of error. Insufficient bias leads to errors from underfitting, whereby the system fails to learn solutions that can effectively separate task-relevant features in representational space. Variance, on the other hand, refers to errors caused by over-sensitivity to idiosyncratic patterns in past experience. Once a representation becomes too complex, it begins to overfit the patterns in past observations and variance drowns out its ability to generalize to new contexts. The great benefit of flexible, empiricist methods of representation learning is that a neural network can learn solutions with lower effective dimensionality than the capacity of the network itself. In so doing, neural systems seek a balance between fast learning and flexible generalization on the one hand and fine-grained performance on the other. Finding the right balance implies that the brain must tradeoff the costs and benefits of low versus high effective dimensionality. High-dimensionality supports more fine-grained separability of complex features, but may come at the cost of generalization. We might thus think of learning in terms of a kind of representational compression that aims at finding the optimal dimensionality for representing the features of a particular task. Because the informational demands of different tasks impose different constraints on this tradeoff, learning gives rise to a task-specific spectrum with respect to the intrinsic dimensionality of representations. The brain's drive to find optimal solutions to the bias-variance tradeoff thus provides a parsimonious explanation of how a pseudo-modular cognitive architecture could emerge from domain-general, associationist learning mechanisms, whereby different ``modules'' exploit different latent effective dimensionalities.

Authority, legitimacy, and the role of the researcher in adapting and validating protocols: an epistemological analysis [Poster]

Ingrid Lamminpää
University of Florence
ingrid.lamminpaa@gmail.com

Elena Niccolai

Amedeo Amedei

Protocols represent a methodological standardization that enables reproducibility, a fundamental premise in science for assessing the validity of an experiment and eventual new discovery. Legitimacy for these protocols is a process that passes through scientific

community acceptance. Acceptance happens whenever the outcome's authenticity corresponds to adherence to conventions and results that are already achieved in other scientific discoveries/studies. However, the legitimacy of experimental practice is not purely objective, it presents in fact both objective and subjective constraints, articulated on a fuzzy background. Protocols may be viewed as recipes (referencing Federico Boem's definition), and the validity of the experiments conducted through protocols relies on researchers having specific expertise and being trained in a context of a shared epistemic culture. Researchers are called upon to use personal judgment and practical skills, such as tinkering, to adapt the given protocol to their specific context, and at the same time to keep the chance of reproducing the experiment. This need for an expertise in adapting, highlights the concept of implicit knowledge, as science is driven not only by codified rules, but also by the principle of learning by doing, drawing a distinction between theoretical knowledge and practical ability. The process of adapting or validating a protocol may be considered a fundamental example of an intersection between knowledge that is explicit and knowledge that is tacit (referencing Michael Polanyi's concept).
In this article, I examine an experimental protocol and its adaptation and validation in a clinical research context. It concerns the immuno-phenotyping of T lymphocytes from frozen Peripheral Blood Mononuclear Cells (PBMCs) of patients with ALS, and is one of many protocols used by some laboratories to perform a broader characterisation of patients with ALS through an in-depth immune and metabolic profile. In addition to its strictly experimental dimension, it has clinical relevance as it provides results relating to extra-experimental dimensions such as health. Therefore, the significance of these protocols also becomes useful to physicians for diagnosis, prognosis and, possibly, treatment. The purpose of adapting and validating this protocol is not only to identify cell type populations, but also to make these populations clinically operable. Although the phenotypic characterisation of cells is not an indicator of their functions, it could be argued that these data are nevertheless necessary because current clinical practice requires scientists to provide this information. Furthermore, as in many similar cases, the epistemic significance of the protocol should not be inferred in isolation but only within other experimental practices aimed at constructing an empirical/manipulative frame of reference. In my analysis, therefore, I will show how the protocol works in concert with others aimed at considering the possibility that the microbiome (in the broad sense) plays a fundamental role in ALS (microbiome-immunity axis).
By correlating these biological signatures with detailed information on diet, medication history and past infections, the ultimate goal is to define a "holobiont phenotype" for each patient, encompassing their unique host biology and microbial ecosystem. This should enable clinicians to stratify patients and guide personalised therapeutic strategies.

From Stories to Epistemology: Preservice Teachers' NOS Ideas Informed by Scientists with Disabilities' Narratives

Mila Rosa Librea-Carden
University of North Texas
milarosa.carden@unt.edu

Jonathan Hall
California State University-San Bernardino
JLHall@csusb.edu

Understanding the nature of science (NOS) is a key goal in science education but remains a challenging learning outcome. Research indicates that learning from, about, and with scientists can give students authentic views of NOS. Although prior studies have examined diverse scientists' NOS views, the experiences of scientists with disabilities (ScWDs) are largely unexplored. In this study, we used our interviews with seven renowned ScWD participants (Bower, Caicedo, , Bowskill, Dennehy, Jacobs, Minkara and Pugeda), self-identified as persons with a disability, and consented to be named in publications and presentations to examine NOS ideas of preservice teachers based on ScWDs interview. These interviews contain ScWDs lived experiences and impact of their disabilities on their work. Therefore, we examined addressed the following research questions: (1) What are PSTs' NOS ideas based on the stories of scientists with disabilities? (2) What are PSTs' insights about scientists with disabilities?

Grounded in an explicit-reflective instructional approach, the intervention was implemented in a science methods course where 30 participated in the study out of 62 PSTs engaged with curated video interviews, followed by guided discussions and reflective writing in Canvas. We used AILYZE, an online AI-driven qualitative analysis tool that supports content, thematic, frequency, and cross-group analyses, to examine participants' discussion responses. The first author has used AILYZE in a recently published project. AILYZE helped enhance objectivity given the authors' familiarity with participants and their bias toward a NOS research agenda. The analysis focused on PSTs' NOS ideas related to the ScWDs interviews, not on assessing their NOS understanding levels. We reviewed the AILYZE-generated themes, revised them, and confirmed accurate coding of participants' quotes. RQ1 findings indicate that PSTs recognize creativity, sociocultural and human dimensions of science, and tentativeness of scientific knowledge. One powerful PST response highlights systemic sociocultural barriers surrounding Dr. Dennehy's (deaf) COVID research: "I do not think being deaf or having a disability gives you credit or degree any less. You earned it, and you are just as knowledgeable as someone who is able bodied." Another PST, who is visually impaired, linked ScWDs' NOS enactment directly to disability: "Dr. Minkara like me is visually impaired and people do tend to believe that if you are visually compromised your ability to observe." This reveals the biased assumption that scientific observation relying on sight, which can be exclusionary. Participants acknowledged the importance of accessibility in science. Exposure to ScWDs' personal narratives deepened PSTs' appreciation of science as a human endeavor and offers an empowering way to highlight ScWDs' largely absent stories in science education.

This study underscores the value of integrating ScWDs' narratives into science classrooms, not only for learning NOS but also for reframing epistemological considerations in practice. The findings align with practice-based epistemology, emphasizing authentic accounts of ScWDs' lived experiences over abstract ideals. We call for future research on how practice-oriented approaches influence NOS understanding and promote equity in STEM.

Investigating Population Health through a Performative Lens

Elena Popa
University of Seville
epopa@us.es

Contemporary approaches to public health have been subject to various types of criticism with earlier origins (e.g., as pointed out by Valles 2018: 'overreliance on narrow biomedical understandings of health and well-being ... cultural and ethical imperialism ... failures to connect with underlying social problems such as food insecurity') and which have become all the more obvious after the COVID-19 pandemic (cf., Broadbent 2022). Using Valles' (2018) arguments for a population health perspective to address these gaps as a starting point, I will investigate how population health can be framed as a performative discipline, e.g., not only concerned with representing health-related behavior, but with altering it. As discussed in the performativity in science literature, in such cases 'we would evaluate the model's adequacy by asking whether its use has changed its target in the intended way, rather than accurately explained or predicted' (Van Basshuysen 2022, 1460).
I will highlight that both public health and population health have the goal of improving overall health outcomes and both target the health of a population, but while public health has used acting on individual behavior as means, population health has kept its focus on acting at population level. I will argue that work in epidemiology on the higher effectiveness of population-level interventions can provide an argument for the better fit for purpose of the latter approach (e.g., Rose 1985). I will explore further implications by sketching out what a performative model of the health effects of inequality would look like.
The chosen case is important because while the health effects of inequality have been pointed out earlier (Wilkinson and Pickett 2009), no clear health measures to respond to them have been brought forward. A performative framing would model health effects of inequality seeking to improve health through more egalitarian social relations and higher levels of trust. Among others, this would involve highlighting relevant phenomena beyond the medical ones (access to education, social support networks), focusing on health-related behaviours more likely to increase trust (such as those following from a notion of collective responsibility instead of a concept of health as something gained at the expense of others), highlighting institutional aspects (scarcity of services; heavily commodified access to healthcare), highlighting the benefits of living in a more egalitarian society among all social classes. Lastly, I will explore the problem of political legitimacy that has been brought forward as a challenge to performative models alongside a new one, whether the value choices can always be made consciously. I will argue that democratic exercises can help move towards a framework where the interests of those involved are better integrated and where the value commitments become more transparent, although they cannot provide quick solutions.

No Atoms, No Chemistry? Dalton's Chemistry Just Before Atomic Theory

Ana Maria Possamai
UFSC - Universidade Federal de Santa Catarina
anamariapossamai@gmail.com

John Dalton's atomic theory is widely recognized and taught as a foundational milestone in the development of chemical science. Although its historical formulation has been extensively discussed, comparatively little attention has been given to the chemical, practical and experimental dimensions that underpinned its development. Contrary to philosophical claims that portray chemistry without atomic theory as a mere "cookbook" discipline, close examination of Dalton's work reveals that his atomic theory emerged from systematic, reproducible, and carefully controlled observations of artificial phenomena. These observations relied on sophisticated chemical apparatus and a well-established body of

knowledge concerning the properties of matter and chemical reactions. Instruments such as Volta's eudiometer and pneumatic flasks, along with methods for identifying, isolating, and synthesizing compounds and understanding their composition and reactivity, were essential to Dalton's most significant contribution to chemistry. From this empirical foundation, Dalton employed abduction as a central cognitive process to infer the submicroscopic structure of matter, grounded in macroscopic observations and mediated by symbolic and mathematical representations. Unlike induction or deduction—often emphasized in philosophy of science, particularly from a physics-centered perspective—abduction played a fundamental role in the early development of chemical reasoning. By abstracting from observable phenomena, Dalton articulated his atomic theory, a process inseparable from the experimental and material practices that were crucial to the maturation of chemistry as a scientific discipline, and revealing that chemistry has always been fundamentally a substance-based science.

Redefining Fitness for Purpose: lessons from fusion energy systems design.

Marc Raguz
University of Edinburgh
s2611929@ed.ac.uk

This presentation draws from case studies in the design of fusion energy systems to motivate a new definition of fitness for purposes that, while accommodating of Wendy Parker's existing notions of adequacy for purpose, diverges in important ways that allow for evaluating models in dynamic and uncertain regimes. Previous definitions of fitness for purpose are either used synonymously with adequacy for purpose or are subsumed under it. I argue that we need to demarcate the two to accommodate real issues of model evaluation dynamics, especially those made salient by first-of-a-kind engineering regimes such as fusion energy. Such first-of-a-kind regimes, especially in their reliance on integrated modelling, introduce two fundamental problems for evaluating model adequacy. Firstly, they introduce instability in the purpose spaces to which the models are compared; the epistemic goals that a model must serve in order to facilitate overarching practical goals, such as informing design choices and policies, are not fully known from the outset. Secondly, the adequacy conditions of these purposes themselves are also indeterminate from the outset; there is uncertainty about what a model has to do in order to be adequate for some purposes.
The significance of these dynamics of adequacy conditions is that if we rely on an unadorned account of adequacy for purpose, which understands fitness as the degree to which adequacy has been achieved for a given evaluative context, we can only use it to evaluate models in terms of the known or possible purposes at a given point in time. This introduces the danger of building models that are locally adequate and optimised for certain static purposes without consideration for the transferability of this adequacy across changes in the purposes themselves or what it takes to be fit for them. In short, relying only on Parker's notion of fitness alone can only be done for short-term, local model evaluation, and relying on it for evaluative contexts that have not fully converged on the complete set of model adequacy conditions risks ignoring what the models sacrifice in terms of long-term fitness when pursuing this short-term adequacy.
In response to this, I offer an extension of Parker's adequacy for purpose account and a different notion of fitness. I define fitness as the malleability and robustness of a model's underlying assumption sets. More precisely, I define a model's fitness for purpose as its

ability to robustly accommodate a range of competing (possible) purpose metrics simultaneously, where maximum fitness is given by a pareto optimal set of model performance scores across these metrics, and where each pareto optimal set of scores corresponds to a different underlying set of possible construction assumptions for that model. Furthermore, I go on to argue that fitness for purpose is as much a measure of how well this synchronic performance snapshot carries over time - as the underlying assumptions of the model are challenged and as adequacy conditions shift and lessons are learned, which are dynamics described by the stability of these pareto frontiers as assumptions are adjusted. Not only does this put the emphasis back on model assumptions when evaluating fitness, it changes the concept of fitness itself in unfamiliar regimes: fitness is a measure of the trust we can place in the model to not get things drastically wrong and/or have its adequacy collapse entirely as things evolve, rather than how well optimised it is for a given purpose in a given evaluative context. Fit for purpose models are those that allow us to target the short-term in ways that don't invite foreclosure of the long-term.

Integrating HPS: Leveraging Complementary Science in the analysis of hydrogen-forming enzymes

Florian Riemer
Ruhr-University Bochum
Florian.Riemer@ruhr-uni-bochum.de

Hydrogenases are enzymes that catalyse the interconversion between protons and electrons on one hand and molecular hydrogen (H2) on the other. They were first described in 1931, but since then there have been substantial shifts in the methods used to characterise these enzymes and in how scientists conceive of them.
My overall goal in this investigation is to develop an account of scientists' changing conception of hydrogenases and their activity that explains changes in research programmes that have occurred over time - and to develop an improved activity assay for hydrogenases that fills the gaps in today's conception of these enzymes.
In the first decades of hydrogenase research the enzymes were primarily thought of as hydrogen-splitters featuring in bacterial metabolism. Around 1960, it was discovered that hydrogenases are involved in photosynthesis, which lead to the idea of utilising them for industrial applications. Concrete proposals in that direction began to emerge shortly before the first oil crisis in 1973, amid a surge of interest in renewable energy. Over the course of the 1970s, investigating hydrogenases became more and more popular, and they were increasingly identified with their hydrogen production capabilities rather than the reverse reaction of hydrogen-splitting. Around the year 2000, several articles extolled the value of hydrogenases in fighting climate change and air pollution, and the majority of publications since then emphasise the urgency of these environmental issues - and the potential role of hydrogenases in addressing them. This conceptual shift has lead to a methodological focus on maximising hydrogen production activities, potentially obscuring other aspects of hydrogenase activity.
The biochemical strand of my investigation therefore aims to provide a standardised and accessible way to determine and compare bidirectional activities of hydrogenases. This development is informed by the Complementary Science approach as described by Hasok Chang - historical developments in methodology and their broader cultural context as well as

considerations on the nature of hydrogenase and enzymatic activity are employed to direct biochemical experimentation and reporting.
The philosophical goal of this study is to sketch a change in the self-conception of molecular biologists over the course of the last century. My conceptual analysis of the epistemic activities in hydrogenase research shows how changes in the methodology of enzyme analysis were not merely a result of changing resources and capabilities - they were also a response to the ontological shift brought on by a desire to address challenges of the energy economy, leading to new purposes and values emerging in this field of research. Crucially, I argue that these changes did not merely occur as a response to external pressure that forced scientific communities to reassess their priorities but that they also reflect a change in the self-image of molecular biology from a question-answering toward a problem-solving discipline. This change has led to a confident modern science, convinced that it can take on an array of social challenges, from the humanitarian to the entrepreneurial.

A Proposed Linguistic Turn in Philosophy of Science

Monika Rydzewski
University of Wisconsin-Madison
mrydzewski@wisc.edu

Natural kinds have been out of style in recent discussions of scientific classification and taxonomy. Ingo Brigandt argues for cease and desist of the use of the natural kind term in "How to Philosophically Tackle Kinds without Talking about "Natural Kinds,"' proposing instead a general methodological framework of distinguishing kinds based on underlying human aims. Acknowledging underlying human aims in classification seems a worthy pursuit. For instance, when looking to classify languages, one can already imagine a human aims-informed taxonomy would work towards accounting for speaker interaction's influence on dynamic linguistic data. Intuitively, then, languages are bad candidates for any natural kind term involvement in their classification, given their commonly held status as social kinds and the apparent lack of fixed boundaries between them. I argue that, nevertheless, there is merit in employing natural kind term talk in language classification, specifically scientific classification of languages in linguistics. While I'm not decisively arguing that languages themselves are natural kinds, I argue that we shouldn't dispense with natural kind talk in linguistics. The vocabulary helps explain why some boundaries, like those between language and dialect, may not be arbitrary, as the multiple aims underlying language classification considers them to be.
Looking at where languages fall in kinds-talk, I will consider the view of languages as social kinds. This view will hold that languages being social kinds blocks involvement of natural kinds. However, arguments for languages as social kinds do not definitively close off natural kind talk in linguistics. I will argue for invoking natural kind talk in language classification, primarily looking at the example of boundary construal between languages and dialects. I will aim to show that under Brigandt's framework, which broadly accommodates languages as social kinds, the classificatory boundary between languages and dialects would be underdetermined: multiple human aims underlying language classification will put forth different, comparable measurement proposals for languages and dialects. That's where adhering to some natural kind talk remains beneficial. Classifying languages as general social kinds may conflate different classifications of speakers in linguistics. Natural kind talk in

linguistics better negotiates speaker involvement in language classification, even potentially distinguishing speaker classification as a separate taxonomy.
Without going so far as claiming that languages are natural kinds, I show that there is use for natural kinds in adjudicating the boundaries of languages. Investigation of language as a scientific kind ought to lend some rigor to its classification, without dismissing the impact of nonepistemic values at any stage in the process. How we talk about nonepistemic aims may influence the classification structure much like people reflexively change their behavior when being measured. So, if linguistics can stand to benefit from natural kind talk in linguistic classificatory practices, natural kind talk may be viable in other domains as well. Finally, philosophy of linguistics itself has faced a 'boundary' issue in its questions, approach, and distinction from philosophy of language. However, taking up issues like the language-dialect boundary, which have been kept up by sociolinguists, would make philosophical questions in linguistics distinct.

Menarche, Menopause, and the Missing Male Milestones: Gendered Kinds in the Classification of Age

Tashmia Sabera
University of Wisconsin-Madison
sabera@wisc.edu

Scientific research and public health communication concerning the human reproductive lifespan overwhelmingly focus on female reproductive milestones, most notably menarche and menopause. In contrast, male reproductive milestones, such as spermarche and andropause, remain largely unrecognized. Reproductive milestones are not merely indicators of significant changes in reproductive capacity; they also play an important role in tracking health trajectories, assessing disease risk, and understanding patterns of morbidity and mortality. In this paper, I argue that the non-recognition of male reproductive milestones is unjustified and reflects gendered power dynamics embedded in scientific classification.
By non-recognition, I refer to any of the following three conditions: (1) scientific communities are largely silent about a milestone; (2) a milestone is acknowledged in relation to certain health outcomes but is not used to mark significant transitions in the reproductive lifespan, despite its relevance; or (3) public health literature frequently characterizes the milestone as a "myth." I argue that spermarche and andropause satisfy one or more of these conditions and therefore count as non-recognized reproductive milestones.
Menarche denotes the onset of menstruation. By contrast, menopause refers to the permanent cessation of menstrual cycles, clinically diagnosed after twelve consecutive months of amenorrhea not attributable to pathological causes. Both are widely recognized as defining transitions within the female reproductive lifespan.
Spermarche, understood as the onset of sperm emission, and andropause, understood as an age-related decline in testosterone often accompanied by physical and psychological symptoms, are not comparably recognized. Spermarche is often treated solely as a marker of puberty rather than as part of the reproductive lifespan, which limits our understanding of male reproductive functioning and its connections to broader health outcomes and longevity. Andropause remains controversial, in part because it is frequently rejected on the mistaken assumption that recognizing it would require a strict biological analogy with menopause.
Drawing on a pragmatic approach to scientific classification, I argue that reproductive milestones or categories should be evaluated primarily in terms of their usefulness for

epistemic and non-epistemic aims. From this perspective, the non-recognition of male reproductive milestones fails to promote either. I critically examine existing justifications for this non-recognition and show that they are often empirically unsupported, conceptually confused, or grounded in implausible assumptions about the goals of scientific classification. I further argue that the costs of non-recognition are substantial. First, it reinforces harmful medical myths about male reproductive invulnerability, including the assumption that male reproductive capacity remains stable throughout life. Second, insofar as we commit to any particular normative account of reproductive responsibility, non-recognition contributes to a skewed distribution of such responsibility. Third, it negatively impacts male reproductive health by obscuring clinically relevant patterns of aging and risk.
By demystifying the politics surrounding the recognition of male reproductive milestones, this paper aims to correct epistemic injustices in reproductive science and to expose how classificatory practices sustain gendered inequalities in reproductive health and responsibility.

Toward a Unified Approach to Philosophy of Science and Clinical Research Ethics: The Values-in-Science Standard and the Social Value Requirement

Ukyo Shimizu
Faculty of Medicine, University of Miyazaki
ukyo.shimizu@gmail.com

I propose a unified approach to value problems in research design that integrates philosophy of science and clinical research ethics. To motivate this approach, I examine the values-in-science (ViS) standard as a guide for clinical research design.
In the controversy over science and values, Steel & Whyte (2012) proposed the values-in-science standard, according to which "non-epistemic values should not conflict with epistemic values in the design, interpretation, or dissemination of scientific research that is practically feasible and ethically permissible." A central feature of this standard is its commitment to epistemic priority, which rejects compromising epistemic values across the key stages of scientific research. Although this commitment has been criticized by several philosophers, in my view these criticisms are not compelling enough.
Focusing on research design, my argument advances a novel line of criticism against the ViS standard. In brief, the standard is in tension with the social value requirement (SVR), which is well recognized in clinical research ethics. The SVR requires that individual research projects have sufficient social value to justify related risks and resources. On a common interpretation, the SVR can justify a design with lower scientific value and lower risks rather than with higher scientific value and substantially greater risks. Scientific value of a research design is normally conceptualized as necessary but insufficient for social value (CIOMS, 2016), that is, not necessarily "prioritized" in clinical research ethics.
How can we integrate the ViS standard and the SVR in the context of research design? As a possible interpretation, ethically permissible and sufficiently valuable research designs must satisfy epistemic priority. On this reading, epistemic priority is supposed to guarantee sufficient social value, so that we, for example, should prioritize larger trials over small ones, even if that means conducting large first-in-human trials, to avoid compromising scientific and social value. I argue that this implication is implausible and defenders of the ViS standard are also unlikely to endorse it. On an alternative, one that they seem to favor, the phrase "ethically permissible" functions as a prior and independent premise of the epistemic

priority. On this view, the ViS standard is applicable after the SVR has been satisfied, thereby preventing any direct conflict with these two norms. However, I argue this interpretation seriously undermines the significance of the ViS standard. If the SVR already coordinates the epistemic and non-epistemic values of the research design, then the ViS standard becomes unnecessary and redundant.
The underlying problem is that, while the defenders of the ViS standard implicitly assume the ethical permissibility of research design can be assessed without reference to epistemic values, this assumption is untenable, at least in the context of clinical research designs. Nevertheless, my aim is not to reject the ViS standard outright, but rather to refine it from the perspective of clinical research ethics. I conclude that a genuinely unified approach combining philosophy of science and clinical research ethics would be mutually beneficial.

An Ethnomethodological Demystification of the Algorithmic

Sabreen Syeed
University of Liverpool
sabreen.syeed@liverpool.ac.uk

Terms such as AI, LLMs and Algorithms have achieved topical significance, despite the growing obscurity and opacity around the actual workings of these systems and the processes that operate 'behind'. This study proposes how an ethnomethodological analysis can unpack the apparent black boxes to respecify the actual practices that underpin these technologies often mistaken as the 'new ghosts in the machine'. Post-analysis of algorithmic systems often is disengaged from the actual course of scientific activity that produced them. Garfinkel argues this is due to the irreducibility of the embodied production of an object to an account of that work that preceded it. Thus, the ethnomethodological emphasis is on understanding the technology from within instead of approaching it from outside accounts prioritizing what data scientists do practically instead of what they say about their practices.
Ethnomethodological Science and Technology Studies understand algorithms as cultural artefacts, praxeological tools and social constructs that are embedded formally, structurally and effectively in human practices and thus are an integral part of our 'life-world'. Demonstrably, such a view is demystifying, deflationary and most importantly realistic in contrast with the corporate run hype on the one hand and the reactionary and alarmist dystopia from critics on the other. The anthropomorphic undertones of elementary terms like 'artificial intelligence and machine learning also are products of the disengaged accounts of the field. Furthermore, the alarmism around the supposed opacity around how LLMs produce outputs can also be resolved via an ethnomethodological breakdown of the actual practices involved in the algorithm.
Ethnomethodology seeks to study the locally produced real-worldliness of ordinary activities related to the production of a scientific output. This being a practical working out of Phenomenology in social settings, sets out to expunge Cartesian polarities usually used to approach scientific outputs as objects created by the observing and acting agent (the scientist) and instead analyse the technology as an in-situ extension of the social agents in the lebenswelt.
And science increasingly incorporates computer-based simulations, software, big data analysis, and machine learning methods, these have provided a new way of doing science and producing scientific knowledge. Related questions like whether simulation is "experimenting with theories" or is it another autonomous form of knowledge production. Thus, the gestalt

approach of Ethnomethodology offers much promise in the fields of Human Computer Interactions (HCI), Computer supportive cooperative work (CSCW) and Critical Code Studies along with the broader discipline of the Philosophy of Science.

Clusters and Clustering: Natural Kinds Meet Statistical Practice

Zina Ward
Florida State University
zina.b.ward@gmail.com

Practice-oriented philosophers of science overwhelmingly favor cluster theories of natural kinds. The most prominent such theory is Richard Boyd's Homeostatic Property Cluster (HPC) account, on which many natural kinds are associated with collections of properties held together by homeostatic mechanisms (Boyd 1989, 1991, 1999). A number of other cluster theories of kinds have been proposed as well, united by three core commitments (Ward 2023, forthcoming): (1) individual members of a natural kind share many properties; (2) many or most natural kinds do not possess intrinsic essences; and (3) there are discontinuities between natural kinds in property space or in their underlying mechanisms. A major advantage of such accounts is that they cast a wider net than essentialist theories: many kinds widely used in science lack essences and yet still seem to deserve the imprimatur of "natural kind."

Another potential attraction of cluster theories of kinds is that they fit well with scientific techniques for identifying kinds de novo. Researchers use statistical clustering algorithms to classify in a wide variety of fields (Hennig et al. 2015). A cluster theorist attuned to scientific practice might well claim that the use of cluster analysis for taxonomy construction in science bolsters cluster theories of kinds in philosophy (Lewens 2012).

As it turns out, however, the story is a bit more complicated. In this poster, I address the various ways in which statistical cluster analysis might be thought to undermine cluster theories of natural kinds. First, a large array of clustering algorithms exist, which often yield different groupings of objects (Hancox-Li 2020). I argue that not all clustering algorithms pick out property clusters in the philosopher's sense, and that some diversity of clustering results is to be expected. Second, clustering algorithms usually deliver a result even if the dataset to which they are applied isn't "really" clustered (Hennig 2015). I show, however, that there are statistical techniques for determining when clustering is present in a dataset and when it is not (Huang et al. 2015). Third, applications of cluster analysis depend entirely on the properties or features used as inputs to clustering algorithms (Sterkenburg 2023). Moreover, scientists sometimes use cluster analysis to identify kinds in a property space whose dimensions are discovered through dimensionality reduction techniques or automatic variable selection methods rather than specified in advance. I suggest that these practices encourage the cluster theorist to adopt a view on which classification of things into kinds involves a process of epistemic iteration (Chang 2004, 2015).

An Enactive Account of Scientific Models and Representations

Yuyou Wu
University of Leeds

pryw@leeds.ac.uk

Philosophers have long sought to explain what scientific models are and how they yield knowledge. Traditional representationalism treats models as representations of target systems, deriving epistemic value from this relationship. However, attempts to explicate this relationship run into persistent difficulties, leading some to declare it a "dead end" (Sanches de Oliveira, 2018): dyadic accounts—e.g., isomorphism (van Fraassen, 1980) or similarity (Giere, 1988)—miss representation's directionality and capacity for error, while triadic, agent-dependent accounts risk arbitrariness, undermining their capacity to explain how we learn about targets from models.

Artifactualism responds by treating models as material tools, emphasizing how concrete engagement scaffolds scientific inference (e.g., Knuuttila, 2010, 2017; Kuorikoski & Ylikoski, 2015; Currie, 2017; Sanches de Oliveira, 2018, 2022). Yet, scientific practice does not involve a random exploration of a model's countless sensorimotor affordances; it is systematically guided and constrained. The selection and coordination of specific sensorimotor engagements, as well as the inferences made about the world based on them, are guided by representational norms and meanings. Existing artifactualist accounts, by deliberately avoiding representational terminology, lack the resources to explain such systematic selection and coordination. Furthermore, supplementing them with traditional representationalism returns us to the aforementioned difficulties.

I propose a solution via an enactive account. Enactivism—a framework within embodied cognitive science—understands cognition not as the manipulation of internal neural representations, but as the exercise of skillful know-how in embodied, situated action, a process termed "enaction" or "sense-making" (Varela et al., 1991; Thompson, 2007). Through this activity, an organism cultivates adaptive skills, becoming attuned to relevant aspects of its environment and thereby transforming that environment into a domain of salience and meaning relative to its embodied features and capacities. Thus, cognition is a history of structural coupling that "brings forth" a domain of significance, where meanings and norms are co-constituted through dynamic interaction rather than being pre-given.

Building upon this foundation, I argue that many scientific models should be understood as enacted representations—representations "brought forth" through the embodied, situated practice of scientific modeling. They are representational artifacts whose specific content, relations, and norms are not pre-specified but instead emerge from this practice—an advanced form of embodied sense-making. This practice is itself situated within and conditioned by the broader history of scientific practice as a structural coupling between humans and nature, through which we explore and become increasingly attuned to nature's varied aspects by developing and refining relevant embodied skillful know-how.

This enactive account reconciles artifactualism and representationalism, incorporating the former's focus on material engagement and embodied skill and the latter's commitment to representational content and norms. Crucially, it explains their integration: representational relations, meanings, and norms are emergent order parameters that arise from and, in turn, impose top-down constraints on lower-level constituent processes—whether sensorimotor, inferential, or interpretive. This explains the selective and coordinated patterns of material engagement with models as concrete artifacts. By grounding representation in embodied practice, the account provides a coherent, naturalistic framework that captures how models, as inextricably artifactual-representational entities, enable scientific understanding.

Abstracts of Contributed Papers

(alphabetical by last name of corresponding author)

The Epistemology of Multi-Messenger Observations

Sarwar Ahmed
Johannes Gutenberg University of Mainz
saahmed@uni-mainz.de

Multi-messenger observation has been celebrated for its epistemic strength, particularly using the notions of "variety of evidence" (Abelson 2022) and "independent evidence" (Elder 2024). Advocates argue that the method conclusively eliminates alternative explanations in the process of inferring the source system of astrophysical events.
Informed by scientific practice in the field, this paper highlights the methodological and epistemic limitations of how these philosophers of science reconstruct the multi-messenger method. It also proposes a general methodological framework that better captures the situation and perspective of practitioners. I highlight two limitations:
Unconceived alternatives: I argue that these authors assume the canonical method of empirical theory confirmation, which does not have the resources to assess the space of alternatives. That is, the hypothesis that a particular kind of event is the source for the data cannot be deduced from the detections. The problem of unconceived alternatives reminds us that completely unknown astrophysical phenomena could produce similar observational signals.
Theory-ladenness: Contrary to what it may appear to these authors; I argue that the multimessenger method is more theory laden. The very idea of independent messengers, the association of the waves, and that they all originate from a single source system are all based on our trust in the theory of the source and established theories. This trust is justified by assessments that go beyond the detection process in question.
The reasoning behind their arguments resembles Ian Hacking's (1983) argument from coincidence. Therefore, to illustrate the limitations, I draw an analogy to Hacking's argument, which defends the veridicality of microscope observations by appealing to the agreement between different types of microscopes. However, as van Fraassen notes, Hacking's argument presupposes a true explanation that is unwarranted due to underdetermination thesis (1985).
Claiming to have observed an astrophysical source system means committing to the idea that the identification of the source system is warranted. However, van Fraassen's fundamental objection challenges this commitment. Therefore, it is essential to address it especially in our case. This is because, firstly, the astrophysical events are not usually subject to intervention. Secondly, astrophysical observations are sensitive to the location and timing of the astrophysical event, in addition to determining the source system. That is, many parameters in the observation and the uncertainty in stabilizing the signal make determining the source system more difficult and recourse to theoretical reasoning essential.
To this end, I propose a general methodological framework for astrophysical observations that can accommodate both single and multi-messenger observations. This is an eliminative inferential process aimed at identifying the source system from two directions, namely from data to phenomenon and from fundamental theory to phenomenon. That is, the process draws from both observational data and theoretical principles to justify the identification of a source system. The framework allows for a case-by-case allocation of the epistemic emphasis given to theoretical and empirical evidence in the process of justifying an observation. Furthermore, the framework addresses the problem of unconceived alternatives in the observation of

astrophysical events by augmenting the meta-empirical assessments developed by Dawid (2013) to extend the notion of evidence beyond the empirical detections.

References

Abelson, Shannon Sylvie (2022). "Variety of evidence in multimessenger astronomy". Studies in History and Philosophy of Science Part A 94 (C):133-142.
Dawid, Richard (2013). String Theory and the Scientific Method. Cambridge University Press.
Elder J. (2024). "Independent evidence in multi-messenger astrophysics". Studies in history and philosophy of science, 104, 119–129.
Hacking, Ian (1983). Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. New York: Cambridge University Press.
van Fraassen (1985). "Empiricism in the Philosophy of Science", in Churchland and Hooker 1985, pp. 245–308.

A tale of two pragmatic values: on the translational gap between biology and medicine

Yin Chung Au
Center for General Education, National Chung Cheng University
10608041@gs.ncku.edu.tw

This study examines the translational gap between biological basic research and clinical medicine. It seeks to shed light on the limited social recognition of basic research, as discussed by many scholars, by reframing it as a problem of inevitably different epistemic designs laden with incompatible pragmatic pursuits. While the two domains often interact and collaborate, and while they both exhibit pragmatism greatly, I maintain that they define 'useful knowledge' and pragmatic values in partially incompatible ways.
I argue that the translational gap originates because explanatory adequacy sought by basic research does not guarantee efficacy of clinical and therapeutic interventions. The latter requires powerful predictions that are not necessarily warranted by causal claims inferred from interventionist success. Such a distinction turns a methodological problem, namely the debated reconciliation of evidence types (mechanistic vs. probabilistic) into a problem of the coherence of different epistemological aims.
Once considered outside the lab, basic research often faces the translational challenge that the research results might be either incapable of guaranteeing external validity or only limitedly extrapolated to ensure general therapeutic efficacy in real-world complex systems. Previous studies have addressed such a challenge in terms of the failure of mechanistic explanations. Methodologically, I consider this challenge to stem from the dominance of Woodwardian interventionist causation seeking in basic research. As variable relativity (Parkkinen 2022) necessarily applies, the resulting causal claims are context-bound to specific variable sets tailored for research purposes and tested in closed systems. Similarly, the mechanistic approach, as the mainstream practice of basic research, is laden with pragmatic values and thus inherently relative and local. This renders the use of the research results incompatible with the required utility of knowledge in clinical contexts.
The core problem is rather epistemological than methodological because the two forms of knowledge justification are disparate. Although interventionist causation ideally manifests in the effectiveness of clinical/therapeutic interventions, the decision-making processes for

these interventions rely on the predictive, not explanatory, power of knowledge inferred from robust difference-making evidence. These two evidence types have been demarcated by the Russo-Williamson Thesis. While scholars have succeeded in clarifying the epistemological contributions of the two types (Illari 2011), this clarification has not distinguished, in real-world translations between biology and medicine, the knowledge pursuits supported respectively by the two evidence kinds. I propose to unify the existing theses and step forward to account for the actual epistemological demarcation that leads to the translational failure.
I will conclude that basic research and clinical practice define 'usefulness of knowledge' differently, resulting in their different epistemic design and the limitation of relative claims. Future analyses can be done to analyse the co-production of translational failure in therapeutic developments by this gap and commercial imperatives. This should help philosophers of science engage social studies of science and science communication scholars.

What's wrong with circles and lines? Early moderns against mathematical sciences of nature

Ovidiu Babeș
Free University Brussels
ovidiu.babes@vub.be

Heroes of modern science made natural philosophy mathematical. Some disagreed: they argued that mathematics is not a proper part of natural knowledge. Recent literature increasingly calls out such early modern oppositions. This presentation reconstructs a specific strand of British anti-mathematical philosophy of the 18th century which is hitherto unacknowledged. John Sergeant, Robert Greene, George Berkeley, and Thomas Reid resisted the mathematization of (parts of) natural science. I claim that they did so for a similar reason: a shared philosophy of perception. They all held that the complexity of nature is a perceptual primitive, (sometimes) to the effect that mathematizing science necessarily involves ideas disconnected from natural perceptions.
There is a growing interest in explicit reactions against mathematizing natural philosophy (Stigler 1992; Gingras 2001; Roux 2011). Schliesser (2017; 2019) puts forward the concept of 'anti-mathematicism' to capture such reactions. While recent studies have investigated several forms of anti-mathematicism, this literature remains confined to isolated case-studies and does not provide a systematic historiographical category. My tentative systematization is the following. Call 'anti-mathematicism' the cluster of beliefs that mathematics should not be constitutive of natural science. Construed this way, the term conflates different worries. Some are metaphysical: [1] Mathematical sciences define their objects by quantities. Yet quantity is an accidental feature of beings. Since natural science deals with causes, natures, or essences, mathematics cannot be a constitutive part of science; [2] Natural bodies do not actually possess properties such as determinate sizes or shapes. Some worries are epistemological: [3] We do not possess clear ideas of quantities, since such ideas depend on faulty mental faculties; [4] Natural phenomena are irreducibly complex. Since mathematics limits knowledge to one type of perception, it illegitimately excludes others, which are part and parcel of the complexity of material beings. Other worries are practical: [5] A natural philosopher needs many other skills aside from observing and manipulating quantities. One worry is political: [6] If natural science will be performed by mathematicians, lay practitioners will be excluded and there is little hope for a public science. Aristotelians held

qualified versions of [1], Leibniz and probably Descartes held [2] (cf. Nelson 2019; Wilson 2019), Hume held [3] (cf. Demeter 2019), Cavendish held [4] and probably [3] (cf. Peterman 2019), Mandeville and Buffon held [4] and [5] (cf. Wolfe 2014, 2019), Diderot held [4] and [6] (cf. Gingras 2001; Wolfe 2019).
In my presentation, I show how Sergeant, Greene, Berkeley, and Reid form a unitary strand of anti-mathematicism that fits within and improves the reconstruction above. I claim that Sergeant held [1] and [4], Berkeley held [3], [4], and possibly [2], Greene held [4], Reid held [3], and [4]. I show that all four share a view on perception which underlies [2], [3], and [4]. This view puts perceptual complexity first, denying that simple ideas are perceptual primitives that compositionally furnish all mental content. By articulating this philosophy, my presentation readdresses the label of 'anti-mathematicism,' thus gaining an analytical tool for understanding more critiques of early science.

What's mechanistic in mechanistic interpretability?

Kristian Barman
Ghent University
KristianCampbell.GonzalezBarman@UGent.be

André Curtis-Trudel
University of Cincinnati

Mechanistic interpretability (MI) is an emerging research program in machine learning (ML) and AI safety where, rather than treating neural networks as black-boxes, researchers try to understand ML model behaviour by identifying internal components and assigning them specific causal and functional roles. MI practitioners borrow experimental and representational tools from the life sciences: they "zoom in" on internal structure, use model organisms such as small transformer language models, and test causal hypotheses via surgical interventions such as ablation, activation patching, and weight editing. This paper analyses case studies from MI to clarify what, in light of the 'virtual' character of the components under investigation, is genuinely mechanistic in its explanatory and causal claims, and what additional epistemic value, if any, is secured by such mechanistic talk.
We start from MI's own self-understanding. Its central claims presuppose that large neural networks contain discoverable and causally meaningful mechanisms (e.g. curve detectors, indirect-object-identification circuits); that these mechanisms are decomposable into smaller units (e.g., induction heads, features); that the resulting structures are in principle intelligible to human investigators; and that they generalise across architectures and training regimes. At the same time, the objects under study are high-dimensional, learned, and heavily distributed. The "parts" that MI identifies are software-localised patterns in activation and parameter space, not spatially discrete components exchanging energy in the familiar biological or engineering sense. Yet causal interventions on these patterns are systematically successful in controlling behaviour and are taken as central evidence in MI practice.
To make sense of this tension, we bring MI into dialogue with contemporary philosophical accounts of mechanisms and with work on modelling and representation in scientific practice. We examine how MI's experimental repertoire and representational devices function, and how mechanistic talk works within their practice. We then evaluate three broad interpretations of their mechanistic vocabulary. On a neomechanist reading, MI discovers genuine mechanisms composed of entities and activities organised across levels. On a

computational-mechanist reading, mechanisms are realised in medium-independent algorithms instantiated by a network. On a heuristic reading, mechanistic talk is merely a pragmatic scaffold for organising inquiry without robust ontological commitment.
We argue that none of these options fully captures the role mechanistic language plays in practice. A purely literal reading struggles with the substrate-independence and with MI's "parts"; a purely computational reading underplays the importance of local interventions and modularity; and a purely heuristic reading cannot explain why the field's methods yield stable, counterfactually robust handles on model behaviour. We develop a fourth, hybrid position: MI is best understood as operating with "fictional mechanisms", i.e., structured models that deliberately misdescribe the ontology of neural networks while nonetheless tracking real counterfactual dependencies via well-calibrated translational keys between representation and target. This framing sharpens questions about evidential standards (e.g. the "streetlight problem" and universality claims), about the relation between laboratory dissection and "in the wild" behaviour of deployed systems, and about what counts as understanding in a domain where mechanisms are virtual, repeatable, tightly coupled to sociotechnical environments, yet afford extremely fine-grained, surgical interventions.

Interface Ontologies for Collaborative Integrated HPS

Hakob Barseghyan
University of Toronto, Victoria College
hakob.barseghyan@utoronto.ca

Dubian Cañas
National Autonomous University of Mexico

Efforts to integrate the history and philosophy of science have typically presumed that such integration should be conducted by individual scholars. Exemplary figures (e.g. Chang, Schickore, Arabatzis) demonstrate how a single scholar can produce both historically sensitive philosophical reflection and philosophically informed historical reconstruction within the same work. However, this individualistic model presupposes unusually broad individual expertise and requires the same scholar to navigate archival study, historical narration, conceptual analysis, methodological reflection, and normative evaluation. This model of integration sets a standard that is possible yet difficult to meet. Consequently, it limits broader participation by scholars who are primarily trained as either historians or philosophers.
This paper suggests that the field requires a structural alternative – a collaborative model that allows for a division of labour where integration becomes a feature of the discipline rather than a burden carried by exceptional individuals. The core proposal is that integrated HPS should be supported by communally developed and curated ontologies that function as an interface across disciplines. An ontology (as understood in information science) is not a grand metaphysical system, but a structured vocabulary of entities, processes, and relations found in a certain domain. In our case, this could include epistemic entities (e.g., questions, theories, methods, norms), agents (e.g., individual, communal), practices (e.g., proposing, measuring, evaluating), stances (e.g., acceptance, use, pursuit), etc. By adopting such an ontology, historians can present their findings in forms that scholars from adjacent fields can readily understand and use, even if the historians themselves have no interest in potential theoretical outputs of their work. Conversely, using the same ontology, philosophers can

articulate their questions and theories (both descriptive and normative) that map cleanly onto historical datasets constructed without philosophical intent.
This approach focuses on institutional preconditions for fruitful integration. A shared ontology requires an institutional mechanism for its collective articulation, evaluation, and revision. Existing models from ontology engineering (e.g. the governance structures of CIDOC-CRM, Scientonomy, SDHSS) provide useful precedents. We propose that the Joint Commission establish a standing committee responsible for developing, maintaining, and modifying such ontologies tailored to integrated HPS. Proposed revisions should be documented, debated, and then accepted/rejected through formal communal procedures (e.g., a supermajority vote), ensuring that these ontologies evolve in an iterative, piecemeal, and transparent fashion.
Such an institutional infrastructure offers several benefits. First, it lowers the entry barrier for integrated work, without necessitating individual cross-training beyond the mastery of the communal ontologies. Second, it scales across research projects, making contributions from historians, philosophers, sociologists, and other scholars interoperable. Third, it provides a broad conceptual framework capable of accommodating diverse scientific cultures including non-Western and indigenous knowledge systems. Fourth, it could enhance HPS teaching by incorporating these shared ontologies in HPS courses. Fifth, shared ontologies supplement existing integrative methods by providing the conceptual architecture within which many forms of collaboration can flourish. In the future, besides shared ontologies, this approach could be developed to provide interfaces of shared questions, theoretical claims, and epistemic standards for HPS.

Overcoming Over Localization of Function: Insights from Circadian Biology

William Bechtel
University of California, San Diego
wbechtel@ucsd.edu

According to mechanists in philosophy of science, an initial step in explaining a phenomenon is localizing the responsible mechanism. Explanation then proceeds by decomposing the mechanism. This strategy has been enormously successful in revealing what processes inside mechanisms are responsible for producing various phenomena. But it is often the case that at the outset the phenomenon was over localized—many components outside the initially charactered mechanism play important roles in generating the phenomenon; in some cases the localized mechanism is not even necessary for the phenomenon. This talk focuses on how researchers discover that their initial localization was an over localization and that the phenomenon is due to a much larger system that overlaps mechanisms responsible for other phenomena. I will argue that it is often the same research strategies used to identify components within the mechanism that enable researchers to supersede their initial localization.
I use research on circadian rhythms to illustrate this process. Starting with a proposal in 1990, most circadian researchers concluded that circadian rhythms in all orders of life are due to a transcription-translation feedback loop (TTFL) in which the protein produced by a gene feeds back to inhibit its own transcription. Research over the next 20 years resulted in identifying many components involved in these TTFLs and provided details as to how they together generated ongoing oscillations.

Two lines of investigation, however, have pointed to this being an over localization. The first showed that a TTFL is not necessary for circadian rhythmicity. In 2005 researchers demonstrated that in cyanobacteria the proteins alone are sufficient. Many circadian biologists viewed this finding as something limited to bacteria, but in 2010 O'Neill demonstrated circadian rhythms in red blood cells that, lacking a nucleus, cannot engage in transcription or translation. These together with other findings in neurospora suggested that the TTFL is not essential for producing circadian rhythms.
The second line of investigation began by seeking to determine the output processes by which the TTFL could regulate metabolism. As metabolic targets were identified, researchers quickly realized that these metabolic processes also acted on the TTFL. Moreover, these components were themselves rhythmic and altering their rhythms could affect key properties of TTFL rhythms such as period and amplitude.
For the most part circadian researchers still characterize TTFLs as the clock mechanisms, viewing them as situated in a larger network. But some researchers influenced by the first line of investigation are exploring other sources of rhythmicity in the metabolic system. Given the numerous feedback loops, there are many potential sources of sustained oscillation. Moreover, there is an evolutionary premium on coupling oscillating conditions in metabolic systems with the light-dark cycle in the environment. Clocking is made possible by multiple feedback loops within the larger metabolic system. A TTFL is important but only one component that supports clocking.
Seeking a mechanistic locus for a phenomenon plays an important heuristic role in enabling inquiry. But the very research that helps to unveil that mechanism can demonstrate that the initially localized mechanism is only one component, and not even a necessary component, in a larger system that generates the phenomenon.

Laboratory Learning: Epistemic Virtues and Teaching Practices in 19th-Century Experimental Chemistry

Pieter T. L. Beck
Ghent University, Department of Philosophy and Moral Science
pieter.beck@ugent.be

Historiographical accounts have repeatedly called attention to the importance of pedagogical innovations in experimental chemistry during the 19th century. A philosophical analysis of the epistemic impact of these innovations however remains lacking. In this paper, I argue that virtue epistemology provides a philosophical framework which shows why teaching practices are epistemically relevant and should thus be of interest to philosophers of science. A general characteristic of virtue epistemology is the focus on the epistemic agent (and sometimes community). This is often described as a "reversal of the direction of analysis": a belief's or epistemic action's status is appraised based on the properties of the knower, not the other way around. This implies that knowledge production is dependent on the epistemic agent and her possession of certain epistemic virtues. These virtues are not abstract entities. In scientific contexts, they are acquired and embodied by agents after long training. A virtue epistemic perspective thus shows that we need to include teaching practices in an analysis of the way scientific knowledge is produced and justified. Following a practice-based iHPS approach, I argue that such an analysis cannot be made in the abstract but should consist in an integrated historical and philosophical investigation. Experimental chemistry in the 19th century is an optimal target for such an integrated investigation. The 19th century was a period in which

the connection between research and teaching was institutionalised in universities Chemistry played an important role in the emergence of the university laboratory as a nexus of teaching and research. Justus Liebig is generally seen as an essential figure for this development. At the university of Giessen, Liebig developed a new model of chemistry training, which emphasised lab work. Although he was not the first to introduce laboratory training, what made Liebig's model successful was the way in which a series of experiments of increasing complexity prepared students to pursue their own research under his supervision. This led to the development of a research school producing a steady output of results. Liebig's pedagogical innovation was closely connected to his development of the "Kali-apparat" for measuring the carbon content of an organic compound. It made the analysis easier, quicker, but also more precise. The "Giessen model" soon spread across Europe. Chemists travelled to Giessen to study with Liebig. Many of his students would take up teaching positions, especially in Britain. I analyse Liebig's pedagogical innovations, showing how they helped students to acquire the relevant epistemic virtues. The spread of the Giessen model in turn lead to the emergence of an international community of chemists that was able to trust and build upon each others' results, thus providing a clear example of how teaching practices were part and parcel of the production and justification of scientific knowledge.

What Are We Measuring When We Measure Age? Accuracy and Actionability in the Clinical Use of Biological Age Clocks.

Alessandro Blasimme
ETH Zürich
alessandro.blasimme@hest.ethz.ch

Oliver Buchholz
ETH Zürich
oliver.buchholz@hest.ethz.ch

Chronological age (CA), while universally used, provides only a coarse proxy for the heterogeneous biological processes that constitute human aging. Individuals with the same CA often display markedly different physiological states, functional capacities, and health risks, revealing a conceptual and practical gap between chronological time and aging as a biological phenomenon. In the last 15 years, Biological Age (BA) clocks have emerged as a way to capture individual variation in the way the passing of time affects biological characteristics, functional capacities, and health. This paper first analyzes the conceptual landscape of BA as a family of measurement practices rather than a singular entity. Over the past two decades, BA clocks have evolved through distinct methodological generations. First-generation clocks used DNA methylation or proteomic data to reproduce CA, treating biological signals instrumentally as predictors. Second-generation clocks shifted the target to health outcomes such as morbidity or mortality, embedding normative and medical assumptions directly into model training. Third-generation clocks use longitudinal datasets to estimate the rate of biological aging, enabling dynamic modeling of individual aging trajectories and organ-specific divergence. These developments illustrate how operational definitions of BA arise from the interplay of available technologies, modeling strategies, and epistemic aims, resulting in multiple coexisting "ages" rather than a unified parameter. We then explore how BA and CA function as classificatory tools in practice. Across primary care, public health, and allocation infrastructures, CA serves as a simple, low-cost trigger for

eligibility (e.g., screening thresholds) or for stratifying risk. BA offers finer-grained, individualized predictions, often outperforming CA in forecasting mortality, disease burden, or functional decline. We further consider how BA intervenes in existing infrastructures of evidence. Introducing BA into medical decision-making requires integrating heterogeneous data types, negotiating standards of accuracy, and reconciling conflicting signals from different clocks. Issues of validation and calibration become central: BA clocks often rely on surrogate endpoints, embed modeling choices that are not theory-neutral, and may lack commensurability. This situates BA within broader debates about machine-learning-based biomarkers, their interpretability, and their ability to track underlying biological processes rather than merely correlate with them. Moreover, BA not only reflects physiological processes but also captures the biological imprint of social and environmental exposures, raising further questions about the epistemological status of BA as a trait combining molecular, developmental, and socio-environmental determinants. Its practical uptake will depend on how researchers and clinicians negotiate the tension between BA's scientific promise and the complexity of standardization, interpretability, and infrastructural integration. We conclude that BA is best understood not as a refinement for CA but as a locus where scientific models, measurement technologies, and practical decision systems intersect. Studying BA thus provides a window into how contemporary biomedicine constructs new kinds of temporal and biological categories—and how these categories reshape our understanding of aging in practice.

Rethinking Interactional Expertise in Gravitational Wave Science

T. K. Blesil
Department of Studies in Science, Technology and Innovation Policy School of Social Sciences Central University of Gujarat
blessilkunju@gmail.com

Harry Collins and Robert Evans developed the concept of Interactional Expertise (IE) through decades of engagement with the experimental science of gravitational waves. They argued that IE constitutes a distinct form of expertise gained through prolonged linguistic immersion in a specialist community, allowing actors, such as sociologists, theorists, or adjacent scientists, to master technical discourse without making direct material contributions to experimental practice. This form of expertise, they contend, is analytically separable from Contributory Expertise and is central to the functioning of complex scientific enterprises. Gravitational wave science provides a particularly fertile site for examining IE. Predicted by general relativity in 1915, gravitational waves were experimentally detected in 2015, following more than five decades of sustained effort that culminated in the Laser Interferometer Gravitational-Wave Observatory (LIGO). Since the 1980s, the field has transformed from small-scale laboratory experiments into a large, globally distributed mega-science project. This transformation has produced multiple highly specialized subfields—post-Newtonian approximations, numerical relativity, data analysis, interferometry, and laser physics, each with distinct yet overlapping technical languages. The coordination of these heterogeneous epistemic cultures raises a central question: how is Interactional Expertise sustained across disciplinary boundaries in such an expansive and linguistically complex scientific enterprise?
This paper addresses this question through a mixed-methods approach grounded in participant comprehension. Empirically, it draws on five months of ethnographic fieldwork at

the Inter-University Centre for Astronomy and Astrophysics (IUCAA), Pune, where I participated in and observed a tabletop interferometer apprenticeship programme designed to train MSc and MTech students for the forthcoming LIGO–India project. This was complemented by participation in a summer school on gravitational wave astronomy at the International Centre for Theoretical Sciences, Bangalore, which offered insight into how theoretical knowledge is communicated to early-career researchers. In addition, semi-structured interviews were conducted with gravitational wave physicists from India, the United States, Europe, Australia, and Japan, enabling a comparative perspective on interdisciplinary communication and language formation.

Analytically, the paper examines three interrelated dimensions of expertise in gravitational wave science. First, it traces the historical evolution of the field's technical languages since the 1980s, showing how increasing specialization has intensified linguistic fragmentation. Second, it revisits the theoretical foundations of Interactional Expertise, identifying methodological tensions that prompted Collins and Evans to introduce the notion of Special Interactional Expertise. Third, it demonstrates how physicists routinely encounter limits in understanding neighboring subfields, despite extensive collaboration, revealing the practical constraints of linguistic immersion alone.

The paper concludes that IE was indispensable to LIGO's success, enabling coordination across epistemically diverse communities. However, it also offers a critical reassessment of Collins and Evans by arguing that language, by itself, may be insufficient to sustain large-scale, long-term experimental programmes. The increasing compartmentalization of knowledge within gravitational wave science suggests that not all contributors possess, or need, comprehensive expertise. While this strategic distribution of expertise facilitates collaboration, it also points to a potential intellectual vulnerability in mega-science projects, where scientific success depends less on shared understanding than on carefully managed epistemic division.

Nancy Cartwright's Philosophy of Quantum Mechanics in Practice

Matthew Brown
Southern Illinois University
matt.brown@siu.edu

In this presentation, I revisit and reassess Nancy Cartwright's interpretation of quantum mechanics as articulated in Chapter 9 of How the Laws of Physics Lie, "How the Measurement Problem is an Artefact of the Mathematics" (1983). I argue that her approach is an under-appreciated alternative in the foundations of quantum physics that still has much to recommend it. Cartwright interprets quantum mechanics as a theory of nondeterministic transition processes: what quantum mechanics does is determine the probability that systems that begin in a specified state will transition to another specified state, under certain conditions typically determined by the environment of the system.

I argue that, despite how she frames her project as "a kind of anti-realism" (p. 2), Cartwright's interpretation is best understood as a realist interpretation of quantum mechanics. Her account eschews vague references to measurement or inappropriate dependence of the microphysical on consciousness, observation, or minds. Cartwright's approach also more faithfully reflects the practice of physics than other interpretations, which begin from abstract issues of theory rather than how the theory is used in practice. It is only with respect to a peculiar, abstract-theoretical presentation of the theory of quantum

mechanics that Cartwright's view seems "anti-realist." But insisting on fidelity to the role various elements of the theory play in practice is not anti-realistic; it is a better form of realism.

Although the "practice turn" in philosophy of science has definitely come to parts of the philosophy of physics, there has been surprisingly little work in the foundations of quantum physics that centers quantum mechanical practice. While the focus on how the theory is used in practice was a central focus of Niels Bohr's interpretations of quantum mechanics, most philosophers of physics have focused on an abstract, axiomatic understanding of the theory. As David Wallace (2019) has recently shown, this understanding is rather at odds with the practice of quantum physicists. I show that Cartwright's interpretation is more suited to the picture of quantum physics in practice that Wallace presents, and more in line with the approach of physicists like Bohr and David Finkelstein (1999), whose interpretations of the theory were closely tied to their understanding of physics practice.

I will also argue that this is all consistent with Cartwright's (2005) later argument that quantum mechanics requires no "interpretation." Cartwright's interpretation is not an "interpretation" in the sense that she criticizes, and this is another way that it improves on the standard options.

References

Cartwright, Nancy. 1983. How the Laws of Physics Lie. Oxford University Press, USA.

Cartwright, Nancy. 2005. "Another Philosopher Looks at Quantum Mechanics, or What Quantum Theory Is Not." In Hilary Putnam, edited by Yemima Ben-Menahem, 188–202. Contemporary Philosophy in Focus. Cambridge University Press.

Finkelstein, D. R. 1999. "Action Physics." International Journal of Theoretical Physics 38 (1): 447–55.

Wallace, David. 2019. "What Is Orthodox Quantum Mechanics?" In Philosophers Look at Quantum Mechanics, edited by Alberto Cordero, 285–312. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-15659-6_17.

When Diagnosis Is Not Classification: Clinical Uncertainty in Rheumatology

Zdenka Brzovic
University of Rijeka
zbrzovic@gmail.com

In medicine, diagnostic criteria are used to determine whether an individual patient has a disease, whereas classificatory criteria are used to define patient groups for research. Ideally, the same criteria would serve both clinical and research purposes (see Solomon 2022). In infectious disease, for example, a positive blood culture can both establish a diagnosis in an individual patient and reliably define a research population, allowing clinical and research uses of disease concepts to align. In rheumatology, however, this alignment often breaks down. A familiar example is rheumatoid arthritis, where up to 30% of patients lack the usual blood markers despite clear clinical disease (see June and Aggarwal 2014). This diagnostic uncertainty has prompted the development of multiple sets of disease classification criteria for research on disease characterisation, epidemiology, prognosis, and the design of clinical trials. Classification criteria consist of disease features used to group individuals into

relatively homogeneous populations for research, rather than to guide diagnosis in individual patients.

This paper examines how to understand the difference between clinical judgement in individual cases and the grouping of patients for research. One initial explanation appeals to the context-sensitive nature of diagnostic reasoning. Diagnostic judgements are made by particular physicians or care teams, at specific points in time, and integrate clinical findings with background information about disease prevalence, referral context, geographic location, and alternative diagnoses. However, context sensitivity alone does not fully explain the divergence, since population-based data can still be applied to individual cases. To further elucidate the distinction in question, the paper turns to the statistical principles governing how criteria are constructed and evaluated in practice, and how these measures are weighted differently in diagnosis and classification.

From a statistical perspective, diagnostic criteria are generally broad and designed to maximise sensitivity and negative predictive value, minimising the risk of excluding individuals who may genuinely have the disease, including those with atypical or evolving presentations. Classification criteria, by contrast, prioritise high specificity and positive predictive value in order to assemble homogeneous groups for research. I argue that this asymmetry is best explained by pragmatic considerations. Diagnostic judgements are action-guiding: they shape treatment decisions, monitoring strategies, and prognostic discussions. In this context, it appears that the costs of false negatives often outweigh those of provisional or revisable false positives. Including more patients within the diagnostic net is therefore a response to the practical demands of clinical care under uncertainty. Some relevant practical considerations include the costs of diagnostic tests, issues of insurance coverage, and even the potential psychological benefits of receiving a diagnosis "with a name" (see June and Aggarwal 2015, Lockshin, Crowand and Barbhaiya 2022).

References

June, Rayford R., and Rohit Aggarwal. 2015. "The Use and Abuse of Diagnostic/Classification Criteria." Arthritis & Rheumatology 67 (7): 1769–1771. https://doi.org/10.1002/art.39191

Lockshin, Michael D., Mary K. Crow, and Medha Barbhaiya, eds. Diagnoses Without Names: Challenges for Medical Care, Research, and Policy. Cham: Springer, 2022. https://doi.org/10.1007/978-3-031-04935-4

Solomon, Miriam. 2022. "A Pragmatic Approach to Diagnostic Categorization." In Diagnoses Without Names: Challenges for Medical Care, Research, and Policy, edited by Michael D. Lockshin, Mary K. Crow, and Medha Barbhaiya, 11–17. Cham: Springer. https://doi.org/10.1007/978-3-031-04935-4_3

Yield: Measurement, Values, and the Aims of Agricultural Science

Julia Bursten
University of Kentucky
jrbursten@uky.edu

Catherine Kendig
Michigan State University
kendig@msu.edu

Jennifer Jhun
Duke University
jennifer.jhun@duke.edu

"Yield" is a term of art in agricultural sciences and farming practices. In its most basic form, it refers to an evaluation of the quantity produced by a given crop. Increasing yield at sustainable cost is both an economic aim of farming and a research frame of agricultural science. Many projects in agricultural science aim at improving commercial yield by developing new cropping practices such as targeted pesticide application, minimizing the frequency of tilling, making planting decisions that aim to increase biodiversity or breeding crop species with traits such as drought or disease resistance in order to minimize loss of crops due to extreme weather events or pest predation.

Given the centrality of yield as a framing concept within both farming and agricultural research practice, we explain why it is a concept worthy of philosophical investigation. Addressing central epistemic questions about yield, such as how yield is conceptualized and measured, and whether yield measurements are effective and reliable, requires examination of the activities of farmers and agronomists. Using recent participatory plant breeding research, we argue that agricultural research's own conception of yield has systematically omitted key culinary practices which when included change both the concept of yield and the articulation of appropriate and effective research questions. Centrally, what counts as the valuable portion of the overall biomass yield of a crop, and who evaluates the yield, turn out to be deeply value-laden questions.

The case considers participatory sorghum variety evaluations in Mali. It examines the ways in which culinary tests have been run and how they might be improved in order to better reconcile the identification of trait-specific genetic varieties that are developed with the cultural and culinary preferences of communities for whom the grains are used. Sorghum is a grain and a main staple crop in Mali. Processing sorghum into food requires significant effort. To transform the harvested grain into the staple dish, tô (or toh), stalks of seed grain must be harvested, dried, removed from the stalk, decorticated, pounded with a mortar and pestle, sieved, and cooked with water and other ingredients over coals. Researchers observed that "improved" plant species, which yield more grain per plant and more disease- and pest-resistant dried grains, are sometimes rejected by communities due to producing less desirable tô. They developed culinary testing procedures to evaluate yield in terms of quality and quantity of cooked food or, in particular the number of tô meals that they were able to prepare with the harvest, and they argue in favor of framing further research in terms of "food yield," a concept identified by the Malian communities as a preferable alternative to standard agronomic yield measures.

In addition to the preconception of yield as useable food, the case highlights gender roles in food production: in the participant communities, harvesting is a co-ed practice while milling and preparing tô is women's work. Adopting culinary testing methods for yield measurement includes whole communities in yield conversations and recognizes the inclusion of women's key role in food yield.

Reference:
Isaacs, K., Smit, M., Samaké, B., Rattunde, F., Cissé, F., Diallo, A., Sidibe, M., & Weltzien, E. (2023). Participatory Evaluation of Sorghum Processing and Sensory Attributes in Mali: Methodology for Improving Food Security Outcomes from Variety Development Efforts. Sustainability, 15(5), 4312. https://doi.org/10.3390/su15054312

Addressing Artificialisation in Scientific Practice

Stefano Canali
Politecnico di Milano
stefano.canali@polimi.it

In this talk I discuss an emerging application of generative artificial intelligence and machine learning in the sciences – synthetic data –, frame it as the extension of surrogative forms of reasoning and discuss their possible radicalisation into a new phenomenon we need to address in philosophy of science in practice – artificialisation.
Synthetic data are data generated with the goal of resembling the characteristics and structure of other data, mostly through machine learning and generative artificial intelligence tools capable of producing synthetic images, text, and videos. In the sciences, synthetic data are increasingly presented as a way to fill in epistemological and as well as ethical gaps. For instance, in the health context increasing contributions present synthetic data as fixes to critical issue of health datasets and science such as bias, lack of representativity, high costs. There are various epistemological implications of synthetic data, but in this talk I will focus on a more general framing: on the basis of a pilot study of synthetic data in scientific practice for electronic health records, I argue that we should frame these and other applications of generative artificial intelligence and machine learning as an extension of surrogative forms of reasoning in science. Surrogative reasoning has received much attention in recent philosophy of science in practice. For instance, the scientific study of health has a long history of using technologies and tools that mediate between scientific interests and world phenomena and work as surrogative systems that are more directly accessible or easier to manipulate experimentally than the target systems under investigation. Prototypical examples of these artificial representations of reality are scientific models, such as animals used to make inferences about the efficacy of pharmaceutical drugs or agent-based simulations to predict the spreading of infectious disease. Data play a similar role: they are mediated ways of looking at reality, epistemic artefacts that reflect human interests, perspectives, and values. Synthetic data give us a clear opportunity to study new tools for surrogate reasoning, but there is more. Synthetic data push towards a new radicalisation of surrogative reasoning in scientific practice that I call ‘ artificialisation’, an extension of surrogative reasoning where generative artificial intelligence and machine learning are used to study reality indirectly and with indirect tools and technologies. Synthetic data give us an opportunity to study artificialisation as a possibly central process of contemporary scientific practice and discuss it from both epistemic and normative viewpoints in philosophy of science in practice. At a descriptive level, we need to understand if artificialisation presents different implications and changes with respect to surrogative reasoning and which tools we have in philosophy of science in practice to study more artificial forms of scientific practice. Getting to a normative level, we need to understand if artificialisation produces recursive scientific forms of practices and how to address the decoupling of data generation from empirical observation, with the risk of generating ‘artificious’ science.

The Complex Interplay Of Early Modern Science And Technological Practices In The Invention Of Newcomen's Engine

Manjari Chakrabarty
Visva-Bharati(A Central University)
manjari.chakraborty@visva-bharati.ac.in

Historical-philosophical studies investigating the character of science-technology interactions – one of the oldest and most fundamental problems in the field – have generally been dominated by the "reductionist view" (Dasgupta, 1996), indicating that technology entails the mere application of scientific knowledge to the making of artefacts. Scholars (e.g., Kitsikopoulos, 2013) who challenge the technology-as-mere-applied-science-reductionist-view often cite how Thomas Newcomen developed the first fully self-acting steam or atmospheric engine primarily in ignorance of the laws of thermodynamics. Once the early steam engines made it possible to obtain work from heat through the medium of steam, they gave a considerable impetus to the development of thermodynamics. Hence goes the old historical saying that science owes more to the steam engine than the other way around. Nevertheless, the connection between the existing stock of scientific knowledge, precisely, the discoveries of the phenomenon of atmospheric pressure by Galileo's pupil Evangelista Torricelli and that of the power of atmospheric pressure over a vacuum by Otto von Guericke, and early steam engine technology is too strong to be ignored. A closer scrutiny of the pre-natal history of Newcomen's engine during the early 1700s reveals a complex interplay between the science of pneumatics (not thermodynamics) and the technological practices related to early steam engines.

Challenging the received 'reductionist view' while acknowledging the indispensability of the science of pneumatics for the Newcomen engine, we aim to dissect the complex interplay of early modern science and technological practices in Newcomen's invention of the steam engine. Our objective in the first section of the paper is to investigate the interconnection between the science of pneumatics and the technological practices related to early steam engines. The function of the science of pneumatics in early steam engine technology, as our review of the relevant literature suggests, was not constructive, i.e., it did not drive the innovation of the Newcomen engine.

The question raised by the foregoing analysis is considered in the final section of the paper: what then was the real potential (if not constructive) of early modern science in steam-engine technology? Recalling a point made by Karl Popper (1944) long ago, we realize that the technological implications inherent in scientific conjectures come into sight only when such conjectures are formulated negatively, in the form of prohibitions. The technological implications of, say, Otto von Guericke's scientific conjecture of the effect of atmospheric pressure over a vacuum become apparent and directly applicable for steam-engine technology when stated negatively as: One cannot prevent the piston (fitted into a cylinder) from being sucked into an evacuated cylinder against the pressure of the atmosphere.

We conclude by emphasizing what historical-philosophical writings on science-technology interactions have rarely noted: the critical (not constructive) potential of science in early steam-engine technology.

References

1. Dasgupta, S. (1996). Technology and Creativity. New York: Oxford University Press.

2. Kitsikopoulos, H. (2013). From Hero to Newcomen: The Critical Scientific and Technological Developments that led to the Invention of the Steam Engine. Proceedings of the American Philosophical Society, 157 (3), pp. 304–344.
3. Popper, K. R. (1944). `The Poverty of Historicism II'. Economica ns XI, 43, pp. 119–137.

Niche Construction Theory as an Explanation-Seeking Strategy in Sociology

Yihui Cheng
Ghent University
yihui.cheng@ugent.be

Niche construction theory (NCT) was originally developed in evolutionary biology to highlight the idea that organisms not only adapt to their environments, but also actively or passively modify those environments in ways that feed back into their own evolution (Odling-Smee et al. 2003). Given that humans might be 'the most creative niche constructors', NCT is particularly promising for studying how human behaviors transform and are transformed by their environments, including social environments.

In fact, a growing body of social science literature has directly or indirectly employed NCT or its variants to explain human social phenomena, such as the origins of agriculture and the formation of markets. Advocates of using NCT in the social sciences argue that it offers a novel theoretical perspective and yields testable explanations (e.g. Laland and O'Brien 2010). Critics, however, contend that NCT fails to provide substantial explanations and new insights (e.g. Spengler 2021). These evaluations, largely from archaeological studies, often extend debates originating in evolutionary biology, leaving it unclear whether NCT's explanatory power extends beyond these contexts or how far such evaluations apply to other social sciences.

Here, we focus on sociology. A preliminary assessment by sociologist Tang (2025) highlights NCT's explanatory potential by arguing that it brings time and space into sociological theorizing. Within philosophy of science, however, little attention has been paid to evaluating how NCT functions within social explanations. This paper addresses the gap by evaluating the explanatory usefulness of NCT in sociological studies.

We first engage with existing evaluations of NCT in biology and archaeology to see how these can be brought into a sociological context. We show that existing evaluations tend to fall into two categories: NCT is either a full-fledged theory from which explanations automatically follow, or a mere relabeling of familiar processes. By contrast, we propose treating NCT as an explanation-seeking strategy: a form of scaffolding that facilitates inquiry and reasoning. Given this role, NCT is particularly useful for explaining cases that clearly show a coevolving process between agents and social environments.

To illustrate this point, we consider a contemporary sociological case. Scholars studying social media often describe influencers' practices to gain visibility as "gaming the system". This narrative implies influencers merely adapt to platform rules. Cotter (2019), however, argues that these practices are better understood as "playing the visibility game," a reframing that emphasizes the interdependence between influencers and platforms, rather than a unilateral view of platform control. We suggest this narrative shift reflects explanatory scaffolding derived from NCT, directing sociologists to conceptualize influencer-platform relations as an environment-behavior feedback loop. In this sense, NCT functions as an explanation-seeking strategy that facilitates sociologists' inquiry and reasoning.

References

Cotter, K. 2019. Playing the Visibility Game: How Digital Influencers and Algorithms Negotiate Influence on Instagram. New Media & Society 21(4): 895–913.
Laland, K. N., and O'Brien, M. J. 2010. Niche Construction Theory and Archaeology. Journal of Archaeological Method and Theory 17(4): 303–22.
Odling-Smee, J., Laland, K. N., and Feldman, M. W. 2003. Niche Construction: The Neglected Process in Evolution. Princeton University Press.
Spengler, R. N. 2021. Niche Construction Theory in Archaeology: A Critical Review. Journal of Archaeological Method and Theory 28(3): 925–55.
Tang, S. 2025. Niche Construction to Social Sciences: Can It Help?. Theory and Society 54(2): 209–41.

Evolution's Bonus, or the Many Ways to be Biased about Female Orgasm

Christopher ChoGlueck
New Mexico Tech
christopher.choglueck@nmt.edu

Evaluating various hypotheses for the evolution of female orgasm, Elisabeth Lloyd (2005) found that nearly every theory was unduly influenced by androcentrism, adaptationism, or both. She contended that the only empirically adequate explanation was the byproduct account, which postulated that female orgasm was not an adaptation per se, but rather an "embryological bonus" for female primates and (cisgender) human women (p. 110). Lloyd maintains that her account is genuinely feminist, in line with Helen Longino's (1990, p. 188) practice-based approach of "doing science as a feminist."

Nonetheless, the byproduct account has been criticized as "androcentric" for normalizing sexual agency as masculine and sexual passivity as feminine (Gannett, 2007; Fausto-Sterling et al., 1997). In his original formation, Donald Symons (1979) troublingly defines orgasm as a male-typical behavior, rendering female orgasm an anomaly (Lee, 2013). Accordingly, one might wonder if Lloyd's account has inherited any latent androcentric bias. Wakil (2021) responds in the affirmative: "Symons's definition of homology—by claiming orgasm is a distinctly male trait and using male orgasm as the proxy for characterizing female orgasm—is dripping with the androcentrism Lloyd defines as 'not treating female sexuality as autonomous from male sexuality or male reproduction'…and 'assuming females' response is like male's response'" (p. 2320).

Defending Lloyd against these critics, this paper makes two arguments related to the values embedded in her account. First, I contend that, while Symons and Lloyd both avoid adaptationism, only Lloyd's account evades androcentrism. Unlike Symons, Lloyd does not gender the standard of human sexual development as masculine. Accordingly, I recommend a distinction between Symons's androcentric byproduct account and Lloyd's "bonus" alternative.

Second, I argue that Lloyd's account supports feminist values, not only because it avoids pathologizing female sexuality, but also because her work has guided various research communities toward knowledge that is sexually liberating. I do so by situating Lloyd's bonus account among other non-androcentric explanations of female orgasm (including Sherfey, Hrdy, and Gould). My analysis contrasts different influences of androcentric values guiding evolutionary theorizing in terms of naturalizing vs. normalizing. The androcentrism in the

adaptation-based explanations (including Freud and Morris) involves naturalizing gender hierarchies. Other non-adaptationist explanations like Symons are nonetheless androcentric in so far as they pathologize female development as an arrested version of "normal" male development.

Works cited:

Fausto-Sterling, A., Gowaty, P. A., & Zuk, M. (1997). Evolutionary Psychology and Darwinian Feminism. Feminist Studies, 23(2), 403–417.
Gannett, L. (2007). Elisabeth A. Lloyd The Case of the Female Orgasm: Bias in the Science of Evolution. Canadian Journal of Philosophy, 37(4), 619–638.
Lee, D. J. (2013). Homology, female orgasm and the forgotten argument of Donald Symons. Biology & Philosophy, 28(6), 1021–1027.
Lloyd, E. A. (2005). The Case of the Female Orgasm: Bias in the Science of Evolution. Harvard University Press.
Longino, H. E. (1990). Science as Social Knowledge. Princeton: Princeton University Press.
Wakil, S. (2021). Objectivity and orgasm: The perils of imprecise definitions. Synthese, 199(1), 2315–2333.

Epistemic rights as a framework for promoting integrity in scientific research

Noa Cohen
Tuscia University
noa.cohen4@gmail.com

Contemporary society confronts a profound epistemic crisis (Friedman, 2023), marked by proliferating unreliable information, diminishing trust in authority, and an eroded notion of objective truth. This information environment, described as epistemically polluted or degraded (Blake-Turner, 2020; Levy, 2021), demands a critical re-evaluation of the conditions of scientific knowledge production and its societal implications. The epistemic environment is changing; science is less isolated and increasingly tightly coupled with media, politics, and commerce, leading to a loss of the historical distance that supported its trustworthiness. Media coverage often prioritizes sensationalism or simplicity over epistemic rigor, shaping public perception of scientific findings (Intemann, 2022). Concurrently, problematic practices underlying scientific research are prevalent and include a widespread reproducibility crisis, flaws in the peer-review system, funding incentives, and cognitive biases affecting both scientists and the public.
Authors such as Elizabeth Anderson (2011) emphasize the role of epistemic authorities in shaping this climate while in addition to others such as Neil Levy (2018) and Regina Rini (2023) consider the insurmountable burden placed on individuals attempting to navigate it for making informed and beneficial decisions, often to little avail. In this environment, teeming with information of unverified sources and reliability, the need for trustworthy knowledge mediation is clear and urgent. Such scholars also point to the tension between individual and social accountability for attaining correct information, investigating what role the environment plays in denying individuals the realization of epistemically responsible beliefs. Purdue Pharma's involvement in driving the opioid crisis in the United States through the propagation of false scientific findings and medical recommendations, exemplifies the dangers of misplaced trust. This case is used by Lani Watson (2021) as a basis for her

advancement of a new class of rights - epistemic rights - and the discussion of the type of actions which amount to their violation. Exploiting the credibility allowance granted to institutions of epistemic authority for financial incentives is clearly a severe form of epistemic harm directly affecting individuals through the provision of inaccurate information. In the case of Purdue Pharma, their manifest violations may also fall under existing legal protections which range from privacy to criminal law; however, it raises the need to identify duties and protections required by knowledge and information practices, especially those overlooked in legal practice.
In bridging the notions of epistemic rights and scientific epistemic authority, this paper addresses the idea of epistemic obligations at an individual and institutional level, framing it within contemporary epistemological perspectives. The specific role of scientific institutions in this context emerges as one both affected-by and responsible-for, holding that advancing relevant awareness and regulation for securing epistemic rights can at once restore their authority and societal role as truth-keepers, as well as support scientific practices in restoring their internal integrity.

References

Anderson, E. (2011). Democracy, Public Policy, and Lay Assessments of Scientific Testimony. Episteme, 8(2).
Blake-Turner, C. (2020). Fake news, relevant alternatives, and the degradation of our epistemic environment. Inquiry (United Kingdom).
Buzzell, A., & Rini, R. (2023). Doing your own research and other impossible acts of epistemic superheroism. Philosophical Psychology, 36(5), 906–930.
Friedman, J. (2023). Post-Truth and the Epistemological Crisis. In Critical Review (Vol. 35, Issues 1–2).
Intemann, K. (2022). Understanding the Problem of Hype: Exaggeration, Values, and Trust in Science. Canadian Journal of Philosophy, 52(3).
Levy, N. (2021). Bad Beliefs: Why They Happen to Good People. Oxford University Press.
Watson, L. (2021). The Right to Know: Epistemic Rights and Why We Need Them. Routledge.

Active Knowledge for Realists

Alberto Cordero
CUNY Graduate Center & Queens College CUNY
acordelec@outlook.com

Over the last two decades, various strands of the pragmatist tradition have gained prominence in the discussion of scientific realism. One feature shared by many of the resulting positions is the abandonment of the idea that successful sciences reveal the "fundamental reality" of the domains they study.
A spirited expression of this new turn is the "active knowledge" realism advocated by Hasok Chang in his book Realism for Realistic People (2022). Active knowledge (AK) seeks to overcome the excesses of the leading camps in the debate on scientific realism, particularly their entrenched focus on the relationship between theoretical representation and the external world at the expense of scientific practice among proponents of realism, and the opposing focus on the social construction of reality at the expense of the external world among critics

of scientific realism. Invoking a perspective of "operational coherence," Chang does not conceive of truth in terms of its metaphysical correspondence with a world external to our thinking and representation, but rather assigns it to specific domains, without claiming universality, and it is evaluated gradually. In Chang's view, a scientific proposition is true "to the extent that there are operationally coherent activities that can be performed by relying on it" (p. 167). Operationally coherent theories are true within their own domain of validity.
The AK proposal outlined above aligns well with many aspects of scientific practice, but it overlooks some arguably crucial ones. Objections to the pragmatist purge proposed by AK include the following: (1) It restricts the potential for deep scientific representation too early. (2) It endorses an ontological pluralism that overlooks or diminishes important differences in epistemic merit among various ontologies. (3) It does not recognize the pragmatic value of scientists' ingrained tendency to pursue deeper, non-utilitarian descriptions of their domains of interest.
This proposed paper clarifies the aforementioned claims and proposes a way to optimize the valuable notions of truth and scientific success offered by Chang. It does so by revisiting the distinction between the fundamental theoretical representations emphasized by classical realism and the restriction of the notion of true descriptions to subdomains (or specific regimes). These latter descriptions have the character of the functional descriptions highlighted in the last decade by selective realists as descriptions suitable for realist interpretation (e.g., ., Alai 2017, Cordero 2017, Egg 2021). A central feature of such descriptions is the explicit specification of their respective levels of abstraction (focused on the aspects considered relevant for describing the empirical domain at the chosen level); their restriction to empirical domains within which the proposed description effectively holds; their explicit coarse-graining; and the face-value ontology of the theories at play. In this regard, a central development in recent selective realism literature concerns the characterization of descriptions suitable for realist interpretation—in particular, their level of abstraction (focused on aspects considered relevant for describing the empirical domain at the chosen level), their explicit restriction to particular empirical domains or "regimes" where the description effectively holds, their explicit coarse-graining, and the face-value ontology of the theory at play. This development is at the heart of the proposed paper. The suggestion offered shares the non-fundamentalism of selective realism but maintains the epistemological and pragmatic value of theoretical projects aimed at ontological deepening and differentiating the epistemic merit of available ontologies.

When Anomalies Become Possibilities: Serendipity and the Co-Emergence of New Affordances in Scientific Discovery

Matteo Costa
Sapienza University of Rome
matteo.costa@uniroma1.it

The philosophy and history of science have paid insufficient attention to a crucial intersection concerning both the dynamics of discovery and the role of tools in scientific research: the relationship between serendipitous discoveries and the emergence of new affordances. Despite sharing significant characteristics, serendipity and affordance have been consistently analyzed as distinct and unrelated phenomena in their respective literatures. This separation, however, obscures a fundamental insight into the nature of scientific discovery itself. Serendipity refers to discoveries occurring 'by chance' when researchers seeking one thing

encounter unexpected observations that prove significant and valuable. The literature emphasizes that beyond the fortuitous element, serendipity requires what Pasteur called a 'prepared mind,' along with specific theoretical, technological, and social contexts enabling anomalies to be recognized as meaningful. Affordances, introduced by James Gibson in ecological psychology, designate the possibilities for action that environments and objects offer to organisms – emerging from dynamic interactions between environmental features and organismic abilities rather than existing as purely subjective or objective properties. Through paradigmatic cases – including Jocelyn Bell's discovery of pulsars and Horace Wells's recognition of nitrous oxide's anesthetic properties – we demonstrate that these concepts share three fundamental characteristics: they are ecological (depending on organism-environment complementarity), situated (embedded within specific embodied, cultural, and historical contexts), and relational (constituted through networks of agents, artifacts, and environments rather than existing as intrinsic properties). Most significantly, we argue that when serendipitous discovery processes arise, they entail the actualization of new affordances that co-emerge and co-evolve with the new knowledge deriving from discovery itself. This co-emergence proves temporally coincident with the Gestalt-switch characterizing serendipitous discoveries: at the moment of discovery, objects remain physically unchanged, yet their functional significance within scientific paradigms undergoes radical transformation. Bell's 'unwanted interference' became evidence of pulsars and the radio telescope's affordances expanded from quasar detection to pulsar observation; likewise, Wells's 'recreational gas' became surgical anesthesia. Serendipity thus produces not merely new knowledge but expands the very 'landscape of affordances' available within particular scientific contexts. By examining this co-emergence through the lens of emergent phenomena – which cannot be reduced to constituent elements or predicted from antecedent conditions – we illuminate how serendipitous discoveries involve not merely recognizing new facts but reconstituting the very possibilities for action and meaning that objects present to investigators.

The Socio-Cultural Origins of Modern HPS

Fons Dewulf
Hong Kong University of Science and Technology
hmfdewulf@ust.hk

Ever since the inception of History and Philosophy of Science (HPS) as a field, its scholars have disagreed about its aims and methods. In these discussions, one mostly finds a similar origin story of HPS, focusing purely on intellectual motivations: in the early 1960s, Hanson, Feyerabend and Kuhn used the history of science to make philosophical theories of science concrete, countering the logical empiricists' tendency to create empty formal models unrelated to actual scientific practice. Supposedly, these efforts cemented widespread interest in HPS and invigorated a turn to practice.
I will argue that this classical origin story is misguided for two reasons. First, it neglects important efforts just after the Second World War, long before Kuhn or Hanson rose to any prominence. Second, it obscures the broader socio-cultural motivations for universities and governments in the post-war period to invest in HPS. These practical motivations, I argue, proved more decisive for the field than the purely intellectual frustrations of Hanson or Kuhn with empty formalizations. Based on an expanded history, I point out that HPS emerged primarily due to its politically and culturally recognized potential as a bridge between the

scientific enterprise and broader socio-cultural developments. HPS as a field, I argue, should take this cultural context of its own emergence to heart in discussions on its aims and methods.

To defend my historical claim, I investigate two important institutional moments for the emergence of HPS as we know it. First, I discuss the appearance of the category "HPS" in the USA. In 1944, atomic physicist Raymond Seeger founded a new section "History and Philosophy of Science" at the American Association for the Advancement of Science (AAAS). As participant in the Manhattan project, Seeger, and many of his colleagues, wanted HPS to aid scientists in their reflection on the meaning and place of their work in human history. By investigating correspondence and institutional documents, I show that, in the initial years after Hiroshima, a need for HPS as a social instrument found important support, largely through Seeger and James Conant, at the National Science Foundation and the American Institute for the Unity of Science. I discuss how this support lost momentum from roughly 1950 onward – most importantly due to disciplinary tensions between historians, philosophers and the natural scientists at the AAAS.

Next, I investigate the founding of the HPS department at Indiana University in 1960 by Norwood Hanson. Based on internal documents and Hanson's inaugural speech, I show that the university funded the program for its potential to prepare scientists-in-training for their culturally leading role in American society. I point to many similarities in the motivation to support HPS at the AAAS. According to Hanson, HPS aimed to shape "tomorrow's minds" in an age where scientific thinking was the dominant cultural force. This socio-cultural motivation for HPS was not unique to Hanson. Similar ideas by Hanson's British associates, Toulmin and Buchdahl, played an important role in the emergence of HPS in the universities of Cambridge and Melbourne.

Imprecise Sexes

Marina DiMarco
Washington University in St Louis
m.dimarco@wustl.edu

Sexes stand accused of many forms of indeterminacy: as used in science, 'male' and 'female' might have multiple meanings, perhaps in different contexts; they might under-specify other relevant details; their boundaries may be blurry, and so on. In light of this, one stream of philosophical work on sexes has been concerned to achieve, or restore, precision and clarity about their meaning(s) as a scientific concept(s) (Richardson 2022), or to abandon them if this can't be adequately salvaged (Watkins & DiMarco 2025). Against this current, however, other philosophers of science hold that conceptual ambiguity can sometimes be legitimate, ineliminable, or even a feature of scientific concepts (e.g., Feest 2017; Neto 2020), including sex (DiMarco 2025). In this talk I'll ask (1) what this means for calls to clarify scientific uses of "sex" and/or preserve their ambiguity, and (2) what we might learn about conceptual indeterminacy in general from careful attention to sexes and the scientific and social practices in which they are embedded. One consideration that cuts across these questions is that sexes are what I call "sticky" kinds: kinds for which social and political baggage, and/or a particular logic of research questions (in Lloyd's 2015 sense), have proven difficult to shed. Because sexes are sticky, I'll argue that forms of imprecision whose benefits depend on transparency about the nature of ambiguity, and/or users' ability to navigate multiple, distinct practices and conceptualizations, are relatively risky. Optimistically, education and advocacy

about the uses of sex concepts both in and out of science might help reduce these risks. This could make conceptual imprecision and its benefits for science an additional motivation for social change.

References

DiMarco, Marina. 2025. "Sex as a Biological Variable and the Pursuitworthiness of Exploratory Inquiry." European Journal for Philosophy of Science 15 (4): 71. https://doi.org/10.1007/s13194-025-00701-0.
Feest, Uljana. 2017. "Phenomena and Objects of Research in the Cognitive and Behavioral Sciences." Philosophy of Science 84 (5): 1165–76. https://doi.org/10.1086/694155.
Lloyd, Elisabeth A. 2015. "Adaptationism and the Logic of Research Questions: How to Think Clearly About Evolutionary Causes." Biological Theory 10 (4): 343–62. https://doi.org/10.1007/s13752-015-0214-2.
Neto, Celso. 2020. "When Imprecision Is a Good Thing, or How Imprecise Concepts Facilitate Integration in Biology." Biology & Philosophy 35 (6): 58. https://doi.org/10.1007/s10539-020-09774-y.
Richardson, Sarah S. 2022. "Sex Contextualism." Philosophy, Theory, and Practice in Biology 14 (2). https://doi.org/10.3998/ptpbio.2096.
Watkins, Aja, and Marina DiMarco. 2024. "Sex Eliminativism." Biology & Philosophy 40 (1): 2. https://doi.org/10.1007/s10539-024-09972-y.

The Road Not Taken: Stephen Toulmin and the Primacy of Practice in (Mid-Century) Philosophy of Science

Max Dresow
Minnesota Center for Philosophy of Science
max.w.dresow@gmail.com

At some point in the not-so-distant past, mainstream philosophy of science experienced a "turn to practice." The "turn" didn't happen all at once, and it didn't affect all parts of the field equally. But at least by the 1980s, "practice" had become the watchword of a reformist movement, which found its epitome in Ian Hacking's famous motto, "Experimentation has a life of its own" (Hacking 1983). Hacking complained that philosophers of previous generations had been obsessively fixated on matters relating to scientific theory. Sometimes it was particular theories, like general relativity. Other times it was theories in general. Anyway, it had led to a massive cataract, in which most of what usually passed as "science" escaped philosophical notice, or at least close scrutiny. It was to redress this situation that the "new experimentalists" of the 1980s launched their reformist campaign.

According to Andrea Woody, the new experimentalists were the first philosophers of science "to advocate explicitly and self-consciously for a turn to practice" (Woody 2014). But this is not so. While the new experimentalists may have been the first philosophers to advocate effectively and self-consciously for a turn to practice, the first person to advocate "explicitly and self-consciously" for a practice-turn was Stephen Toulmin— a Cambridge-trained philosopher and disciple of the later Wittgenstein. He did this in a slender book titled The Philosophy of Science: An Introduction (1953). This was his attempt to reorient philosophy of science around a very different set of themes than those that prevailed in the 1950s. Although the book was favorably reviewed, it had little influence on the direction of the field;

and while Toulmin went on to a successful and eclectic career, he played little part in the practice-turn that began to crackle during the 1980s.
This talk has a modest aim. It is to critically review this early practice-oriented philosophy, which even sympathetic commentators have tended to overlook, or else to treat as a preamble to the more overtly historicist philosophy Toulmin began to develop during the 1960s. A secondary aim is to display its basically Wittgensteinian character, which in fact Toulmin repeatedly acknowledged in the text.
I also wish to broach a larger issue. Paul Franco has recently argued that "ordinary language" philosophers like Toulmin played an underappreciated role in the historicization of philosophy of science in the 1960s (Franco 2021). The implication is that, by attending to these oft-neglected contributions, we can better understand why philosophy of science developed as it did in the decades following World War II. I agree with this. Still, when it comes to Toulmin's early work, I'm inclined to emphasize a different point. Far from illuminating how philosophy of science actually developed during the twentieth century, Toulmin's philosophy is more useful for drawing attention to an unrealized possibility: a road not taken. Philosophy of science in the post-war period may bear the stamp of ordinary language analysis, but of Toulmin's philosophy of science in practice, very little survived.

References

Franco, P. L. 2021. Ordinary language philosophy, explanation, and the historical turn in philosophy of science. Studies in History and Philosophy of Science Part A 90:77–85.
Hacking, I. 1983. Representing and Intervening: Introductory Topics in Philosophy of Science. Cambridge: Cambridge University Press.
Toulmin, S. 1953. The Philosophy of Science: An Introduction. New York: Harper & Brothers.
Woody, A. 2014. Chemistry's Periodic Law: Rethinking Representation and Explanation after the Turn to Practice. In Soler, L. Lynch, M. and Isreal-Jost, V. (eds.) 2014. Science After the Practice Turn in the Philosophy, History, and Social Studies of Science, 123–50. New York: Routledge.

Different Theories of Constitutive Relevance Relation as Conceptual Tools for Mechanistic Discovery in Neurobiology

Yichu Fan
University of Edinburgh
yichu.fan@ed.ac.uk

The project of understanding mechanistic explanations in neurobiology calls for an adequate account of constitutive relevance, i.e. the relation between a phenomenon and the components of the mechanism underlying the phenomenon. Building on Craver's (2007) mutual manipulability account of constitutive relevance (and its criticisms), two lines of recent approaches to accounting for constitutive relevance have gained popularity, namely, the "no decoupling" approach (Baumgartner and Casini, 2017; see also Romero, 2015; Baumgartner and Gebharter, 2016) and the "in-betweenness" approach (Menzies, 2012; Harinen, 2018; Krickel, 2018a; 2018b; Prychitko, 2021; Craver et al., 2021). These two accounts are often pitched against each other as competing philosophical theories, fueling an ongoing debate (e.g. Baumgartner and Casini, 2025). In contrast, this paper reconciles these

two accounts of constitutive relevance by arguing for a pluralist view. Specifically, I argue that different philosophical theories of constitutive relevance can function as different conceptual tools, guiding mechanistic discoveries in different empirical settings in neurobiology.

In particular, I will first highlight a conceptual difference between the two accounts of constitutive relevance, namely that the NDC accounts model the explanandum phenomenon with a single variable, whereas the in-betweenness accounts model the phenomenon in terms of a causal chain between input and output variables. I argue that the two accounts do not directly conflict because there is not predetermined way to translate these two conceptual pictures. Then, I argue that in experimental contexts, the same phenomenon indeed can be subject to different conceptual frameworks and be modelled by different variables, partly depending on the pragmatic considerations of the researchers. Therefore, I argue that which account of constitutive relevance is more applicable to a given case depends on whose conceptual picture matches better with the conceptual framework adopted by the researchers for the empirical investigation at hand.

To support my argument, I examine case studies from the mechanistic explanation of the circadian rhythm in Drosophila. I argue that during what I call the "localizing" stage of the mechanistic discovery, the circadian phenomenon was represented by a single variable (e.g. the behavioral events of Drosophila as detected by infrared light) due to the endogenous nature of the phenomenon. Therefore, the NDC approach to constitutive relevance offers a tailored reconstruction of the experimental works conducted at this phase of research. On the other hand, during what I call the "enriching stage" of the mechanistic discovery, the circadian phenomenon was then represented by a chain of variables connected by causal relations. Given this conceptual framework, the in-betweenness account serves better as the guiding principle for the experimental studies at this stage. In sum, I argue that different philosophical accounts of constitutive relevance should be understood as conceptual tools, shedding light on different empirical scenarios. From this instrumental perspective, we should strive for a larger rather than smaller philosophical toolbox with a constellation of theories ready to facilitate the progress of mechanistic discovery.

Epistemic Injustices in the Use of Hormone Replacement Therapy Following the Women's Health Initiative (WHI) Study

Manuela Fernández Pinto
Universidad de los Andes
m.fernandezp@uniandes.edu.co

Allison Wolf
Universidad de los Andes
a.wolf@uniandes.edu.co

Alejandra Bermúdez Jiménez
Universidad de los Andes
a.bermudezj2@uniandes.edu.co

Laura Vargas Santiago
Universidad de los Andes
lv.vargass1@uniandes.edu.co

By the end of the 20th century, the use hormone replacement therapy (HRT) to manage and ameliorate menopausal symptoms was quite common and large observational studies associated estrogen use with lower incidence of coronary disease, osteoporosis, and overall mortality. Despite this, the lack of randomized clinical trials in relation to menopause management along with increasing feminist activism through the Women's Health Movement, demanded better medical knowledge from, for, and about women. In response, in 1992 the National Institutes of Health launched the Women's Health Initiative (WHI). As part of its work, the WHI conducted a large-scale three-arm randomized controlled trial on HRT comparing the safety and effects of two treatments (conjugated equine estrogens combined with medroxyprogesterone acetate and estrogen alone in hysterectomized women) with a control group, in relation to coronary heart disease, stroke, venous thromboembolism, and cancers of major public health relevance, particularly breast and colorectal cancer.
In 2002, an increased risk of breast cancer, coronary heart disease, stroke, and venous thromboembolism in some study participants led to the premature termination of the WHI HRT study, marking a turning point in menopausal medicine. The immediate dissemination of these results through official communications and press conferences triggered a global reaction characterized by alarmist headlines, restrictive regulatory measures, and a sharp decline in HRT prescriptions. Many women and their doctors, immediately began to associate HRT with breast cancer, a perception that continues today despite the fact that the ten- and twenty-year reassessments of the WHI concluded that the original interpretation had overstated these risks. Consequently, many perimenopausal and menopausal women think that they must simply suffer symptoms from menopause while many are denied HRT by providers who still rely on the WHI study.
There are various critiques of the current state of affairs from physicians, researchers, and activists, but as of yet there has been little explicit philosophical investigation into the relationship between menopause care, HRT, and epistemic concerns. This paper aims to change that. More specifically, we will analyze the case of HRT research to reveal the epistemic injustices towards patients who seek treatment for menopausal symptoms. We will do so first by identifying some of the more obvious testimonial and hermeneutical injustices those seeking menopause treatment endure, related, for instance, to doctor's paternalist attitude towards women and difficulties accessing medical literature. Then, we will deepen this analysis by revealing how other forms of epistemic oppression manifest in these circumstances. These include, for example, issues of epistemic marginalization and cultural imperialism, which specifically affect patients and physicians in historically marginalized groups or in poor countries. Confronting these injustices, however, requires going beyond identifying cases of epistemic oppression or classifying them. Accordingly, we uncover the underlying ontological and epistemological conditions in medical research that generate these injustices in the first place, including the predominant individualistic and compartmentalized approach to medical knowledge. We argue that a primary cause of these problems is the failure of medical research to embrace a feminist epistemology capable of confronting the value-free ideal of science. Consequently, reducing epistemic injustice in menopause care will require rejecting such an ideal in researching HRT.

In vitro models and replacement alternatives in toxicology: exemplification, representation, and disruption

Grant Fisher
Korea Advanced Institute of Science and Technology
fisher@kaist.ac.kr

New approach methodologies (NAMs) offer potential replacements or complements for animals used in scientific experiments. NAMs are attracting attention from philosophers of science in practice as well as scholars working in animal studies due to their potential epistemic, ontological, and policy significance. For example, Ankeny et al (2024) argue that the humanities and social science can help to inform practitioners on the difficulties in establishing epistemic standards in a proposed shift from data based on animal data to NAMs. Leonelli (2023) probes the repurposing of existing forms of model organisms from biology to precision medicine and toxicology (Leonelli 2023). In this paper, I focus on models repurposed from biomedicine to toxicology using or derived from human cells: in vitro models that are, at least in part, constructed from the same stuff as their human target systems. They are envisaged as surrogates that stand-in for their human targets that cannot for obvious ethical reasons be used as test subjects in toxicology. As replacement alternatives, they are also proposed as stand-ins for the animals they are intended to replace in toxicological risk assessments. They do so, I will argue, by exemplifying specific properties in support of this role. For example, replacement alternatives would be expected to be capable of providing mechanistic evidence in a modular form compatible with the adverse outcome pathways (AOP) anchoring framework in toxicology. This paper probes the epistemic relationship between exemplification and representation and argues that these models occupy a contested space of transition in which novel and repurposed models are offered as candidates for potential validation and eventual regulatory approval. As the proposed models increase in organismic complexity, the greater the ontological and regulatory ambiguities that result which may undermine their surrogative roles as stand-ins for their targets, becoming ambiguous ontologies of regulatory and moral concern of their own.

References.

Ankeny, Rachel. A., Davies, Gail., Kirk, Robert G.W, Whittaker, Alexandra L., and Johnson, Jane. “Lessons for the Future of NAMs from History, Philosophy and Social Studies of Science”, Alternatives to Laboratory Animals 52(5) (2024): 276-278.
Leonelli, Sabina. “The Future of the Model Organism Repertoire”, lecture presented at the symposium Model Organisms: Materiality, History and Politics, ICI Berlin, 21 March 2024, video recording, mp4, 48:47 <https://doi.org/10.25620/e240321-1>

A philosophical toolbox for improving measurement in cognitive neuroscience

Jolien C. Francken
University of Amsterdam
j.c.francken@uva.nl

Cognitive neuroscience aims to uncover the neural mechanisms underlying cognitive capacities, such as memory or consciousness. However, 'what' we measure is not independent of 'how' we measure it: different ways of measuring e.g., 'consciousness' often capture different things. Moreover, complex experimental manipulation and measurement techniques are used, producing a heterogenous set of results from different methods and levels of investigation. Yet interpretation of the meaning of these measurements, and their integration, is underdeveloped, resulting in conceptual and methodological inconsistency and fragmentation. Extending previous philosophical work, here I propose a novel approach in the form of an applicable 'philosophical measurement toolbox', with the ambitious goal of resolving this situation. The project targets the problems of inconsistency and fragmentation by directly addressing the root of the problems, i.e., the interdependence of conceptualization and measurement.
In this paper, I will discuss the 'Integration' measurement tool and focus on the question of how to integrate evidence from different measurement methods. Cognitive neuroscience currently lacks a systematic method for integration, impeding multilevel convergence and understanding. First, the field is specialized and organized along specific methods with their own training, journals, and conferences, and integration is often not their primary goal. Second, cognitive neuroscientists do sometimes attempt to integrate findings (e.g., behavioural data, neuroimaging data and computational models) yet often in an intuitive way, without a fully developed theoretical basis and consideration of philosophical assumptions. To fill this gap, I introduce an interdisciplinary concept/method for using multiple approaches to study a phenomenon, called 'triangulation', and investigate and adapt this theoretical concept and the methodology of triangulation to fit cognitive neuroscience practices. Specifically, I focus on the integration of different cognitive neuroscience methods, and measurements/data obtained at different levels. How should we weigh different types of data, e.g., are intracranial data more informative than electro-encephalography (EEG) data because measurements are more direct?
The philosophical analysis of the concept of triangulation prepares for development of a corresponding 'Integration' tool (i.e., a concrete protocol), and its application and testing in a cognitive neuroscience case study. In this research project, human participants and animal subjects perform similar behavioural tasks manipulating and measuring consciousness, and simultaneously human EEG data and mouse intracranial electrophysiology data are obtained. In collaboration with the cognitive neuroscientists involved, I will employ the 'Integration' tool in the data analysis and interpretation stages of this research project. This should provide a more coherent answer to their research questions in a standardized way, explicitly connecting to cognitive neuroscience's epistemic goals, and reveal points of divergence creating opportunities for new research directions. Iterative cycles of conceptual and methodological 'engineering' ensure that the 'Integration' tool is continuously improved to fulfil its intended practical function in alignment with the underlying triangulation concept. The resulting measurement toolbox will equip cognitive neuroscientists with a practical yet conceptually rigorous methodology to solve their measurement problems from the ground up, perform more effective experiments, and achieve a more meaningful integration of different studies.

What Counts as Evidence? Reconsidering Evidence-Based Policy for Climate Change Adaptation

Niklas Gärtner
Université Grenoble Alpes
niklas.gaertner@univ-grenoble-alpes.fr

Evidence-based policy (EBP) builds on the principles of evidence-based medicine (EBM), advocating for the systematic use of high-quality scientific evidence for decision-making. EBP promises structured methods for evaluating evidence and systematic guidelines for decision-makers, making it especially appealing for climate change adaptation (CCA), where scientific data is complex and, due to inherent uncertainties, difficult to interpret for non-experts. While EBP is widely accepted, there is little discussion about what constitutes evidence in policymaking for CCA. This paper addresses that gap by examining how evidence hierarchies function in practice and why they fail when transferred from medical to climate contexts.

EBP originated from EBM's use of evidence hierarchies, which rank systematic reviews and randomized controlled trials above observational studies and expert judgment (Sackett, 1998; Burns et al., 2011). However, applying this hierarchy directly to CCA is problematic. First, some evidence types in the traditional hierarchy simply do not exist in climate science, and vice versa. Second, ranking climate evidence according to EBM's quality criteria produces an inadequate hierarchy. Evidence hierarchies are question-dependent: the EBM hierarchy is designed for intervention-focused causal questions about treatments, while climate change assessment requires not only causal attribution but also predictions and policy guidance, which are tasks the EBM hierarchy was never designed to support (Jerker, 2021).

The relevance and availability of evidence in policy decision-making differ significantly from clinical decision-making. I argue that for every context in which EBP is applied, practitioners must determine what constitutes evidence, what is relevant, what is available, and how different evidence types should be integrated. I suggest three modifications of EBP that I think are necessary to effectively apply it in CCA contexts.

First, EBP must incorporate epistemic pluralism by integrating multiple knowledge systems rather than privileging a single empirical methodology (Longino, 2020 Harding, 2013; Wylie, 2013). CCA requires synthesizing diverse types of evidence (climate models, observations, expert judgment, and local knowledge) each contributing complementary and partial insights. Different systems capture phenomena at different scales, and integrating across systems enables more complete understanding than any single approach could provide (Teng¬®o et al., 2014; Orlove et al., 2023). This has practical implications for how adaptation assessments are conducted, and also whose knowledge is included (Fricker, 2007).

Second, evidence hierarchies must treat relevance as an epistemic criterion alongside methodological quality (Montuschi, 2009; Parkhurst and Abeysinghe, 2016). If evidence is not relevant to the decision at hand, it provides no practical value (Anderson, 2004). Systematic reviews may rank high on methodological quality but offer limited guidance for specific adaptation decisions, whereas local expert judgment may prove essential despite its subordinate position in traditional hierarchies. This suggests the need for multiple flexible hierarchies depending on decision context.

Third, EBP must abandon value-neutrality claims and acknowledge the legitimate role of values in evidence generation and assessment. A practice-oriented framework recognizes values as constitutive of knowledge production (Intemann, 2005; Douglas, 2009). For example, expert judgment derives its usefulness from contextual knowledge that helps

manage uncertainties in specific decision settings (Majszak and Jebeile, 2023). The key distinction lies not in whether values are present, but in whether they serve the goals of the practice at hand and are made explicit for scrutiny.
These modifications point toward a reconceptualization of EBP that takes serious how evidence functions within actual decision-making practices, rather than treating evidence assessment as an abstract methodological exercise detached from context and use.

References

Anderson, E. (2004). Uses of value judgments in science: A general argument, with lessons from a case study of feminist research on divorce. Hypatia, 19 (1), 1–24.
Burns, P. B., Rohrich, R. J., & Chung, K. C. (2011). The levels of evidence and their role in evidence-based medicine. Plastic and reconstructive surgery, 128 (1), 305–310.
Douglas, H. E. (2009). Science, policy, and the value-free ideal. University of Pittsburgh Press
Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.
Harding, S. (2013). Rethinking standpoint epistemology: What is "strong objectivity"? In Feminist epistemologies (pp. 49–82). Routledge.
Intemann, K. (2005). Feminism, underdetermination, and values in science. Philosophy of science, 72 (5), 1001–1012.
Jerkert, J. (2021). On the meaning of medical evidence hierarchies. Philosophy of Medicine, 2 (1).
Longino, H. E. (2020). Science as social knowledge: Values and objectivity in scientific inquiry.
Majszak, M., & Jebeile, J. (2023). Expert judgment in climate science: How it is used and how it can be justified. Studies in history and philosophy of science, 100, 32–38.
Montuschi, E. (2009). Questions of evidence in evidence-based policy. Axiomathes, 19 (4), 425–439.
Orlove, B., Sherpa, P., Dawson, N., Adelekan, I., Alangui, W., Carmona, R., Coen, D., Nelson, M. K., Reyes-García, V., Rubis, J., et al. (2023). Placing diverse knowledge systems at the core of transformative climate research. Ambio, 52 (9), 1431–1447.
Parkhurst, J. O., & Abeysinghe, S. (2016). What constitutes "good" evidence for public health and social policy-making? from hierarchies to appropriateness. Social epistemology, 30 (5-6), 665–679.
Sackett, D. L. (1998). Evidence-based medicine. Spine, 23 (10), 1085–1086.
Teng¬®o, M., Brondizio, E. S., Elmqvist, T., Malmer, P., & Spierenburg, M. (2014). Connecting diverse knowledge systems for enhanced ecosystem governance: The multiple evidence base approach. Ambio, 43 (5), 579–591.
Wylie, A. (2013). Why standpoint matters. In Science and other cultures (pp. 26–48). Routledge.

When farmers make experiments, they sometimes bet on the unexpected -- and this is not serendipity

Paul-Edouard Gerard
UniLaSalle Beauvais
paul-edouard.gerard@unilasalle.fr

Nicolas Brault
UniLaSalle Beauvais
nicolas.brault@unilasalle.fr

Michel Dubois
UniLaSalle Beauvais
michel.dubois@unilasalle.fr

In interviews with farmers conducting self-initiated experiments on their farms, a recurrent pattern emerges. A question is posed--how to improve soil health, reduce external inputs, or reorganise labour. Hypotheses are formulated, and an experimentation process begins. Yet in some cases: (1) the initial question is not resolved; (2) the experiment nevertheless generates valuable knowledge; (3) this knowledge remains difficult to formalise within conventional agronomic publications. Yet farmers consistently describe these processes as successful. We may characterise them as "bets on the unexpected"--similar to serendipity, while their recurrence suggests a more systematic epistemic dynamic.
To understand this, we propose adopting an ontology of relation rather than one of structure. Drawing on Simondon's philosophy of individuation, we suggest these experiments do not apply fixed models to given contexts. Rather, they operate within evolving systems where problems, knowledge, and field of action co-emerge. The farmer does not seek to verify a hypothesis, but engages with a milieu in tension--where neither problems nor solutions are fully determined in advance.
To formalise this, we introduce a tripartite framework. The farm is first understood as a milieu (M): a heterogeneous, historically constituted mesh of agents--soils, tools, plants, infrastructures, norms, and farmers themselves. This milieu is rough, noisy, and inherited, resisting standardisation and generating delayed effects. Experimental action within this milieu simultaneously transforms conceptual trajectories (C), knowledge (K), and the milieu itself. This process, which we term CKM co-individuation, captures the entangled evolution of concepts, knowledge, and contextual conditions.
This frame helps to explain why answers are not solely validated by the experimentation, but emerge from it: farmers do not resolve predefined questions but transform them. Knowledge is not extracted from a stable reality, but produced through engagement with a moving field. Moreover, this CKM process embeds farmers' practice within an ontology of relation. In this regime, and following Simondon's allagmatic, the knowledge created cannot take the form of operational schemata. Rather, it corresponds to schematisms of thought--cognitive figures that orient perception and guide situated engagement. This matches what farmers report. In contrast, as Cohen (2017) notes, classical agronomic science relies on a nomologico-prescriptive epistemology, rooted in an ontology of structures. The divergence between these regimes helps explain the apparent misfit between farmers' knowledge and academic dissemination formats.
In this light, the "bet on the unexpected" is not serendipity. It is the expression of a mode of knowledge production grounded in co-individuation conducted by farmers within a lively, partially resistant milieu. The un-expected emerges not from randomness, but from relational

engagement where the farmer is not the sole agent of change, but a pilot among many forces. This results in specific--yet rigorous--forms of knowledge, that resist the usual scientific frameworks.
Showing that knowledge can emerge from the co-evolution of concepts, knowledge, and milieu--not by resolving predefined problems, but by transforming them--CKM framework invites a practice-based philosophy attentive to both reasoning and material entanglement.
Cohen (2017). Des lois agronomiques à l'enquête agroécologique.
https://doi.org/10.4000/traces.6989

Promissory Science and the Commodification of Aging

Sara Green
University of Copenhagen
sara.green@ind.ku.dk

Antonia Kaspar
University of Copenhagen
antonia.kaspar@stud.uni-regensburg.de

Biological aging is not only a rapidly expanding field of scientific research but also the target of a growing commercial market of self-tests, health counseling, and lifestyle optimization services. Promoting novel opportunities to quantify and manage the pace of aging, companies market molecular biomarkers—such as DNA methylation "clocks" and glycan patterns—as personalized indicators of biological decay and health status. These products promise actionable insights and longer, healthier lives, while emphasizing their scientific legitimacy and evidence-based character. Yet, how is "science" mobilized, interpreted, or even distorted in this emerging bioeconomy of aging? This paper analyzes how an image of scientific authority is constructed in the marketing and practice of commercial aging biomarker testing. Combining online ethnography with a review of biomedical literature, we analyze how promotional narratives, data visualizations, and health recommendations align—or fail to align—with current scientific understanding of aging. Our analysis reveals that commercial actors deploy a range of rhetorical techniques that produce an image of scientific certainty and clinical utility, despite the lack of consensus on the ontology, reliability, and interpretive scope of aging biomarkers. We argue that aging biomarker testing exemplifies a contemporary form of biomedical promissory discourse, where the authority of science is simultaneously invoked and reshaped through market dynamics. This promisory discourse allows for new actors to influence scientific and market developments. For example, on social medica, the figure of the "scifluencer" (scientific influencer) to play a key role in legitimizing these technologies by blending lifestyle coaching with scientific discourse, thereby blurring the boundary between expertise and entrepreneurship. Moreover, we stress the ethical implications of this trend: how consumers are subtly nudged toward non-evidence-based interventions—from costly supplement regimens to unnecessary and potentially harmful diagnostic screenings—under the banner of "science-backed" self-care. We conclude by calling for a more critical attention to the epistemic and moral consequences of premature translation of complex science into seemingly straightforward predictors, with important implications for how aging and health are perceived.

Has AlphaFold solved the protein folding problem?

Gregor Greslehner
University of Vienna
gregor.greslehner@univie.ac.at

The protein folding problem has been receiving a lot of attention, especially since the predictive success of the AI tool AlphaFold and the 2024 Nobel Prize in chemistry. The prediction results by the deep machine learning program made big headlines – its achievements were sometimes even appraised to have solved the protein folding problem. But what would it mean for the protein folding problem to be solved? Here, I investigate and develop an account of what "solving the protein folding problem" would entail. In particular, I argue that it would have to include an understanding of the folding mechanism, insight into chemical and physical principles, being able to predict from scratch for different types of proteins in different kinds of environments and conditions– and not just those it has been trained on with available experimental structural data (which, for example do not contain many transmembrane proteins or intrinsically unstructured proteins).
Other theoretical problems exist that a solution of the protein folding problem would entail: among others, a proper solution would need to explain how small differences in amino acid sequence can have a huge impact on the three-dimensional folding structure, how intrinsically unstructured proteins (or their intrinsically unstructured parts) behave, what the free energy of different folding states is, what the individual steps and dynamics of the folding process are, etc. Predictive success of the final protein, even if it were to give an accurate structural model to each amino acid sequence, would not solve the protein folding problem. Many open questions remain that AlphaFold cannot provide but that one would expect from a proper solution to the protein folding problem.
What AlphaFold did do: it provides some sketched ideas of hitherto blank spots of the protein universe, predicted models that can easily be worked with (under fitting caveats). Proteins whose structure has not been determined experimentally, for whatever limiting reason, can now at least have a predicted model structure that can work as a starting point for further modeling and determination of the actual structures. This is indeed an impactful achievement with a lot of scientific value. Saying that AlphaFold would have solved the protein folding problem, however, would be declaring a premature victory.
These general considerations about AlphaFold also have implications for other deep learning technologies and their applications in other areas of science. However, the basic scientific puzzle of how and why proteins fold remains to be a fundamental problem that is unlikely to be solved by AlphaFold or any other prediction. By applying the criteria of what it would mean to solve the protein folding problem, we can then answer the title question negatively– that is: no, AlphaFold has not solved the protein folding problem. However, it is a powerful data generation and prediction tool for modeling, crafting hypotheses about mechanisms, providing benchmark and reference data, and understanding the guiding principles of the folding process, which– despite AlphaFold's achievements–remains an open scientific problem.

Supporting the validity of the 'Variety of Evidence Thesis': evidential pluralism in evidence-based management

Paride Del Grosso
University of Antwerp
paride.delgrosso@uantwerpen.be

The Variety of Evidence Thesis (VET) states that, ceteris paribus, a more varied body of evidence confirms a hypothesis better than a less varied body of evidence, because evidential variety increases the robustness of the hypothesis. The VET has received criticisms (often from the Bayesian epistemological perspective) and its validity is still discussed (Landes 2021). The purpose of this paper is to argue that the VET holds as a 'rule of thumb' (i.e. a practical heuristic) for practitioners.
I will focus on evidence-based management (EBMgt) (Barends and Rousseau 2018), an approach to management according to which predictions about the effectiveness of managerial interventions should be supported by the best available scientific evidence. Causal hypotheses in EBMgt tend to be highly context-specific and, to support them, practitioners often need to use evidence of both general and single-case causal relations. Despite being pivotal also in other evidence-based practices, the interplay between evidence on different causal levels has remained quite unexplored in the literature on causality. This motivates my choice to adopt EBMgt as a case study in which to explore the practical validity of the VET.
First, after introducing the VET, I will focus on causal hypotheses and on evidential pluralism (EP), an approach to evidence for causation that endorses evidential variety by prescribing the triangulation of two types of evidence – difference-making and mechanistic evidence (Shan and Williamson 2023). EP has been promoted in several evidence-based practices (like evidence-based policy), but neither in relation to the VET, nor to EBMgt.
Second, I will delineate what I call the orthodox approach (OA) to evidence for causation in EBMgt. I will show its application in a real case study concerning remote working and employees' performance. Regarding the VET, the OA promotes evidential variety in terms of sources. This allows to effectively combine evidence of general causal claims with context-specific evidence, thus increasing hypothesis robustness. However, the OA does not promote evidential variety in terms of types, since it mainly prescribes the use of difference-making evidence, possibly from randomised controlled trials (RCTs). This limits hypothesis robustness, especially in cases of extrapolation.
Third, I will point out the epistemic advantages of integrating difference-making evidence with mechanistic evidence (as prescribed by EP). In line with the VET, increasing type-variety in such a way improves single-case-hypothesis robustness because mechanistic evidence mitigates the downsides of difference-making evidence and allows for better extrapolations.
Fourth, I will highlight the practical advantages of adopting the VET: employing a varied body of evidence helps develop more cost-effective solutions and allows to employ methods (e.g. quasi-experimental methods) that are less costly and time-demanding than RCTs.
Given these epistemic and practical advantages, I will conclude that the VET works as a 'rule of thumb' in EBMgt and, potentially, also in other evidence-based practices in which single-case causal hypotheses are central.

References

Barends, E., Rousseau, D. (2018). EBMgt. Kogan Page.

Landes, J. (2021). The VET and its independence of degrees of independence. Synthese, 198: 10611-10641.
Shan, Y., Williamson, J. (2023). EP in the Social Sciences. London: Routledge.

Explaining with networks in the social sciences

Chen Long Guo
University of Cambridge, Department of History and Philosophy of Science
clg65@cam.ac.uk

Existing philosophical accounts of network-based explanations are predominantly developed by philosophers of the life sciences. Within this body of literature, Huneman offers arguably the most systematic and total defense of the position that network explanations offer a largely unique type of (topological) explanation that is 1. not reducible to other models of explanation and 2. independently explanatory in virtue of mathematical entailment by network topology.
I offer a social scientific perspective to the philosophy of network science by arguing that, in practice, the explanatory relevance of social network analysis necessarily depends on contextually variable as well as non-structural social and behavioural factors external to network topology. I flesh out this claim in three parts, weaving an extended case study of the network scientific research on contagion diffusion throughout.
First, I assess whether Huneman's framework adequately characterises how social network analysis explains, concluding that his account fails to capture how the explanatory role of social scientific background knowledge goes beyond merely guiding and justifying the researcher's network reconstruction of the target system. I argue that within social network analysis, non-structural knowledge plays the crucial role of evaluating the explanatory relevancy of competing topological measures. Crucially, given that topological measures only entail a social phenomenon when combined with a set of non-trivial social and behavioural assumptions, non-structural social scientific knowledge is indispensable to any full social network explanations.
Second, I demonstrate how my criticism of Huneman's account bears on the practice of social network analysis itself. Since the explanatory relevance of topological measures relies on context-dependent psychological and social factors, we must temper the ambitions of the formalist approach to social network analysis, which seeks to develop generic topological measures that can convey structurally important information across diverse social systems. Given that the explanatory relevance of such measures depends on non-structural information they cannot capture, I conclude that 1. their success as explanatory tools will drastically vary across different empirical target systems and 2. even given initial success in statistically accounting for a phenomenon, their explanatory relevance may subsequently be screened off by non-structural factors.
Finally, based on the preceding discussion, I develop an account of social network explanations that builds on José Díez's conception of scientific explanations as ampliative, specialised, embeddings (ASE). I emphasise that since a necessary condition of scientific explanations is to "capture" the explanandum within a continuously developing "theory-net" rooted in a set of guiding principles with branches of increasingly specialised context-specific nomological regularities, the question of interest is not a generic question of how network structures explain, but the specific question of how network analyses can provide uniquely social scientific explanations congruent within existing disciplinary knowledge. Based on an

in-depth analysis of the actual process of development for a number of network centrality measures, I argue that social network analysis successfully explains when it expands on existing social scientific theory-nets through demonstrating novel implications of their constituent regularities when a structural pattern of tie-distribution is accounted for.

Embodied Scientific Practice: Identity Development and Epistemic Innovation Among Scientists With Disabilities

Jonathan Hall
California State University, San Bernardino
jlhall@csusb.edu

Mila Rosa Carden
University of North Texas
milarosa.carden@unt.edu

Scientists with disabilities (ScWDs) occupy a meaningful and underexamined position within scientific communities. Their experiences reveal how scientific practice is shaped by embodied constraints, institutional norms, and the social organization of scientific work. While research in STEM education and identity development has shown that recognition as a scientist can profoundly affect academic persistence and career trajectories, little attention has been given to how disability intersects with the practical activities through which scientific knowledge is produced. This paper adopts the Society for Philosophy of Science in Practice commitment to integrating theory, practice, and the world by examining how ScWDs navigate, modify, and sometimes transform scientific practice.

Drawing on a conceptual framework grounded in identity development and intersectionality, we investigate the life histories, present-day professional experiences, and reflections of seven influential ScWDs across diverse scientific disciplines. Participants engaged in three in-depth interviews that traced their formative years, early encounters with STEM, professional trajectories, and reflections on the norms and expectations of scientific work. Analysis of these lived experiences illuminates the complex interactions among ableist assumptions, disciplinary practices, and the cultivation of a scientific identity.

Three themes from the data were selected for presentation. First, participants described their STEM figured worlds as shaped by implicit and explicit expectations about ideal scientific bodies based on ableist perspectives, productivity standards, and participation practices that positioned disabled scientists as marginal or exceptional. These findings highlight how STEM structures, social expectations, and material environments mediate scientific practice, influencing who can engage in particular activities and how they are recognized as contributors.

Second, despite these constraints, participants articulated science figured worlds as spaces for exercising their curiosity about the natural world. These STEM spaces sometimes allowed alternative modes of problem-solving and novel forms of engagement. Their lived experiences reveal how disability shaped practical experimentation, conceptual reasoning, and collaborative work in ways that challenge standard assumptions about how scientific knowledge is generated. Their experiences suggest that embodied perspectives can serve as epistemic resources rather than deficits, contributing meaningful approaches to modeling, communication, and collaboration.

Third, participants described forms of ongoing identity work that included practice-based adaptation. They employed strategies to navigate inaccessible environments, negotiate credibility, and reconcile their embodied identities with professional expectations. These adaptive practices illuminate the lived epistemology of scientific work by showing how scientists coordinate embodied constraints and disciplinary norms to achieve scientific goals. Together, these findings advance a practice-based philosophy of science by showing how identity development is intertwined with the activities that constitute scientific inquiry. Disability serves as a lens for examining how scientific practice is employed and how actors of STEM figured worlds determine whose contributions count as knowledge. Attending to ScWDs' experiences enriches our understanding of experimental realism, operationalism, and the pragmatics of scientific reasoning. Practical implications for building more inclusive and epistemically productive scientific environments and perspectives will be discussed.

What is Patienthood in Clinical Practices? A Tool-First Approach to Conceptual Engineering

HyeJeong Han
Jeju National University
hhj29@kaist.ac.kr

Yun-Ying Kuo
University of South Carolina
yunyingkuo17@gmail.com

Karen Yan
National Yang Ming Chiao Tung University
karenyan@nycu.edu.tw

Conceptual engineering aims to revise existing concepts, yet philosophers have struggled to implement the concepts revised in actual practices (Jorem, 2021; Machery, 2021; Landes, 2025). In this presentation, we propose a "tool-first approach" to conceptual engineering that addresses this implementation challenge by examining how concepts operate through tools-in-use within situated practices. Rather than stipulating revised concepts from the armchair or testing them through experimental manipulation, our approach employs ethnographic inquiry to analyze the cognitive and material tools that structure reasoning in practice. This focus of the tools-in-use recasts the problem of implementation to the "operativity" of revised concepts.

Our approach builds on a Deweyan view of concepts as rules for action (Westerblad, 2024) while incorporating insights from practice-oriented philosophy of science that emphasize the material and normative constraints of the world (Hacking, 1983; Rouse, 2015; Chang, 2022). By tracing these tools-in-use, we can investigate how concepts function in real contexts and evaluate whether philosophical revisions are capable of becoming operative. Conceptual engineering, on this account, proceeds not through abstract stipulation but through the analysis and possible reconfiguration of the architectures that sustain reasoning.

We demonstrate this method through an ethnographic case study of patienthood in clinical reasoning. Patient-centered care (PCC) is commonly framed through the biopsychosocial model (BPSM), which defines the patient as an integration of biological, psychological, and social dimensions. Yet whether these categories actually organize reasoning has rarely been

examined empirically. Drawing on 125 hours of ethnographic observation and 17 interviews in a specialized cancer hospital, we show that clinicians did not reason through the BPSM's tripartite schema. Instead, oncology, surgery, and palliative care each employed distinct patient models generated by specialty-specific tools: staging systems and expectation management configured patients as therapeutic trajectories; anatomical diagrams and probabilistic framing configured operative bodies and risk-bearers; symptom scales, narrative recapping, and informational scaffolding constituted experiential subjects and epistemic agents.

These findings reveal what we call the "operativity problem" in conceptual engineering: theoretical concepts may be philosophically refined yet fail to fit the structures that organize reasoning in practice. Our tool-first approach addresses this problem by treating tools-in-use as empirical traces of operative concepts. This methodological orientation offers distinct advantages. Unlike experimental approaches that test conceptual intuitions under controlled conditions, ethnographic analysis examines how concepts function in ecologically valid contexts. And unlike purely ameliorative projects that prescribe conceptual change, our approach first establishes whether proposed categories operate as intended before attempting reform.

Overall, our study demonstrates how philosophy of science in practice (PSP) can productively engage with conceptual engineering. By analyzing tool-mediated reasoning, we show that the BPSM's categories are not operative in clinical practice, and that plurality in PCC is not a failure of integration but a structural consequence of distinct tools constituting different forms of patienthood. The tool-first approach thus reorients conceptual engineering from refining abstract definitions to understanding how concepts are enacted through tools in practice. This shift makes the implementation of reformed concepts both empirically tractable and philosophically accountable.

Content and process models as a tool in (philosophy of) science

Rafaela Hillerbrand
KIT - Karlsruher Institute of Technology
rafaela.hillerbrand@kit.edu

Claudia Eckert
Open University
claudia.eckert@open.ac.uk

Models play an important role in science and in engineering. In the last decades, models took centre stage in philosophy of science in practice, and different ideas of models are discussed, ranging from models as representation (e.g. van Fraassen, 1980, Hughes 1997), to models as mediators (Morgan and Morrison, 1999) or models as fiction (e.g. Frigg, 2010), just to name a few. While the philosophy of engineering community has paid considerably less attention to model, models themselves are enormously important in engineering. Design research investigates, amongst others, the use of models in engineering practices in industry. For most products designers, engineers, or manufacturers mostly interact with the final 'product' through models. Thereby it is common in engineering research and development to use a wide variety of models: from physical prototypes, to computer-aided design (CAD) models, mathematical or simulation models. Only a small fraction of models is taken over or developed in analogy to a scientific model. Process models help to deal with complex

organisational structures and high uncertainties in development and production. Engineering models have a complex relationship between the models and its target system (Eckert and Hillerbrand, 2018). Many models exist to bring a product into being, other record once held intensions superseded by multiple design changes. Many models have other models as target systems, for example, prototype built from a CAD model. Notably engineering also uses models to plan engineering processes and has recognised the importance of so-called process models for the management and execution of processes (Eckert/Hillerbrand 2018).
In this paper we contend two things: Firstly, philosophy of science in practice can profit from design research in so far as the latter pays close attention to the way that models are used in engineering practice. Secondly, process models as known in engineering may proof useful for (analysing and shaping) complex scientific settings.
Design research takes the complex realities of engineering modelling in industrial settings as a starting point. The relationship between the model and its target system is for the working engineer not evident from the model itself. This may create tensions as many engineers reuse existing models for which they are not aware of modelling assumptions. While products have the eventual product as an anchor, process model can only relate to the lived experience of the process and the snap shots of its deliverables (Stacey, 2020).
We argue that models in engineering are first and foremost decision-making aids which advance the product development process (Eckert and Hillerbrand, 2022). Understood as decision making aids, process models lend themselves to shape and understand complex scientific practices in the techno sciences such as in high-energy physics (e.g. CERN). But also other large-scale scientific collaborations, e.g. in health sciences, may profit.

Trust in and the authority of science – why it is worth distinguishing

Hannah Hilligardt
Universität Bern
hannah.hilligardt@unibe.ch

Hanna Metzen
Universität Konstanz
hanna.metzen@uni-konstanz.de

For some decades now, philosophers of science have raised concerns about the so-called erosion of scientific authority and a loss of trust in scientific expertise. But what does it mean for science and scientists to have authority in the first place? We show in this talk that in answer to this question, philosophers of science tend to focus on the extent to which the public accepts scientific claims and acts based on them. In contrast, we argue that more attention needs to be paid to science's political authority, i.e. scientists' influence on coercive and binding decision-making. Doing so brings to the fore that science being trusted and having authority are not necessarily the same thing and that value-management strategies in particular need to take into account what kind of authority the research in question has.
Both authority and trust are central concepts within the values in science debate (e.g. Douglas 2009, Kitcher 2011, Holman & Wilholt 2022). Yet despite their prevalence, it is unclear how the concept of authority is used in this context, what kind of authority scientists have and should have in politics and how this relates to trust. Indeed, authority is usually understood as epistemic authority and appears interchangeably with trust. However, using resources from epistemology and political philosophy (Friedman 1990, Moore 2017, Zagzebski 2012), we

argue that in many democratic contexts science does not only have epistemic authority, but indirect political authority and possibly even direct political authority. Trust mostly correlates to epistemic authority. For indirect and direct political authority, trust and authority can go apart.
We will discuss the different kinds of authority scientists can have at the hand of cases from public health research. We argue that it is of importance to clarify when and how science has some form of political authority, not because this is the only way in which science is impactful, but because, according to most democratic theories, authoritative, coercive decision-making demands specific mechanisms of legitimation and control. Characterising science's authority primarily in epistemic terms and equating it to trust and voluntary compliance risks overlooking this dimension. Consequently, we argue that the more political science's authority is, the more important it becomes for values influencing the scientific research process to be politically legitimated.

References

Douglas, H.E. (2009). Science, Policy, and the Value-Free Ideal. University of Pittsburgh Press.
Friedman, R.B. (1990). On the Concept of Authority in Political Philosophy. In Raz, J. (ed.), Authority. New York University Press, p. 56–91.
Holman, B. and T. Wilholt (2022). The New Demarcation Problem. Studies in History and Philosophy of Science 91, 211–220.
Kitcher, P. (2011). Science in a Democratic Society. Prometheus Books.
Moore, A. (2017). Critical Elitism. Cambridge University Press.
Zagzebski, L.T. (2012). Epistemic Authority. Oxford University Press.

Not what, but how is epistemic plurality in engineering? A cross-cultural comparison of irrigation practices

Clara Jing Hohmann
Technical University Berlin
c.hohmann@tu-berlin.de

The understanding of engineering as an applied science is increasingly regarded as outdated. Studies examining engineering practices found sufficient epistemic coherence to call for an epistemology of engineering (Zwart, 2022). Yet this turn away from methods and values long inherited from the rich history of philosophy of the natural sciences has left a void that has only begun to be addressed. Most existing approaches regarding this problem focus on the heuristics of engineers, such as technical know-how and decision-making techniques. In this paper, while still sharing a practice-oriented starting point with previous literature, I argue that a shift of focus from the engineer as agent to the material arrangements brought into function by them, can reveal new insights into the epistemic practice of engineering. Here, by material arrangements, I refer to the set of physical and non-physical elements and their boundaries that are treated as a technological system.
To this end, I examine irrigation systems, as a historically prominent and cross-culturally relevant issue of engineering practice. For this I will draw on Chang's (Chang, 2012) account of an action-based epistemology derived from the history of chemistry and thermodynamics and transfer it to the practical field of engineering. This offers a basis for empirical

reconstruction based on observations of technological practices, that give insights on how solutions are recognized as adequate and how assumptions about material boundaries shape engineering knowledge .
In the first part of the talk, I will present four case studies on irrigation systems in different cultural contexts, concretely in Europe and South-East Asia. Instead of treating these systems as technical designs or artifacts, they are analyzed through their corresponding practices, referring to the context specific relations between task formulation, task solution and how they come together in usage. The perspective reveals different material configurations in how irrigation practices are constituted, particularly in how boundaries between problem and solution are materially set in practice, i.e. through varying modes of water flow manipulation and system integration (field layout, pipe system, data monitoring devices).
In the second part, I will discuss the observed discrepancy between a seemingly constant task and the contrasting wide range of engineered solutions developed in response to it. The variety of effective practices contradicts standard narratives of technological efficiency and can also not be explained on the basis of economic constraints alone. I argue that through a cross-cultural comparison of technological practices, epistemic structures can be made visible by decoupling such from contingent technological trajectories and prominent epochal styles of reasoning (Hacking, 1994). Finally, I will elaborate how the plurality of assumptions about the material boundaries of (1) problem, (2) solution, and (3) the configuration of both in practice suggest a shared epistemic operation, the negotiation process of material boundaries itself.

References:

Chang, H. (2012). Is water H2O?: Evidence, realism and pluralism (Vol. 293). Springer Science & Business Media.
Hacking, I. (1994). Styles of scientific thinking or reasoning: A new analytical tool for historians and philosophers of the sciences. Trends in the Historiography of Science, 31-48.
Zwart, S. (2022). Engineering epistemology: Between theory and practice. Engineering Studies, 14(2), 79-86.

The Elusive Catalogue: Mechanisms and Generality in the Practice of Analytical Sociology

María Jiménez-Buedo
Dept of Logic, History and Philosophy of Science. UNED
mjbuedo@fsof.uned.es

Saúl Pérez-González
Dept of Philosophy. Universitat de Valencia
saul.perez@uv.es

Analytical sociology is a meta-theoretical and methodological framework that highlights the relevance of mechanisms in the social sciences (Hedström, 2005; Hedström & Bearman, 2009; Hedström & Ylikoski, 2010). Inaugurated by sociologists such as Hedström, Swedberg, or Gambetta, and also by economists such as Schelling or philosophers like Elster, analytical sociology is currently is one of the main empirical research programmes in some

important sociological subfields and it often leads the debate around questions such as class stratification, the transmission of socioeconomic inequality, or the study of racial segregation. From the outset, and at the heart of the analytical proposal, is the idea that mature social science should provide of a catalogue of mechanisms that guides sociologists' practice. The centrality of the catalogue for analytical sociologists is explained by its attributed functions: the catalogue would provide a toolbox for building mechanism-based explanations and it would increase their explanatory power by making these explanations more general. It would thus account for knowledge accumulation, which would amount to an ever-growing list of explanations that fit the different types of mechanisms contained in the catalogue.

Despite its centrality to the analytical sociological project, and although loose references to the idea abound both in the theoretical literature around analytical sociology and in the discussion of empirical works within the programme, the idea of the catalogue has not been the subject of direct theoretical scrutiny, neither by practitioners, nor by philosophers of the social sciences. Relevant questions thus remain open regarding the exact form of the catalogue, the precise nature of its individual components, the modes in which it should be employed by practitioners, and the current degree of development of the project of constituting a catalogue of mechanisms.

In this paper, building on the postulates and theoretical assumptions of analytical sociology (many of which are based on and have commonalities with, the precepts of the New Mechanical Philosophy), we reconstruct and operationalise the project of the catalogue of social mechanisms to then contrast it to analytical sociologists' actual practices.

Based on our analysis and operationalisation of the notion of a catalogue of mechanisms, we study several case studies in analytical sociology to contrast our reconstruction and operationalization with relevant mechanistic explanations that are considered canonical examples within analytical sociology. With a particular focus on the case of a recent contribution identifying the Trojan Horse Mechanisms (Arvidsson et al.'s 2021study on labour market mobility patterns), we show the difficulties for the project of the catalogue when put to practice.

In particular, we focus on analytical sociologists' notion of generality of explanation as embedded in the idea of a catalogue of semi-general explanation. This conception of generality, based on Merton's notion of middle-range theory and in Machamer's concept of mechanism schemata, fits poorly the actual practices of analytical sociologists: although classic exemplars in the literature are often cited in subsequent works, we do not find evidence of significant reutilization of semi-general mechanism schemata across different cases. We show how in this case, the practices of analytical sociologists fail to fit the ideas embedded in their manifestoes, even if (or perhaps in part due to the fact) analytical sociologists are very careful in their use of many of the concepts developed by the New Mechanical Philosophy, often developed in close observation of biological mechanisms. We show how, in practice, for any given social mechanistic explanation, it is very difficult to separate the particular details of the singular case with the general features of mechanistic models that could in principle work well across different cases.

References

Arvidsson, M., Collet, F., & Hedström, P. (2021). The Trojan-horse mechanism: How networks reduce gender segregation. Science Advances, 7(16), eabf6730.

Hedström, P. (2005). Dissecting the social: On the principles of analytical sociology. Cambridge University Press.

Hedström, P., & Bearman, P. S. (2009). What is analytical sociology all about? An introductory essay. In P. Hedström & P. S. Bearman (Eds), The Oxford handbook of analytical sociology (pp. 3–24). Oxford University Press.
Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. Annual Review of Sociology, 36, 49–67.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. Philosophy of Science, 67(1), 1–25.
Manzo, G. (2021). Does analytical sociology practice what it preaches? An assessment of analytical sociology through the Merton award. In G. Manzo (Ed.), Research Handbook on Analytical Sociology (pp. 1–47). Edward Elgar Publishing.

Null-hypothesis testing in preprints

Mikkel Willum Johansen
University of Copenhagen
mwj@ind.ku.dk

Helene Vindmar
University of Copenhagen
tkz227@ind.ku.dk

In 1991, arXiv was established as the first digital preprint server. Initially focused on physics, arXiv later expanded its scope to include mathematics, computer science, and related fields Ginsparg P. (2021). Other research disciplines were slower to adopt preprint servers, but the 2010s saw the emergence of several field-specific platforms, including bioRxiv (2013), ChemRxiv (2017), and medRxiv (2019). Today, preprints are widely used across all STEM research fields.
The increased use of preprints has fundamentally changed scientific practices related to the dissemination of results. Preprints are increasingly assuming the role once played by peer-reviewed papers in communicating findings within academia and to the public. While they enable rapid and open dissemination, this speed comes at a cost. The vetting and quality control associated with traditional research papers are reduced or nonexistent for preprints, prompting debates about their trustworthiness (e.g., Ni, R., & Waltman 2024; Browman et al 2025).
In this talk, we contribute to this discussion by investigating the quality of preprints from three biomedical preprint servers: medRxiv, Research Square, and Preprints.org. Based on existing investigations of publication rates in scholarly journals for preprints from these servers, we expected medRxiv preprints to be of high quality, and preprints from Research Square and Preprints.org to be of middle or low quality (Drzymalla et al., 2022).
To focus the investigation, we examined the reporting of null hypothesis testing (NHT), a fundamental statistical tool, in preprints published in the second quarter of 2025. In total, approximately 27,000 preprints were published on the three servers during this period. We investigated the sample using a mixed-methods approach combining large-scale quantitative investigations with qualitative close reading of selected preprints. First, all preprints in the sample were scanned using a Python script to determine whether they used NHT. This provided an estimate of the prevalence of this statistical tool. Then, 30 preprints using NHT were randomly selected from each server and subjected to qualitative close reading. For each preprint, we assessed whether it contained the necessary information needed to interpret the

stated test results (e.g., significance level, statistical model) and whether it included relevant methodological reflections concerning the tests (e.g., justifications for choosing a particular significance level or model).
In the talk, we will present results of the investigation and discuss implications for scientific practices related to the use of preprints.

References

1. Browman, H., Leopold, S. S., & Anderson, K. R. (2025). Preprinting does not meet science's duty of care responsibility to society. Journal of Trial and Error. https://doi.org/10.36850/b6fa-4513
2. Drzymalla, E., Yu, W., Khoury, M.J. et al (2022). COVID-19-Related manuscripts: lag from preprint to publication. BMC Res Notes 15, 340. https://doi.org/10.1186/s13104-022-06231-9
3. Ginsparg P. (2021). Lessons from arXiv's 30 years of information sharing. Nature reviews Physics, 3(9), 602–603. https://doi.org/10.1038/s42254-021-00360-z
4. Ni, R., & Waltman, L. (2024). To preprint or not to preprint: A global researcher survey. Journal of the Association for Information Science and Technology, 75(6), 749–766. https://doi.org/10.1002/asi.24880

Coordination and Instrumentation in the Discovery of Seasonal Fluctuations in the Length of the Day

Shaul Katzir
Tel Aviv University
skatzir@tau.ac.il

What is a measurement and what is measured is not always trivial. In 1934 Adolph Scheibe and Udo Adelsberger assumed that they were measuring the rate of their new quartz clocks by comparing them to astronomical time. By analysing their results, however, they inferred that they were actually measuring the rate of the earth's daily rotation, and concluded that it fluctuates along the year. In other words, they transformed their yardstick into the object of study, supposing that their clocks are actually steadier than terrestrial rotation, as manifested by astronomical time.
Known as the coordination problem, the question of establishing which of the phenomena or instruments is constant (or steadier than the other) has received philosophical attention. However, the specific type in which the researchers exchange the standards during the research has not gained much attention. Moreover, unlike most of the other examples in the literature, this case involved two very different phenomena: the quartz clock could be manipulated, whereas the earth's rotation could not. Theoretical considerations could not help (as it is usually the case); there was no strong argument for the invariability of either. Actually the rate of both were known to drift, and the question was whether their rate corrected for their empirically known drift was steady, and how reliable the new instruments were.
To establish the steadiness of the quartz clock, Scheibe and Adelsberger relied on the agreement among their clocks, which were based on two different mechanisms and should therefore be dissimilarly affected by common physical influences. Shortly later, Nicolaus Stoyko strengthened this argument by demonstrating that their clocks agreed with pendulum

clocks based on an entirely independent physical process. The epistemic power of this kind of inner agreement is known also in other cases (e.g. temperature). Following a classical common cause argument, the assumption of fluctuations in terrestrial rotation was preferred over that of separate causes leading to the same deviation in three distinct periodic motions that control the clocks. This indicated that the deviation was of the astronomical rather than the clock time.
The advantages and limitations of the common cause inference will be discussed by comparing this case of measurement instruments with the case of the deviations of the theoretical motions of a few planets from observation in the previous decade. In the latter case, establishing a correlation between the planets pose a challenge, whose difficulties along with the wish to employ the common cause explanation had led to erroneous conclusions. Criticism of the way it was applied to reach these conclusions indicates that there are no a priori criteria for successfully implementing the common cause argument.
The irregularity of the earth's rotation showed that it could not be used to define time. Neither could the quartz clock play that role since each resonator that controlled the clock was different from the others, and in addition its rate drifted.

Participatory Research as Epistemic Repair: The Case of R2D2-MH

Daan Kenis
University of Antwerp
daan.kenis@uantwerpen.be

Patient Engagement (hereafter, PE), a specific form of participatory research, has proliferated in data-intensive biomedicine (Beier et al. 2018). Across policy contexts, incentives to include extra-academic stakeholders aim to ensure that the epistemic goods produced are relevant and trustworthy. In this sense, PE is often explicitly promoted as addressing historical epistemic injustices (Fricker 2007; Grasswick 2017) – redressing the exclusion of marginalized communities from biomedical knowledge production and restoring damaged epistemic trust relations. Yet, despite their emancipatory intent, such top-down mandates for diversity in science often are ineffective (Koskinen 2022); they tend to reproduce rather than address epistemic injustices (Madden and Speed 2017). This is particularly alarming for data-intensive research, where – as the recent controversy surrounding the Autism Research Centre's Spectrum10K signaled – sentiments of epistemic distrust may inform communal decisions to disengage with particular types of research completely. This raises a pressing question: how must PE be structured to promote epistemic justice?
Drawing on field observations in a participatory preclinical project in autism genetics (R2D2-MH), I argue that the effectiveness of PE in promoting epistemic justice depends on whether it targets the right level of intervention. Existing approaches tend to focus primarily on procedural reforms – altering rules of engagement for including marginalized voices in research decisions (Arnaud 2025; Catala 2024). These are necessary but insufficient in contexts marked by histories of epistemic exclusion. What is damaged in such contexts are not simply discrete epistemic goods, but the conditions for epistemic relations between research and participant communities. Borrowing from restorative justice theory (Almassi 2018; Altanian 2024; Walker 2006), I argue that epistemically just participatory research must commit to both procedural and substantive requirements. PE should proceed from genuine engagement with community interests and values, and in the process, identify and

substantively target the epistemic norms and infrastructures that sustain problematic epistemic relations.
The R2D2-MH case illustrates how both conditions may unfold. Here, a central outcome is the co-design of an explicitly neurodiversity-affirmative ICF Core Set for data harmonization, positioned as an "authoritative" epistemic resource to "facilitate neurodiversity-affirmative investigation" (Black et al. 2025). This infrastructural intervention addresses the abundance of deficit-focused cohort data and standards, and creates conditions for future research to proceed along neurodiversity-affirmative values. In addition to engaging the autistic community in co-designing the Core Set, the project provides substantive repair by renegotiating embedded exclusionary epistemic norms; specifically, the classification schemes and data standards that govern how autism can be studied. To the extent the Core Set finds uptake in the broader research community, the proposed data infrastructure may function to restore epistemic relations at a systemic level.
This paper then contributes to the philosophy of participatory science by revealing an underexplored dimension of value. Rather than assessing participatory research on its political legitimacy, or by its capacity to be knowledge-tracking (Potochnik 2024), this lens recognizes its knowledge-promoting role (Currie 2019): PE may also repair damaged epistemic relations and restore trustworthiness, thereby creating conditions for sustainable and just knowledge production.

Putting a leash on homeostasis with constraints from scientific practice

Jan Pieter Konsman
ImmunoConcept, CNRS & University of Bordeaux
jan-pieter.konsman@u-bordeaux.fr

Martin Zach
Institute of Philosophy of the Czech Academy of Sciences
zach@flu.cas.cz

Homeostasis has been coined by Walter Cannon in the early 20th century to refer to “the coordinated physiological reactions which maintain most of the steady states in the [animal] body” (Cannon 1929). It has “played the role of a central, all-encompassing phenomenon forming the principal object of physiology” and continues to be influential beyond this field of study (Lemoine & Pradeu, 2018). The mechanisms by which organisms maintain themselves have long been thought to involve some kind of balance of opposing forces, initially between humors for the ancient Greeks and later, for example for Cannon, between sympathetic and vagus or parasympathetic nervous systems. After Cannon, the cybernetic movement provided a type of explanation for the relative stability of living systems in the form of negative feedback loops that could involve one detected variable without the need to postulate opposing variables. Such explanations of homeostasis have become very popular, up to a point that nowadays homeostasis is often equated with the existence of negative feedback loops. In contrast, and possibly, in response, a plethora of other concepts, such as allostasis, rheostasis, etc., have since been put forward to explain biological regulation in metazoans.
Moreover, It has recently been argued (i) that the cybernetics view does not do justice to the notion of homeostasis, and (ii) that instead of adopting alternative notions it would be better to adopt a broader vision of homeostasis. The latter would subsume alternative concepts

while providing an organismal perspective rather than a focus on molecular parameters (Bechtel & Bich, 2024; 2025). Indeed, Bechtel and Bich have been rejecting individual variables and the notion of set points for understanding homeostasis.
While we in general agree with some specific points, we take issue with such an approach to homeostasis. Our argument is two-fold. First, we argue that Bechtel and Bich's interpretation of Cannon as reflecting a whole organism perspective needs to take into account that Cannon considered individual variables Second, even if this issue is resolved, their insistence on using homeostasis as an umbrella concept risks obscuring the diversity of biological responses playing a role in biological regulation. We use the example of inflammatory responses in acute infection and their role in biological regulation to show why the notion of homeostasis cannot be expanded as Bechtel and Bich would have us believe.
Finally, we argue that Bechtel and Bich's insistence on the organism level of analysis for homeostasis may lead to the dismissal of scientific discussions concerning cell and tissue-level homeostasis, and overlooking important questions about how variables at different levels contribute to overall maintenance, thereby narrowing the scope of inquiry to the organism alone.

Philosophers and Scientists in Practice: Building local research cultures with field philosophy

Anatolii Kozlov
Université de Genève, Département de Philosophie
anatoly.kozlov@unige.ch

Helene Scott-Fordsmand
University College London, STS
h.scott-fordsmand@ucl.ac.uk

In its initial formulation, the PSP movement aimed not only "to learn from the experiences and understandings of [scientists] but encourage them to become more reflective about their practices, as well as the underlying assumptions and the implications of these practices" (Ankeny et al. 2011). This latter intention has manifested, for example, in work under the banner PinS as 'philosophy-of-science in practice' (see Kosolosky 2012) where philosophers participate in research, particularly biomedical and biological research, working with scientists on conceptual questions and co-publishing in recognized science journals (Pradeu et al. 2024). While these tight collaborations have fostered important contributions both to core concepts in philosophy of the life sciences and to the life sciences themselves, the reflectivity that arises from these partnerships has mainly been about the content of the scientific work, and much less about the practices themselves.
In this paper, we suggest that there are meaningful ways to combine the two aims: to philosophically explore aspects of scientific practice and to foster scientists' reflexivity about them. We describe, [name of intervention], a 1-year 'field' intervention at a biomedical department in a big UK university from 2024/25. The project was led by us, two philosophers of science, and funded as an intervention for the improvement of the local research culture. The core of the intervention was a series of workshops set up to provide a space for collaborative, reflexive exchange between postdoctoral researchers. These workshops were organised around topics from philosophy of science (e.g., emotions and ideals in science). To facilitate this pedagogical field philosophy work, we used creative and embodied methods

under the modality of "serious fun", as this allowed a lowering of disciplinary expectations and reliance on jargon amongst participants, bringing everyone onto a level playing field. In the paper, we provide an epistemic narrative of the project throughout its course, including two spin-off interventions in neighbouring departments. Drawing on Wylie's four forms of field philosophy – (1) facilitative practices; (2) technical contributions; (3) embedded field philosophy; and (4) transgressive ("gadfly") field philosophy of science (Wylie 2025) – we highlight how this intervention mobilised philosophy of science in order to co-create reflective work on the local research culture and built a better, shared understanding of the people and the communities engaged in producing science. Halting at some empirical examples from the workshops, we demonstrate how this work provides input on central questions in PSP, for example, on the epistemic functioning of biomedical science, on assumptions on the role and identity of 'a scientist', on the epistemic agency of a scientist within the structure of a biomedical lab, as well as on the value of 'unstructured time'. Ending with the spin-off interventions in the neighbouring departments, we outline a few meta-reflexive points about the research culture in philosophy of science, and the beneficial reflexive lessons on our own disciplinary practices that we took from our work in this intervention.

Bibliography

Ankeny, R., H. Chang, M. Boumans and M. Boon (2011). Introduction: philosophy of science in practice. European Journal for Philosophy of Science 1(3), 303-307.
Kosolosky, L. (2012). Philosophy-of-Science in Practice vs. Philosophy of Science-in-Practice. Newsletter SPSP, Winter 2012, 9-10.
Pradeu, T., M. Lemoine, M. Khelfaoui, and Y. Gingras (2024). Philosophy in Science: Can Philosophers of Science Permeate through Science and Produce Scientific Knowledge?. The British Journal for the Philosophy of Science, 75(2), 375-416.
Wylie, Alison. 2025. 'Field Philosophy of Science: Externally Motivated and Externally Engaged'. In Methods in the Philosophy of Science: A User's Guide, edited by Sophie J. Veigl and Adrian Currie. MIT Press.

Measurement as epistemic scaffolding: The role of patient-centered measurement in clinical practice

Yun-Ying Kuo
University of South Carolina, Department of Philosophy
yunyingkuo17@gmail.com

Measurement in medicine is value-laden and context-dependent, raising challenges for standardization and objectivity. While philosophers of measurement have debated representational accuracy and the iterative measurement practice of research (Alexandrova, 2017; Chang, 2022 ; Rodriguez Duque et al., 2024), few accounts examine how measurement functions within clinical practice. In this paper, I argue that patient-centered measurement in clinical practice should be understood as epistemic scaffolding: a temporary and structured support for inquiry, interpretation, and coordination under epistemic and ethical uncertainty. To development my argument, first, I situate this proposal in relation to two dominant philosophical concerns. On the one hand, Hausman (2015) has argued that constructs like well-being are too heterogeneous to be measured, especially for purposes of health policy. On

the other hand, Alexandrova (2017) has responded by advocating contextual or mid-level theories of well-being, constructs tailored to specific populations or conditions, thereby restoring measurability to a specific groups of people. Yet both positions remain focused on representational adequacy. Drawing on McClimans (2024)'s view, I shift attention from the outcome of measurement to the cognitive and epistemic processes of measurement: facilitating epistemic dialogue between patients and clinicians.
Second, I argue that clinical practice is understood as an epistemic but also ethical activity. It is a practice where knowledge about illness, patient experience, and therapeutic options is co-constructed through clinician-patient interaction. In this context, patient-centered measurements function as temporary and structured supports that enable clinicians and patients to engage in inquiry, coordinate perspectives, and arrive at shared understanding. These temporary but structured supports are what I call epistemic scaffolding. Borrowing from Feng Lin & Sadhana Puntambekar (2024)'s theories in science education, epistemic scaffolding is used to support learners through the learning process but is removable once the epistemic sophistication is stable. In clinical practice, the epistemic scaffolding is important in the first few times of clinical encounters before mutual understanding is firmly established. These scaffoldings can be adjusted or removed as dialogue unfolds and inquiry becomes more individualized to the clinician-patient interaction.
Lastly, I argue that if we conceive measurement in clinical practice as epistemic scaffolding, it offers three key advantages: (1) It facilitates epistemic dialogue, enabling shared understanding between patients and clinicians. (2) It allows for holistic and interpretive engagement, responsive to heterogeneity. (3) It retains pragmatic utility in structuring therapeutic discussion and guiding care. Through a case study on the Edmonton Symptom Assessment Scale (ESAS) in palliative care, I show how measurement as epistemic scaffolding functions, offering both a descriptive and normative account of its role in clinical reasoning and care.

Rethinking Improvements to the Clonal Evolution Model Through Its Epistemic Purposes

Anaïs Lamoureux
CNRS, UMR8590 IHPST, University Paris 1 Panthéon-Sorbonne, Paris, France
anais.lamoureux@etu.univ-paris1.fr

The clonal evolution model is the dominant model for explaining cancer progression (Greaves & Maley, 2012). It propose that tumor cells form an heterogenous population through the accumulation of genetic and epigenetic alterations that modulate their relative fitness. This heterogeneity allows for expansion or extinction via natural selection. The model is facing both scientific and philosophical criticism. Recent literature has emphasized its limitations and the need for conceptual refinement, whether through the incorporation of non-genetic mechanisms of variation (Landau et al., 2014), the integration of additional evolutionary processes such as neutral evolution (Williams et al., 2016) or multilevel selection (Laplane et al., 2025; Lean & Plutynski, 2016), or the recognition that natural selection alone may be insufficient to account for the apparent complexity of adaptation in cancer (Germain, 2012). These interventions have produced a numerous modified versions of the clonal evolution model. Those versions coexist without clear criteria for determining whether such modifications genuinely improve the model. This article examines how to genuinely improve the clonal evolution model ? We argue that modifications to the model are

beneficial only relative to the specific epistemic purpose the model is intended to support. This view is supported by the adequacy-for-purpose model evaluation (Parker, 2010, 2020; Currie, 2017). In this framework a modification that enhances one epistemic purpose may be irrelevant or even detrimental to another. In this work, we identify four such roles that the clonal evolution model currently fulfills in scientific practice: explanatory, either as an explanans, when the model is used to explain cancer cell progression or as an explanandum, when the model itself is in need to be explained. There are also non-explanatory purposes : descriptive, when the model functions as a classificatory or representational schema; and applicative, when it is considered as true and used to produce applications such as treatment. Each purpose involves its own conditions for what counts as beneficial modification. For instance, when the model is used as an explanans for the metastatic process, revisions that extend its explanatory resources, such as incorporating mesenchymal–epithelial transitions to account for metastasis, may be warranted. Conversely, additions that lack causal relevance to the target phenomenon, such as parameters that do not affect tumor dynamics, do not constitute epistemically valuable improvements. Acknowledging these multiple purposes makes clear that many critiques arise when a given modification is evaluated against purposes it was never intended to meet. The article's second part reconstructs the historical evolution of these epistemic purposes, tracing how shifts in scientific practice, ranging from early theoretical formulations to the rise of genomic profiling and contemporary ecological (Maley et al., 2017) or systems-oriented perspectives (Plutynski & Bertolaso, 2018) have guided successive reconceptualizations of the clonal evolution model. This analysis reveals that the model's transformation is driven not only by empirical discoveries but also by changing epistemic expectations. We conclude by offering a set of conditions for assessing and guiding model modification in light of the purposes it is meant to serve.

Model-based interpretations of field study results in geology and the risk of idealization-alignment

Clément Lasselin
Lund University
clement.lasselin@tg.lth.se

Alexis Geisler-Roblin
École Normale Supérieure
alexis.geisler-roblin@ens.psl.eu

The philosophy of science has extensively examined the role of models in contemporary scientific practice (Bailer-Jones, 2009; Godfrey-Smith, 2007; Knuuttila et al., 2024). A prominent research trend investigates the use of models in geosciences (Bokulich, 2021; Bokulich & Oreskes, 2017; Ohnesorge & Watkins, 2024). This observation applies to geological models used in subfields such as geochemistry, hydrogeology, or seismology (Linde et al., 2017; Nearing et al., 2016; Zanetti et al., 2023). Geology indeed provides valuable case studies for the philosophical study of models as it investigates systems that are complex—featuring multiple interacting components and subsystem dynamics—and that are difficult to access empirically (Craig-Thompson & Kuchler, 2025). Even when conducting field experiments, geologists must rely on subsystem-specific, indirect measurements (e.g., place-specific core samples) which create opportunities for errors and require researchers to interpret results through empirically under-determined hypotheses. In this context, models

frequently serve to interpret limited field results. These models are typically grounded in previous field investigations and laboratory experiments and may take the form of numerical simulations of geological dynamics. They can be extremely helpful in interpreting field results while providing an appearance of systematicity (Bokulich, 2020). However, models are constructed with parameters and input data that reflect modelers' choices, which are in turn influenced by a range of factors—including consistency with existing data, disciplinary conventions, and resource availability. This often results in models that are highly idealized and that bear non-trivial relationships to truth (Spagnesi, 2025). For instance, three-dimensional geological models may consist of single deterministic representations that do not fully account for the full range of plausible geological dynamics under study (Madsen et al., 2022).

The paper seeks to demonstrate how the combination of (1) field results that have high evidential value but remain limited, (2) models necessary for interpreting those field results, and (3) models that are highly idealized, can create pressures to align of field result interpretations to the constraints of model idealizations. This phenomenon is referred to as "idealization-alignment", defined as a specific form of model-based inference bias (Cartwright, 1983). This bias involves adjusting interpretations to preserve field evidential value while advantages of model use—echoing classical theories of the "theory-ladenness" of observation (Hanson, 1985). The paper argues that idealization-alignment may occur, intentionally or unintentionally, as a result of scientists' desires to produce coherent interpretations of field results, to report findings that appear robust and non-disputable, or to maintain model stability and reusability across studies. Beyond strictly epistemic motivations, there may be non-epistemic values at stake—particularly in geology, a discipline historically intertwined with industrial and policy interests. Notably, idealization-alignment can preserve the evidential value of field results while directing interpretation in ways that align with models chosen in part to support certain practical interests. This argument is illustrated with examples from geological engineering research related, particularly studies on nuclear waste storage and carbon storage, where such practical interests can significantly influence modeling choices and field result interpretations.

The Ontology of Engineering a Quantum Computer: Connecting the Interpretations of Quantum Mechanics to Engineering Practice

Thijs Latten
TU Delft
t.m.k.latten@tudelft.nl

No consensus exists on how to interpret quantum mechanics, and the theory is often discussed as difficult or impossible to understand (e.g., Feynman, 1985). Yet physicists and engineers in quantum technologies (such as quantum computing) are finding innovative ways to actively create, manipulate, and exploit quantum behaviour for practical benefit. While there are increasing calls in the philosophy of science to bring our ontology into closer contact with successful experimental and technological practice (Chang, 2009, 2022; Hacking, 1983; Vermaas, 2005), the practical success of quantum mechanics is rarely discussed in relation to the expressed difficulty in interpreting quantum mechanics. In this paper, I explore one way to bring the largely empirically equivalent interpretations of quantum mechanics closer to the empirical world: I assess the ontological assumptions of engineers who aim to build a quantum computer.

Hasok Chang argues that for any activity to be intelligible, certain ontological principles need to be assumed (Chang, 2009, 2022). These ontological principles are construed as necessary preconditions for the intelligibility of an activity, such as how the activity of narration presupposes persistence of entities over time, or how intervening only makes sense if one presupposes some notion of causality. I analyse a successful engineering practice for physically realising qubits, namely, quantum dot spin qubits, and reflect on the ontological principles that are assumed in some of the activities of these engineers.
I highlight a prima facie tension between a common explicit interpretational commitment of engineers (textbook quantum mechanics) and some of the presuppositions in the practices these engineers engage in (such as the presumed persistence of electrons through the process of initialisation, manipulation, and readout). Although some interpretations are potentially better suited to support the operative ontology over others, I note that the debate on the interpretations of quantum mechanics seems largely not set up to accommodate the operative ontology of engineers. This study opens up a way to bridge this gap between the interpretations of quantum mechanics and the technological success and formulates a challenge to the interpretations of quantum mechanics to accommodate the ontological assumptions of engineers. Lastly, I warn against a literalist interpretation of the operative ontology that arises from the practices and argue that much care is needed to avoid building an ontology on outdated forms of understanding.

Getting the Antenna Inside the Theory: How the CMB Was Detected Before It Was Discovered

Anastasiia Lazutkina
University of Wuppertal
lazutkina@uni-wuppertal.de

In philosophy of science, empirical evidence is often treated either abstractly through formal accounts of confirmation and inference (Earman 1992; Howson and Urbach 2006; Mayo 1996), or else through detailed accounts of how data are stabilized and structured so as to play an evidential role (Leonelli 2016; Boyd 2018). What is less often spelled out is the mechanism that allows structured data to become comparable with theoretical models in the first place. Boyd (2018) identifies resources needed to adjudicate a theory's adequacy with respect to evidence while also stressing that such resources are not sufficient on their own. Building on this, I propose that what is missing in current accounts of evidence is precisely what Stein (1994) and Curiel (forthcoming) call schematization of the observer: "in a model of an experiment, to provide a representation of something like a measuring apparatus, even if only of the simplest and most abstract form, that allows us to interpret the model as a model of an experiment or an observation" (Curiel forthcoming). Consequently, calibration practices, background subtraction, and rules for treating residuals help to constitute evidential relations rather than merely refine them.
I develop this claim through a reconstruction of the historical context of one of the most significant discoveries of the twentieth century: the detection of the cosmic microwave background (CMB). The CMB is an isotropic blackbody radiation that was predicted by Big Bang cosmology already in the 1940s but only discovered in 1965. I show that the same radiation was in fact measured several times between the 1940s and 1950s but was not interpreted cosmologically (Kragh 1996). What changed in 1965 that allowed the same kind of measurement to count as evidence for a cosmological theory?

I argue that this transformation was twofold. First, it involved what Boyd (2018) calls enrichment: the careful work by Penzias and Wilson (1965) that made their residual antenna temperature a stable, reproducible result. Second, it required what the paper by Dicke, Peebles, Roll, and Wilkinson (1965) supplied: a schematization of the observer that defined where and how to look for the CMB, what kind of apparatus would register it, and what would count as contamination or noise. Only when these two layers, empirical enrichment and observer schematization, were combined did the result become comparable with the theoretical model of relic radiation from the early universe. In Stein's terms, the CMB case illustrates how the "laboratory" must be integrated into the theoretical framework for data and theory to become genuinely comparable, so that evidential relations can be established at all.

Progress in Psychology: A Practice-First Reconstruction

Yen-Tung Lee
Academia Sinica
ylee734@uwo.ca

Karen Yan
National Yang Ming Chiao Tung University
karenyan@nycu.edu.tw

Feest (2022) rejects the standard model of scientific progress—accumulating true theories about natural kinds—arguing that psychology advances instead through accumulating epistemic resources for conceptual development. Her framework rests on three interconnected concepts. First, research objects in psychology—capacities like memory, attention, and perception—are not mind-independent natural kinds but historical kinds whose identity conditions are partially constituted by evolving socio-historical contexts and folk-psychological practices. Second, minimal conceptual continuity is maintained across theoretical changes because folk-psychological vocabulary continuously individuates the research objects, providing stability despite theoretical shifts. Third, epistemic resources—the totality of theories, methods, instruments, and experimental paradigms developed throughout the history of psychology—accumulate over time, enabling researchers to track and conceptually develop their understanding of psychology's research objects.
However, Feest's framework faces three critical challenges. First, her account of conceptual continuity remains fundamentally ambiguous—does it refer merely to vocabulary persistence (trivially true, as languages are continuous) or to genuine referential continuity (which she has not established)? Second, if it is to genuine referential continuity, she owes us an account of what the stable referents of research objects could be, given her commitment to how research objects are "socio-historically contingent" (2022, 195). Third, most problematically, her characterization of epistemic resources is circular: resources are deemed epistemic simply by virtue of their involvement in psychological research, providing no independent criterion for determining which resources genuinely contribute to progress.
This paper proposes a practice-first reconstruction of progress in psychology that addresses these challenges while preserving Feest's core insights. Drawing on recent philosophical work on experimental individuation, we argue that we should replace Feest's what-questions ("What are psychological objects?") with how-questions ("How do psychologists experimentally individuate their research objects?"). Instead of asking what attention or

consciousness is, we should examine how these phenomena are operationally separated from background processes, systematically manipulated in controlled interventions, and maintained as unified objects across experimental contexts.
This practice-first approach dissolves Feest's theoretical tensions. Research objects are neither mind-independent natural kinds nor mere social constructions, but process-stable entities successfully individuated through experimental practices. Their reality is neither a metaphysical presupposition nor a social convention, but an experimental achievement. Conceptual continuity should be measured not by vocabulary persistence but by continuity or translatability of individuation practices across research programs. When James, Broadbent, and Posner all study "attention," we can trace how their experimental operations evolved—from introspective selection tasks to dichotic listening paradigms to spatial cueing—revealing both genuine continuities in individuation practices and significant ruptures where new methods enabled more stable individuation. Epistemic resources are tools that support more stable, reliable, or fine-grained individuation practices, providing a non-circular criterion: a method becomes an epistemic resource when and because it successfully enables experimental individuation, not merely because it is used in research.
Through historical case studies tracing attention research from James through cognitive neuroscience, we demonstrate how this practice-first framework better captures psychology's genuine achievements and provides actionable criteria for evaluating scientific progress without presupposing fixed natural kinds or lapsing into social constructivism.

Can conceptual analysis be scientifically relevant?

Maël Lemoine
Université de Bordeaux
mael.lemoine@u-bordeaux.fr

"Philosophy in Science" (PinS) has recently been introduced as the attempt to do science with philosophy, which involves meeting 3 necessary, and jointly sufficient, conditions: "raising a scientific problem, using philosophical tools to address it, and making a scientific proposal" (Pradeu et al. 2024). A key distinction is between a scientific proposal and a scientific contribution. In short, a proposal becomes a contribution only if it "happens to be accepted as useful by scientists", that is, with "the explicit and substantial recognition, by scientists, that a philosophical work advances science." (Pradeu et al. 2024). It is fair to object that what truly matters, to define contributive PinS, is not recognition by a scientific community, as characterized by an externalist criterion (e.g., bibliometric), but production of novel scientific knowledge, as defined in an internalist perspective, which may come to be ignored by scientists for bad reasons. So, even if Pradeu and colleagues are right to claim that such recognition is the best criterion to detect contributive PinS, it is not the best way to define 'contribution'. In the present paper, I focus on how scientifically contributive conceptual analysis, philosophers' typical tool, can be. While it is fair to say that "It would be absurd to consider all of the almost infinite instances of conceptual analysis proposed by philosophers of science as scientific contributions" (Pradeu et al. 2024), a difference should be made between a proposed conceptual analysis that cannot produce knowledge and one that can. I propose and examine the hypothesis that conceptual analysis may or may not produce scientific knowledge depending on whether it reduces the contrast class of a scientific question, that is, the set of expected answers within a scientific field. Scrutinizing a paradigmatic example of contributive conceptual analysis according to Pradeu and

colleagues, "Individuality and Selection" (Hull 1980), I will show why this influential paper does not, properly speaking, constitute a contribution to scientific knowledge, although it constitute a contribution in the "recognition" sense. I will conclude with phrasing some conditions for conceptual analysis to successfully contribute to scientific knowledge, which is not its only application, but definitely one of its potential venues.

Demarcation in Action: Negotiating the Scope of Scientific Cosmology, 1948–1965

Nichole Levesley
Department of History & Philosophy of Science, University of Cambridge
nml46@cam.ac.uk

A topical issue in the philosophy of cosmology concerns the scope of cosmology: what problems and questions are within the bounds of scientific inquiry. Scholars have recently been careful to emphasize the limits of what cosmology can assuredly claim, for instance arguing that inquiries into the nature of the laws of physics, boundary conditions, and the multiverse belong to philosophy rather than science. What has sometimes been overlooked, however, is that the boundaries between philosophy and science are not static: some issues that were once considered beyond the scope of scientific investigation have since come to form essential parts of our current worldview, and there are various ways the boundaries of science can shift, many of which are not linear, frictionless transpositions of a problem from 'metaphysics' to 'science'.

This paper provides a historical perspective on the changing nature of the limits of scientific knowledge by investigating the founding of physical cosmology in the mid-twentieth century, identifying the practices by which physicists delineated the scope of cosmological research. In the late 1940s, the scientific community largely agreed that questions about the age and large-scale development of the universe were unanswerable scientifically; but by the mid-1960s, the hot big bang theory was accepted as a philosophically unproblematic part of the cosmological worldview. This change came about in part due to a shift in perceptions of the limits of cosmological knowledge. In the first half of the presentation, I explicate two of these limits — the extrapolation and creation limits — and chart the practices by which they were introduced, their applications in cosmological research, and how physicists eventually negotiated their resolutions. In the second half, I draw parallels between these discussions and those concerning the scope of cosmology today, arguing that the problems faced by current cosmologists are not new and that an analysis of historical actors' actions can inform our current debates.

The intended upshot is twofold. To philosophers of cosmology and practicing cosmologists, the paper aims to demonstrate that the scope of science has previously changed in more complicated, multidirectional ways than is sometimes assumed; we should neither view the boundaries of science as static, nor consider the scope of cosmology to expand linearly. More generally, I hope to illustrate that accounts attuned to scientists' practices and activities can enhance discussions of the demarcation problem and limits of science: the scopes of scientific disciplines are not given, but constantly and actively negotiated by practicing scientists' pursuits.

Epistemic iteration and measurement of aging [Paper/Poster]

Amirreza Liaghat
Bordeaux University
amirreza.liaghat@u-bordeaux.fr

Epistemic iteration, as described by Chang in "Inventing Temperature" (1), is a process in which imperfect concepts, measurements, or classifications are progressively improved through repeated cycles of use and revision. Later, in his article Epistemic Iteration and Natural Kinds (2), Hasok Chang shows how the method of epistemic iteration can be usefully applied to fields such as medicine and in particular, psychiatry. Importantly, Chang emphasises that in empirical sciences, iterations often do not converge on a single point. Instead, they tend to stabilise at what can be described as local minima. Chang argues locally stable solutions can still be epistemically productive and practically effective, even if they fall short of idealised notions of convergence.
In this article, I propose to test Chang's assessment of local minima by applying it to the contemporary measurement of aging. Due to the growing prevalence of age-related diseases, a wide range of methods has been developed to measure biological aging with the aim of prediction, prevention, and intervention. These include biological age estimators based on biomarkers, epigenetic clocks, functional assessments, and clinical frailty indices.
I argue that the case of aging poses certain difficulties to Chang's claim that convergence on local minima is "good enough." Aging is a complex phenomenon, characterised by heterogeneous trajectories, population-level variation, and multiple irreducible levels of description, ranging from molecular damage and cellular senescence to organismal function and social context. Moreover, the history of aging measurement can show us how different values have been applied to this field. Therefore, different measurement practices often stabilise around distinct local minima that are not merely complementary, but partially incommensurable. When we adopt a pluralistic approach (3) to theories of aging, the accumulation of such local minima does not necessarily yield epistemic progress. Instead, it risks producing a fragmented landscape of measures that cannot be straightforwardly integrated or compared, thereby limiting their practical and explanatory value.
To address these problems, I propose to draw on the concept of Pins (philosophy in science), developed by Pradeu and Lemoine (4). Using philosophical tools, this method allows us to distinguish among different ways of moving between levels to make sense of convergence on multiple endpoints. By identifying which passages are epistemically legitimate, we can make sense of progress in aging research through relying on the additive accumulation of local minima. In doing so, I aim to show how epistemic iteration can be applied, rather than rejected, to better accommodate pluralism and complexity in the biomedical sciences.

1- Chang, Hasok.¬†Inventing temperature: Measurement and scientific progress. Oxford University Press, 2004.
2- Chang, Hasok. "Epistemic iteration and natural kinds: Realism and pluralism in taxonomy."¬†Philosophical issues in psychiatry IV: Psychiatric nosology¬†(2017): 229-245.
3- Mc Auley, Mark T. "The evolution of ageing: classic theories and emerging ideas."¬†Biogerontology¬†vol. 26,1 6. 29 Oct. 2024,
4- Pradeu, Thomas, et al. "Philosophy in science: Can philosophers of science permeate through science and produce scientific knowledge?."¬†The British Journal for the Philosophy of Science¬†75.2 (2024): 375-416.

Geopolitics, diplomacy and international collaboration: expanding the boundaries of HPS in ocean sciences

Pablo Lima
University of Seville - ERC DEEPMED project
plima@us.es

The dynamics of interdisciplinary and international cooperation have long been a topic of concern among historians and philosophers of science. Several epistemological frameworks, ranging from 'trading zones' (Galison, 1997) to 'repertoires' (Ankeny & Leonelli, 2016), have been proposed to analyse how experts, scientists and technicians from diverse research cultures, intellectual traditions, and academic backgrounds have created shared instruments of dialogue to assess unprecedentedly complex research problems. However, the ways in which the diplomatic and geopolitical dimensions of these collaborations have shaped the research methods and practices of the scientists involved remains understudied. In this paper, I delve into the history of international oceanographic cooperation in the Eastern Mediterranean Sea over the last third of the 20th century to explore how the existent network of international relations, institutional connections, and national interests constrained both the execution of such collaborations and their resulting research outputs.

After the Second World War, the institutionalization of 'big' science constituted a decisive epistemic shift towards an era of large-scale collaboration that prompted methodological innovation across disciplines. In the field of oceanography, this shift became particularly pronounced following the establishment of dedicated science governance bodies, such as the Intergovernmental Oceanographic Commission (IOC) of UNESCO in 1960. Nevertheless, the 'big' oceanographic enterprises took decades to reach regional seas like the Mediterranean, which was marked by large technoscientific gaps and geopolitical tensions among its bordering countries. The situation changed with the launch of the IOC's Physical Oceanography of the Eastern Mediterranean (POEM, 1982-1995) program, which, for the first time, coordinated modelling and fieldwork across national boundaries to produce the first fine-grained numerical circulation models of the Eastern Mediterranean basin.

The modelling and data practices deployed by POEM participants spanned from the coordinated collection of physical, chemical and ecological data of the Eastern Mediterranean ecosystem aboard multiple cruises to their subsequent processing, calibration, standardization and assimilation into interdisciplinary oceanographic models. I will show how the data journeys (Halfmann, 2020) undertaken within the POEM project and their symbiotic relation to model-building (Edwards, 2010) were subjected to the web of interests of the actors involved, from the predictive purposes of the US Navy to the needs of UNESCO and other countries to protect the Mediterranean against growing ecological threats. Each participating nation, research group and institution attached a different epistemic meaning to the data, and such plurality of aims was reflected in the final modelling output of the project.

Seeking to move beyond the current methodological boundaries of integrated history and philosophy of science, I bring insights from geopolitics and science diplomacy—long central to histories of oceanography (Oreskes, 2021)—to bear on the philosophy of international and interdisciplinary scientific collaboration. In line with recent work showing that oceans provide privileged sites for examining the intertwinement of scientific and non-scientific values and forms of knowledge (Jones, 2022; Massimi, 2025), I argue that the Mediterranean Sea, conceived as an integrated object of study, intersected with military, territorial, diplomatic, and environmental concerns in ways that decisively shaped both collaboration and knowledge production.

A Plea for History and Philosophy of Statistics and Machine Learning

Hanti Lin
University of California, Davis
ika@ucdavis.edu

For statistics, the integration of its history and philosophy was initiated at least by Hacking (1975, 1990) and further developed by Mayo (1996) and Zabell (2005), yet it has not received sustained follow-up. This lack of integration is increasingly problematic. Recent advances in artificial intelligence have been driven largely by machine learning, a field that developed historically alongside statistics and now overlaps with it in both methods and aims. Today, the boundary between statistics and machine learning is increasingly blurred. What is urgently needed, therefore, is integration twice over: an integrated history and philosophy of two scientific fields that are themselves deeply integrated---statistics and machine learning.

In response, this paper investigates an epistemological idea that, I argue, has played a central role in both frequentist statistics and machine learning but has remained largely implicit and under-articulated. For want of a standard term, I call it achievabilism: the epistemological thesis that the frequentist standard for assessing even just the truth-finding performance of a non-deductive inference method should not be fixed in advance, but should instead be sensitive to what is achievable within a given problem context. On this view, epistemic evaluation is constrained not by an absolute ideal of inference, but by the highest standard in frequentist statistics achievable under the structural and informational limits of the inquiry at hand.

I argue that achievabilism can be traced at least back to Neyman and Pearson's 1936 work, an often overlooked follow-up to their far more frequently cited 1933 paper on frequentist hypothesis testing. Although rarely articulated, achievabilism informed the methodological innovations of frequentist statistics. Over time, the idea was often practiced without being explicitly recognized, occasionally forgotten, and later reinvented---most notably by machine learning theorists in the 1980s and 1990s, such as Devroye (1982) and Devroye, Györfi, and Lugosi (1996). By reconstructing these episodes and related ones, the paper argues that achievabilism is one of the core principles guiding the practice of frequentist statistics---and machine learning---but has been largely omitted from philosophical accounts of statistics, including those by Hacking (1975, 1990), Mayo (1996), and Zabell (2005).

The paper also develops two philosophical applications of achievabilism. First, achievabilism offers a response to Norton's critique of formal theories of induction by motivating a conception of formal epistemology that is local and context-sensitive rather than universal---an epistemology that, I argue, represents what many statisticians and machine learning theorists have been doing. Second, achievabilism provides a unifying framework for understanding the many local versions of Ockham's razor operative across scientific practice, drawing on statisticians' and machine learning theorists' context-sensitive justifications of different inferential razors.

Finally, beyond integrating the history and philosophy of statistics with machine learning, the paper also pursues a further methodological integration. It brings together two strands of philosophy of science often treated as opposites: history and philosophy of science, with its focus on historical case studies and scientific practice, and formal epistemology, which helps uncover epistemological ideas embedded in the highly mathematical theories of statistics and machine learning.

The Medicalization of the Cyborg: How Smart Wearable Devices Reshape Our Understanding of Health

Qiman Liu
Peking University
lqm@hsc.pku.edu.cn

In contemporary technological societies, the definition of health—once the core debate between naturalism and normativism—has undergone a profound transformation. Naturalist models traditionally define health in terms of biological functions, while normativist perspectives emphasize its embeddedness in social and cultural values. However, the rapid proliferation of smart wearable devices equipped with artificial intelligence (AI) has blurred this long-standing boundary. These devices, functioning as extensions of the body, continuously collect and quantify physiological data, transforming every aspect of bodily existence into measurable health indicators. Under the discourse of "health risk," AI systems interpret these data through medicalized algorithms that classify even minor deviations—such as sleep duration, eating habits, or activity levels—as potential disorders. In this process, human bodies, as newly emerged Digital Intelligent Cyborgs, are subjected to a comprehensive medicalization that extends beyond traditional medicine's domain.

This total medicalization is not merely a matter of monitoring but represents a redefinition of what counts as "health." As algorithms translate complex, lived bodily experiences into standardized metrics like BMI, BMR, or other risk indices, health becomes a computational construct shaped by the logic of data optimization. The medical gaze, once limited to clinical spaces, now operates through wearable sensors and cloud-based analytics, rendering every individual a site of continuous medical surveillance. The result is a hybrid model of health that fuses biological facts with socially constructed values—where "being healthy" is no longer a natural state but a technologically mediated one.

Yet, this transformation produces what may be called the "data health paradox." The more precisely health is measured, the further individuals drift from genuine well-being. Constant self-tracking and algorithmic comparison foster anxiety, dependency, and an illusion of control, while the pursuit of "normal" medical reference values can paradoxically generate new forms of pathology. Within this paradox, the cyborg body becomes both the subject and object of medical regulation—its autonomy eroded by technological determinism, its individuality subsumed under the universal logic of data.

Ultimately, the AI-driven wearable revolution does not merely enhance our capacity to monitor health; it redefines what health is. By medicalizing every aspect of life through the language of risk, it dissolves the boundaries between the biological and the social, the healthy and the pathological, the living and the machinic. In doing so, it compels us to confront a new philosophical and ethical question: in a world where data define the body, can health still be human?

Integrating omics in the practices of molecular biology small labs: experimental units and mechanistic explanations

Emmanuelle Maciel
independent scholar
emaciel02@gmail.com

The development of omics in molecular biology has triggered extensive discussions about their impact, both among biologists and philosophers. Notably, the role of databases and the scientific practices associated with their curation, and the packaging and travel of data involved were thoroughly described by Leonelli (Leonelli 2014). In addition, large-scale projects, such as the Human Genome Project or ENCODE, have been widely commented on both by biologists and philosophers of science.
Yet, although these types of biological research participate in shaping the way biologists understand their research object, they do not follow the typical organization of most research in molecular biology. Traditionally, indeed, research in molecular biology is organized in "small labs" (Cetina 1991). In small labs, molecular biologists are organized in groups around a laboratory leader. Knorr-Cetina wrote in 1999 that large-scale projects had not abolished the small lab culture in molecular biology. More recently, Ratti and Stoeger showed how small lab research is influenced by large-scale projects. For example these projects provide useful tools to help small lab scientists develop their hypotheses (Ratti and Stoeger 2021). However, although it has been discussed by biologists, it is often overlooked in philosophical discussions that small labs themselves also carry out their own omic experiments. By this, I refer to the fact that small research units perform omic experiments for the needs of their own specific projects. In this work, I examine how the advent of omics impact small lab biology. Based on my own experimental work in a molecular biology laboratory, I propose a framework to conceptualize omics without relying on large classical principles such as deduction, induction or abduction. Instead, I suggest characterizing what omics bring to the construction of a mechanistic explanation in molecular biology, starting from the way experimentation is structured both in traditional molecular biology and in omics. It appears that in molecular biology small labs, omics and traditional molecular biology experimentation are structured in the same way, and are thus similarly contributive to the construction of a mechanistic explanation. In particular, I propose that experimental work is divided into experimental units. These experimental units consist in the smallest research work that yields an observable result, which is directly involved in the justification of a part of the mechanistic explanation. More precisely, experimental units yield an entity or an activity of the mechanism, or any associated features -localization, interactions between entities, etc. Experimental units comprise four steps: the intervention in the biological system, sample
preparation, production of information on the component of interest, and revelation. Their articulation generally follows the mechanism schemata, but is more complex. In brief, molecular biologists have integrated omics into their own epistemic culture: omics belong to molecular biologists' experimental toolbox as any other type of experimentation. This does not mean that there are no epistemically relevant differences between omics and traditional molecular biology experimentation. Indeed, omics have particular features that have consequences on scientific practice and on the way molecular biologists conceive experimental design.

Extrapolating from Tissue Models

Patrick McGivern
University of Wollongong
patrickm@uow.edu.au

Biomedical research on tissue models such as organoids has exploded in the last decade. These models involve artificially constructed complexes of cells derived from human donors, and are able to mimic features of complex organs including the pancreas, liver, and brain. Such models raise various questions: how should artificially constructed tissue models be validated, and to what extent can observations of tissue models be extrapolated to their natural biological counterparts?

The problem of extrapolation is familiar from work on model organisms, where the primary concern is with the extent to which information obtained from the study of organisms of one species can support inferences concerning organisms of another species (Ankeny and Leonelli 2020; Kaplan 2017; Schaffner 2001; Steel 2008). A common theme in discussions for extrapolation involves the need to supplement superficial factors such as structural similarities with phylogenetic factors indicating the degree of evolutionary relatedness between model and target organism (Levy and Currie 2015).

The aim of this paper is to consider a set of problems concerning extrapolation from model systems in biology that arises when the appeal to phylogeny is no longer straightforward. This occurs in particular in cases involving synthetic systems such as organoids and other tissue constructs (Lancaster and Huch 2019), as well as in cases involving synthetic bioengineered systems (Davies and Levin 2023; Levin 2023). Such systems are often used as model systems in biomedical research and other areas, despite the fact they are significantly dissimilar from the intended target system, and despite the fact that their synthetic nature makes their phylogenetic status problematic since they are hybrids of engineered and evolved systems.

To address biological models in these novel context, I consider the possibility of understanding the role of such systems as biomedical models by drawing an analogy with so-called minimal models in physics (Batterman and Rice 2014; McGivern 2020). One of the key factors in the analogy is the idea that minimal models can capture many explanatorily significant features of a system while substantially misrepresenting other features. In the case of minimal models in physics, one way of justifying inferences across this 'misrepresentational gap' involves a rigorous process of demarcating classes of systems that are sufficiently similar in relevant respects despite significant variation in irrelevant respects. I argue that the appeal to phylogenetic relatedness plays a similar epistemic role in the case of model organisms, but one that can also be fulfilled in other ways in biological model systems. Describing this involves clarifying several distinct types of similarity that are often emphasised in biomedical contexts.

References

Ankeny, R.A. and Leonelli, S., 2020. Model organisms. Cambridge University Press.
Batterman, R.W. and Rice, C.C., 2014. Minimal model explanations. Philosophy of Science, 81(3), pp.349-376.
Davies, J. and Levin, M., 2023. Synthetic morphology with agential materials. Nature Reviews Bioengineering, 1(1), pp.46-59.

Kaplan, D.M., 2017. A bridge too far? Inference and extrapolation from model organisms in neuroscience. In The Routledge handbook of philosophy of animal minds (pp. 448-457). Routledge.
Lancaster, M.A. and Huch, M., 2019. Disease modelling in human organoids. Disease models & mechanisms, 12(7), p.dmm039347.
Levin, M., 2023. Darwin's agential materials: evolutionary implications of multiscale competency in developmental biology. Cellular and Molecular Life Sciences, 80(6), p.142.
Levy, A. and Currie, A., 2015. Model organisms are not (theoretical) models. The British Journal for the Philosophy of Science.
McGivern, P., 2020. Active materials: minimal models of cognition?. Adaptive Behavior, 28(6), pp.441-451.
Schaffner, K. F., 2001. Extrapolation from animal models. In McLaughlin, P., Machamer, P., and Grush, R. (eds.), Theory and method in the neurosciences , pp. 200-230. Pittsburgh: Pittsburgh University Press.
Steel, D., 2008. Across the boundaries: Extrapolation in biology and social science. Oxford: Oxford University Press.

From Surfacing to Shaping: Rethinking the Toolbox Dialogue Method in the Philosophy of Scientific Practice

Julie Mennes
Ghent University
julie.mennes@gmail.com

Chad Gonnerman
University of Southern Indiana
cgonnerm@icloud.com

Stephen Crowley
Boise State University
stephencrowley@boisestate.edu

Chet McLeskey
Michigan State University
mcleskey@msu.edu

Michael O'Rourke
Michigan State University
orourk51@msu.edu

Brian Robinson
Texas A&M-Kingsville
prof.brobinson@gmail.com

The Toolbox dialogue method (hereafter, "the method") is a widely used philosophical intervention designed to support crossdisciplinary collaboration in the sciences and beyond. It is typically described as serving a diagnostic or illuminative function: by responding to Toolbox prompts and engaging one another in dialogue in a workshop setting, collaborators

are depicted as uncovering their philosophical commitments—including their implicit assumptions, tacit values, and unspoken worldviews—in ways that enhance both their own self-understanding and their understanding of one another. But this way of describing the method overlooks an important aspect of its enactment.
A closer look at how a Toolbox workshop unfolds shows that the method does more than surface pre-existing philosophical commitments. At the heart of the Toolbox method is a survey-like instrument with features designed to foreground project-relevant issues, invite exploration of ambiguous or contestable terms, and encourage metacognitive reflection on team alignment. These features, together with the dynamics of the facilitated dialogue itself, invite participants to articulate their responses, explain their reasoning, and engage with alternative interpretations, elaborations, and challenges. As a result, Toolbox workshops do more than reveal and clarify what is already there: by creating conditions that encourage participants to develop supporting reasons and work out the implications of their views, they both enrich participants' philosophical commitments and stabilize them epistemically and socially, while also revising those commitments and sometimes generating new ones. In this way, the method becomes a vehicle for co-constructing commitments, demonstrating how organized, goal-directed activities can shape the orientations that guide crossdisciplinary work in the sciences and beyond.
This analysis contributes directly to the development of a philosophy of scientific practices that considers theory, practice, and the world together. It shows how reflective tools can be constructive components of scientific work, helping to coordinate philosophical commitments, collaborative practices, and real-world problems both by clarifying existing alignments and differences and by developing new ones. This expands the scope of what counts as epistemically significant in science by showing that activities aimed at facilitating collaboration—in addition to experiments, models, and data analysis—warrant philosophical attention, since they too help constitute how scientific knowledge is produced, justified, and made actionable. The paper develops this account through examples drawn from Toolbox workshops and considers its implications for how philosophers of science engage with scientific practices from within.

Participation and Immersive Imagination in Thought Experiments

Daniele Molinari
University of Parma
daniele.molinari@unipr.it

This paper explores the epistemic role of immersion in thought experiments ("TEs"), framing it as a form of imaginative participation that shapes both their interpretation and persuasive power. Building on Walton's theory of make-believe, I distinguish between physical and psychological participation, each involving commitment to a game's rules and self-projection into a fictional scenario. Unlike other forms of engagement in make-believe, participation involves taking on a role through an exercise of experiential imagination rather than detached analysis or disembodied observation of a fictional world. Although Walton and Waltonian fictionalist accounts maintain that TEs do not depend on participation, I argue that immersive psychological participation plays a central role in their epistemic function.
To clarify this role, I draw on aesthetic theories of performance. Adopting Hacking's theatrical metaphor, I present thought experimentation as akin to interpreting a part in a play. While Hacking employs this metaphor to argue that TEs cannot evolve, I adopt it to

emphasize the performative dimension of immersive imagination. From this perspective, TEs reveal two intertwined aspects: their dual nature as both scripts and performances, and their status as public acts rather than private experiences. As a result, immersive features of TEs are best understood as epistemic moves that put their scripts into play through a variety of potentially divergent interpretations, which are critically evaluated by their audience – the research community.

Two case studies – Thomson's Dying Violinist and Einstein's Chasing a Beam of Light – illustrate how immersive strategies can prompt ethical judgements or highlight theoretical limits. In Thomson's case, immersion invites readers to inhabit a dilemma in which their constrained agency becomes prominent in moral reasoning. Chasing a Beam of Light, on the other hand, encourages readers to imagine an impossible experience that forcefully conveys the urgency of a paradigm shift in physics.

Although immersive features can enhance the epistemic function of TEs, the very same mechanisms can also bias their reception and argumentative impact, as what sustains psychological participation may conflict with what is commonly regarded as more reliable intellectual processing – such as the careful analysis of the argumentative structure of TEs. Therefore, the epistemic dimension of immersion emerges as ambivalent: its compelling power can be either productive or misleading, prompting negotiations and resistance. By examining critics who seek to reduce or reconfigure the immersive appeal of Thomson's Dying Violinist, I argue that the epistemic value of immersive psychological participation lies not in the accuracy of the experiences it recreates, but in its capacity to establish a shared space of contestation – thus functioning as a dialectical strategy within scientific and philosophical inquiry.

The outcome of this analysis is that immersion should be regarded not as a mere embellishment of TEs, but as a method through which theories are collectively tested, contested, and transformed. Taking immersive TEs seriously requires that philosophy of science and metaphilosophy expand their epistemic toolkit, recognizing style, performance, and audience reception as key dimensions that can both enrich and distort our ways of exploring and understanding the world.

Spotting Planets: Appearance, Motion, and the Making of the Planetary in 19th Century Astronomy

Janna Müller
University of Cambridge
jkm64@cam.ac.uk

The early 21st century debate on Pluto and other small solar system objects focused on one central question: what exactly makes a planet? The decision, which demoted Pluto from its planetary status to that of a dwarf planet, highlighted tensions between observational complexity, astronomical taxonomy, and public imagination, and has become a standard example in discussions of conceptual engineering and scientific classification (see, for example, Bokulich 2014 and Jorem 2022). In this paper, I will focus on a period about 200 years earlier, asking how planethood was defined by astronomers in the early 19th century. Between 1801 and 1807, four new celestial objects were discovered, now known as asteroids or minor planets. While they moved on elliptical orbits like the other known planets, they only appeared as tiny, blurry spots of lights, challenging traditional understandings of the taxonomy of solar system objects and sparking an explicit debate among astronomers about

their classification. In the studies of planets, meanwhile, high-resolution telescopes allowed the detailed studies of their disc and appearances, especially surface details such as spots and belts. By analysing these studies and discussions and their practitioners, this paper argues that different astronomical practices produced distinct concepts of what it meant to be "planetary."

Two competing approaches to grouping and distinguishing celestial objects exemplify this claim. The first, which I call a motion-focused practices, emphasized positions and orbital characteristics. Within this framework, asteroids and classical planets were treated as members of a shared category because they followed similar elliptical paths around the Sun and were observed and calculated through the same observational and mathematical procedures. Motion-focused astronomers thus prioritized how objects moved as a classificatory characteristic, following the traditional Keplerian concept of a planet as an object moving on elliptical orbits.

In contrast, an appearance-focused practices reoriented attention toward the observations of surface features visible through telescopes and making inferences about the physical nature of those bodies. Focusing on the study of spots on planets, I will show how appearance-focused astronomers used these observations make assumptions about the nature of planetary atmospheres and its physical and chemical properties. In that regard, asteroids and planets as categories were made more distinct, with the "planetary" becoming akin to bodies as (livable) worlds not unlike Earth.

By tracing how these two kinds of practices shaped the objects of study, I argue that what it means to be planetary is not inherent in celestial objects but emerges from the practices that scientists enact: from the instruments they use, the questions they pose, and the criteria they prioritize in their communities of practitioners.

Making Epistemic Perspectives Explicit: The Case for Explicit Causal-Variable Assumptions in Psychedelic Trial Reporting

Christina Ni
Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University
cni.sh11@nycu.edu.tw

Karen Yan
Institute of Philosophy of Mind and Cognition, National Yang Ming Chiao Tung University
karenyan@nycu.edu.tw

Recent debates in psychedelic science concern whether the preparatory, monitoring, and integration procedures used in psilocybin trials should be classified as forms of psychotherapy (Goodwin et al., 2023; Villiger, 2025). At first glance, this dispute appears to be a substantive disagreement over the nature of psychotherapy. We argue, however, that the disagreement is better understood as a methodological conflict arising from different epistemic perspectives—implicit background assumptions that structure how researchers conceptualize interventions, identify causal variables, and interpret clinical trial results. By reconstructing these perspectives and examining their role in experimental practice, we show that the debate is less about classificatory boundaries and more about mismatched assumptions regarding what psychedelic trials are designed to investigate.

Our analysis draws on the framework introduced by Yan et al. (2025), who distinguish two epistemic perspectives in psychedelic research. The drug-centered perspective (DCP) treats

the psychedelic compound itself as the primary causal variable, with non-pharmacological procedures functioning as controlled background conditions. The drug-assisted perspective (DAP), by contrast, treats the relationship between drug-induced experience and therapeutic engagement as the causal variable. Our analysis shows that Goodwin's argument presupposes DCP. Under DCP, non-directive support is treated as a background control rather than as part of the active intervention; psilocybin's effects are therefore interpreted as pharmacological. Villiger operates within DAP. Under DAP, the causal variable under investigation is the interactional process linking the participant's drug-induced experience with therapeutic engagement, and this is why the same support procedures are classified as psychotherapeutic. The two sides thus talk past one another not because they disagree about the meaning of psychotherapy, but because they begin from different assumptions about what the trial is meant to test.

We then extend this analysis beyond the Goodwin–Villiger exchange by examining a broader range of psychedelic trials, including the classical psychedelic studies catalogued in Uyar et al. (2025). This reveals a more diverse landscape of epistemic perspectives than is typically acknowledged. In addition to DCP and DAP, some trials target the temporal profile of drug action (e.g., acute versus enduring effects), while others foreground neural-level mechanisms. These examples show that psychedelic science is not organized around a single unified experimental logic but instead around multiple, often implicit, models of what counts as the relevant object of investigation. This conceptual heterogeneity has concrete methodological consequences: trials cannot be straightforwardly compared, aggregated, or meta-analyzed when they are designed to investigate different kinds of causal targets.

To address this challenge, we propose a perspective-explicit reporting protocol for psychedelic clinical trials. The protocol requires researchers to (1) state the epistemic perspective guiding their trial design, (2) operationalize and manipulate the variables that this perspective identifies as causally relevant, and (3) interpret results accordingly. Making epistemic perspectives explicit clarifies what a given trial is intended to test, improves the consistency of causal inference, reduces interpretive conflicts across research groups, and enhances the transparency needed for regulatory, clinical, and ethical decision-making.

More broadly, psychedelic research provides a rich case for philosophy of science in practice. It illustrates how experimental design is shaped by implicit background assumptions; how methodological tensions arise when relational or temporally extended causal processes are investigated using standardized trial formats; and how conceptual clarification can guide better scientific practice. By analyzing psychedelic trials through the lens of epistemic perspectives, we offer a generalizable approach for understanding methodological pluralism and conceptual heterogeneity in contemporary clinical research.

Towards a pluralistic account of social norms' influence for behavioral policy

Tatu Nuotio
University of Helsinki
tatu.nuotio@helsinki.fi

Social norms are widely recognized for their ubiquitous influence in philosophy and social sciences. Yet, there is persistent disagreement over what they are and how they influence behavior (Legros & Cislaghi, 2020). For example, while most accounts agree that social sanctions are central in causing conformity, only some take them to be also necessary, and many exclude motivations based solely on internalized normative attitudes, which others are

happy to include. Facing this theoretical diversity, some authors have adopted a pluralistic view, arguing that norms come in many guises and influence behavior through various ways (e.g., Young, 2015). Others aim to rather fix the answers to both questions at once, defining social norms in terms of why they are followed (e.g. Bicchieri, 2017). I argue that the pluralistic view is preferable at least in one important context where theories of social norms are applied: social norm interventions.

Recently, there has been an increasing number of different interventions that aim to leverage or alter the influence of social norms in order to solve health-related, social, or environmental problems, (see Bicchieri, 2017). Despite their popularity, these social norm interventions (SNI) have often yielded heterogeneous results, evidence on long-term effectiveness is lacking, and their ethical aspects are rarely addressed. Philosophers have argued that addressing such issues of effectiveness and ethical acceptability requires knowledge about the causal mechanisms through which interventions affect behavior (see Grüne-Yanoff 2016). To understand these mechanisms, theoretical explanations of social norms' influence are a promising starting point, since supposedly, they describe the processes that these interventions intervene on. However, there are many explanations that are plausible and empirically compatible, and thus prima facie relevant for understanding SNIs, but which do not all fit under a single account of social norms. Working too strictly within the conceptual boundaries of any such account thus risks neglecting explanations that could prove relevant for understanding intervention mechanisms, which is why we should instead embrace the pluralism and treat different theories as being relevant for different context. The lack of conceptual rigidity is acceptable, because we are ultimately not interested in what happens to social norms per se in SNIs, but to the processes that cause and sustain norm-related behaviors. Whether different explanations truly capture these processes should be treated as an empirical rather than conceptual issue.

References:

Bicchieri, C. (2017). Norms in the Wild: How to Diagnose, Measure, and Change Social Norms. Oxford University Press.

Grüne-Yanoff, T. (2016). "Why behavioural policy needs mechanistic evidence". Economics and Philosophy, 32(3), 463–483.

Legros, S., & Cislaghi, B. (2020). "Mapping the social-norms literature: An overview of reviews". Perspectives on Psychological Science, 15(1), 62–80.

Young, H. P. (2015). "The Evolution of Social Norms". Annual Review of Economics, 7(1), 359–387.

An integrative perspective on models for multi-risk research

Malvina Ongaro
Politecnico di Milano
malvina.ongaro@polimi.it

Lorenza Petrini
Politecnico di Milano
lorenza.petrini@polimi.it

Multi-risk research aims to provide tools to describe and manage pluralities of interconnected natural risks. As it is a comparatively recent and strongly multi-disciplinary field, there exists a multitude of alternative approaches providing instruments to guide risk assessment and risk management (Ciurean et al. 2018, Hochrainer-Stigler et al. 2023). However, a systematic framework to address multi-risk is still missing. Indeed, the definition of common ontological and conceptual frameworks is one of the main challenges faced by practitioners in multi-disciplinary fields (O'Rourke, Crowley & Gonnerman 2016).

Philosophers of science therefore have a crucial role to play in the development of multi-risk research, as they can provide robust analyses of the core concept of the field and help in the definition of a shared ontology. Moreover, multi-risk research provides an interesting field for the study multi- and inter-disciplinary work, as it intrinsically involves multiple perspectives on a variety of interconnected risks. Indeed, even the contribution presented in this paper is born of an inter-disciplinary collaboration between philosophy of science and seismic engineering.

In this work, we develop a conceptual framework to understand multi-risk. To do that, we expand on existing definitions of (single) risk, putting the focus on the role played by the specification of spatial and temporal boundaries, the different causal interactions that are possible between co-existing hazards, and the ways in which both natural hazards and human action can change exposure and vulnerability to multiple hazards. This variety grounds the identification of a typology of multi-risk contexts.

From this framework, we derive an operational tree that suggests when a multi-risk approach may be necessary in general, and we identify multiple ways to model multi-risk contexts serving different and complementary purposes. In doing that, we define the conditions under which it is justified (and sometimes recommended) to adopt each of them, and we propose an integrative perspective that allows us to visualize consistently different types of models and their roles in understanding multi-risk.

We illustrate this approach with discussion of two types of models. In most cases, the ultimate objective of a multi-risk model is to identify the expected impacts and allow to calculate variations in these impacts as a consequence of potential interventions. This can be done at varying levels of completeness and resource demand (Liu et al. 2015). Comprehensive models aim at the construction of assessments providing precise forecasts of the expected impacts considering all relevant hazards (e.g. Marzocchi et al. 2012). However, these models are particularly demanding in terms of data, models, and calculations. On the other hand, scenario-based approaches (e.g. Zuccaro et al. 2018) propose to pick a specific chain of possible events, rather than modelling all possible interactions, and construct a timeline describing how expected impacts evolve after each event. Our framework suggests that neither of these modelling approaches is uniquely better than the other: rather, the relative adequacy of each of them will depend on the purposes of the model and on the resources available.

References

Ciurean, R.; Gill, J.; Reeves, H.J.; O'Grady, S.; Aldridge, T. (2018). Review of multi-hazards research and risk assessments. Nottingham, UK, British Geological Survey.

Hochrainer-Stigler, S., Trogrli*f*á, R. ≈†., Reiter, K., Ward, P. J., de Ruiter, M. C., Duncan, M. J., Torresan, S.,Ciurean, R., Mysiak, J., Stuparu, D., & Gottardo, S. (2023). Toward a framework for systemic multi-hazard and multi-risk assessment and management. Iscience, 26(5).

Liu, Z., Nadim, F., Garcia-Aristizabal, A., Mignan, A., Fleming, K., & Luna, B. Q. (2015). A three-level framework for multi-risk assessment. Georisk: Assessment and management of risk for engineered systems and geohazards, 9(2), 59-74.
Marzocchi, W., Garcia-Aristizabal, A., Gasparini, P., Mastellone, M. L., & Di Ruocco, A. (2012). Basic principles of multi-risk assessment: a case study in Italy. Natural hazards, 62, 551-573.
O'Rourke, M., Crowley, S., & Gonnerman, C. (2016). On the nature of Cross-disciplinary integration: a philosophical framework. Studies in history and philosophy of science part C: studies in history and philosophy of biological and biomedical sciences, 56, 62-70.
Zuccaro, G., De Gregorio, D., & Leone, M. F. (2018). Theoretical model for cascading effects analyses. International journal of disaster

---

Reflexivity and the Ethics of Assertion: Revisiting 'Values and Science'

Jakob Ortmann
Leibniz University of Hannover, Centre for Ethics and Law in the Life Sciences (CELLS)
jakob.ortmann@cells.uni-hannover.de

Kevin C. Elliott
Michigan State University
kce@msu.edu

Speech acts by scientists come with consequences. For instance, a claim about a drug's efficacy may turn out to be false, risking the health of patients who depend on those assessments. This simple yet pervasive fact about the pragmatics of science has arguably been underpinning the core of what is today known as the 'values and science' literature. Where it matters, the argument from inductive risk (Rudner, 1953; Douglas, 2000, 2009) is not about what hypothesis to reject, privately, or what proposition to believe, cognitively, but about what claim to assert, publicly, and with what foreseeable perlocutionary consequence (Franco, 2017).
Despite the breadth of the values and science literature, however, it has so far paid almost no attention to value judgments that pertain to a different kind of linkage between science and speech act theory, which we here refer to as 'reflexive performativity'. The concept is part of a distinction between 'generic performativity', which occurs when a scientific assertion has just any perlocutionary effect, and its subset 'reflexive performativity', which occurs when perlocutionary effects specifically bear on the referent of the assertion, such that the changes caused by the assertion feed back on the assertion's veracity (Ortmann, 2025). A stylised example of the latter is a 'self-fulfilling prophecy' (Merton, 1948), where, for instance, an economist might inadvertently bring about a bank run by publicly asserting "A bank run is imminent!". In such a case, the assertion does not merely function as a description or an epistemic service to society, but effectively as an intervention on the referent – the act of uttering was conducive to making it 'true' in the first place. Over the years, many related cases under various terms have been discussed in the social sciences (Hacking, 1996; MacKenzie & Millo, 2003; van Basshuysen et al., 2021; Godman & Marchionni, 2022; Khosrowi, 2023; Koskinen, 2022; Marchionni et al., 2024; Ortmann & Veit, 2023) but also in biology (Cooper, 2004; Khalidi, 2010; Laimann, 2020; Fagerberg, 2022).
In this paper, we argue that the distinction between generic and reflexive performativity marks a line between what could be considered the familiar territory of 'values and science'

and its understudied subdomain. Under non-reflexive generic performativity, arguments of error management are compelling precisely because one can assume that a stable target system exists about which true or false things can be said. Value-ladenness here pertains to decisions about what one takes to be true, given the available evidence. Under reflexive performativity, value-ladenness pertains to the question of what systems or practices should be sustained, changed, or brought into being by way of scientific practices, raising a distinct class of non-epistemic value questions that are not well-conceive of in terms of error management. As we argue, this puts pressure on core points of 'values in science': (i) the universal prohibition against wishful thinking, (ii) the distinction between (in)direct roles of values, (iii) the desiderata for managing values according to the New Demarcation Problem, and (iv) tuning value management to specific concerns.

References

Cooper, R. (2004). Why Hacking Is Wrong about Human Kinds. The British Journal for the Philosophy of Science, 55(1), 73–85.
Douglas, H. (2000). Inductive Risk and Values in Science. Philosophy of Science, 67(4), 559–579. https://doi.org/10.1086/392855
Douglas, H. (2009). Science, Policy, and the Value-Free Ideal. University of Pittsburgh Press.
Fagerberg, H. (2022). Reactive natural kinds and varieties of dependence. European Journal for Philosophy of Science, 12(4), 72. https://doi.org/10.1007/s13194-022-00500-x
Franco, P. L. (2017). Assertion, Nonepistemic Values, and Scientific Practice. Philosophy of Science, 84(1), 160–180. https://doi.org/10.1086/688939
Godman, M., & Marchionni, C. (2022). What should scientists do about (harmful) interactive effects? European Journal for Philosophy of Science, 12(4), 63. https://doi.org/10.1007/s13194-022-00493-7
Hacking, I. (1996). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds), Causal Cognition: A Multidisciplinary Debate. Oxford University Press.
Khalidi, M. A. (2010). Interactive Kinds. The British Journal for the Philosophy of Science, 61(2), 335–360. https://doi.org/10.1093/bjps/axp042
Khosrowi, D. (2023). Managing Performative Models. Philosophy of the Social Sciences, 53(5), 371–395. https://doi.org/10.1177/00483931231172455
Koskinen, I. (2022). Reactivity as a tool in emancipatory activist research. European Journal for Philosophy of Science, 12(4), 65. https://doi.org/10.1007/s13194-022-00487-5
Laimann, J. (2020). Capricious Kinds. The British Journal for the Philosophy of Science, 71(3), 1043–1068. https://doi.org/10.1093/bjps/axy024
MacKenzie, D., & Millo, Y. (2003). Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange. American Journal of Sociology, 109(1), 107–145. https://doi.org/10.1086/374404
Marchionni, C., Zahle, J., & Godman, M. (2024). Reactivity in the human sciences. European Journal for Philosophy of Science, 14(1), 8. https://doi.org/10.1007/s13194-024-00571-y
Merton, R. K. (1948). The Self-Fulfilling Prophecy. The Antioch Review, 8(2), 193. https://doi.org/10.2307/4609267
Ortmann, J. (2025). Performative Paternalism. European Journal for Philosophy of Science, 15(25). https://doi.org/10.1007/s13194-025-00651-7
Ortmann, J., & Veit, W. (2023). Theory Roulette: Choosing that Climate Change is not a Tragedy of the Commons. Environmental Values, 32(1), 65–89. https://doi.org/10.3197/096327122x16452897197784
Rudner, R. (1953). The Scientist Qua Scientist Makes Value Judgments. Philosophy of Science, 20(1), 1–6.

van Basshuysen, P., White, L., Khosrowi, D., & Frisch, M. (2021). Three Ways in Which Pandemic Models May Perform a Pandemic. Erasmus Journal for Philosophy and Economics, 14(1), 110–127. https://doi.org/10.23941/ejpe.v14i1.582

Non-Causal Unificationist Explanation: A Philosophical Methodology for the Epistemology of Artificial Intelligence

Enrong Pan
Sun Yat-sen University
panenrong@mail.sysu.edu.cn

Ziming Wang
Sun Yat-sen University
wangzm68@mail2.sysu.edu.cn

The pervasive integration of artificial intelligence (AI) into scientific practice and knowledge generation has motivated philosophers to seek a solid epistemological foundation for AI as a new, legitimate instrument. However, Paul Humphreys's epistemic opacity has become an obstacle to this new trend.

This paper discusses how epistemic opacity threatens the epistemological status of AI in knowledge generation, ethical contexts, and automated science. It then reviews two attempts to overcome this obstacle, i.e., seeking justification and explanation. The former approach bypasses model opacity and, from an externalist perspective, specifies the conditions under which AI-generated outputs qualify as knowledge, drawing on reliabilism, computational reliabilism, and functionalism. The latter aims to extract scientific explanations from AI models to foster scientists' understanding of model behavior and real-world systems, namely by constructing path dependencies, mechanistic explanations, and causal abstractions using explainable AI (XAI) techniques.

After reviewing these studies, the paper identifies three key pitfalls in existing approaches: the untimely normative approaches, the neglect of internal machine knowledge, and the conflict between "non-anthropocentric epistemology" and scientific explanation. The untimely normative approaches presuppose a "consensus view" on scientific knowledge, failing to accommodate the dynamic and evolving nature of scientific practice. The second pitfall highlights that current research pays insufficient attention to model structure. Externalist approaches, while enabling quantification, promote an "avoiding strategy" by overemphasizing easily measurable external indicators and systematically overlooking difficult-to-quantify structural indicators. This tendency further leads externalists to ignore the possibility of seeking scientific knowledge within the model. Finally, the third one shows that explanations grounded in human psychology and common sense, though useful for prediction and satisfying curiosity, have limited value for automated science.

XAI appears to be a potential approach for exploring internal machine knowledge, yet it still faces significant challenges. First, XAI relies on causal explanation of AI mechanisms, but extracting causal abstractions or constructing structural equation models is extremely difficult, and there remain ongoing philosophical debates concerning causality. Second, the value and role of causal explanations in automated science remain to be further discussed. Finally, existing XAI techniques either use surrogate models or relevant background knowledge as explanans. As a result, they neglect the possibility that the model itself can

serve as an explanans and overlook the objective explanatory relations between the model and the input/output data, as well as between the model and the training/test sets.
To address these challenges, this paper proposes a philosophical methodology and, in the final section, presents a case of machine knowledge with non-causal unificationist explanatory power. Drawing on Humeanism, we examine AlphaFold and investigate its internal "implicit mechanism of protein folding," contrasting it with the "explicit mechanism" produced by mathematical modeling. We argue that an implicit mechanism explains a fact insofar as it integrates that fact into an existing unified Humean mosaic and successfully predicts it. This paper shows that this approach both fills the gap in seeking scientific knowledge within AI models and provides an explanatory framework compatible with automated science, while also highlighting the importance of exploring machine knowledge for scientific practice.

Iterating Concepts in Endocrinology

Piet Fritz Pankratz
University of Vienna
piet.pankratz@univie.ac.at

Philosophers of science increasingly emphasize that concepts are not fixed representational containers but evolving tools that structure inquiry. Yet we still lack detailed accounts of how scientific concepts develop through practice in historically situated, methodologically heterogeneous fields—particularly in the life sciences. This paper offers such an account through a case study of the emergence of the concept of "sex hormones" in endocrinology (1905–1940), analyzed through Hasok Chang's framework of epistemic iteration. I argue that conceptual epistemic iteration (CEI) provides a powerful model for understanding both the formation and the subsequent transformation of scientific concepts, especially when these concepts are simultaneously empirically responsive and normatively laden.
I examine six stages in the development of early endocrinology: from Starling's introduction of "chemical messengers," through the rise of sex antagonism theories, to the biochemical destabilization of sex-specific taxonomies, proliferating classificatory alternatives, and eventual institutional standardization into "estrogens" and "androgens". Drawing on historical work by Nelly Oudshoorn and Anne Fausto-Sterling, I show that these transitions were not episodes of straightforward discovery but iterative refinements shaped by experimental practice, disciplinary negotiation, and entrenched cultural assumptions about sex and gender.
CEI illuminates this process in three key ways. First, it captures how inquiry begins with vague yet operationally fruitful concepts; Starling's early notion of hormones exemplifies the kind of imperfect starting point that CEI treats as epistemically indispensable. Second, CEI explains how empirical anomalies (such as the discovery that so-called "female" hormones circulate in men) generate conceptual stretching, pluralistic taxonomic experimentation, and eventual partial convergence. Third, CEI foregrounds how conceptual change is constrained by what Chang calls "engagement with nature" (Chang 2022, 242) while remaining permeable to prescientific assumptions. In the case of sex endocrinology, binary gender norms were not merely external biases but iterated components of the conceptual scaffolding guiding experimentation, interpretation, and standardization.
The resulting picture challenges two common narratives in the philosophy of science. Against discontinuous or revolution-based models of conceptual change, the case illustrates how

scientific progress can be simultaneously cumulative, messy, and normatively entangled. Against essentialist accounts of natural kinds, the history of "sex hormones" demonstrates that classificatory stability emerged not from intrinsic essences but from the iterative success of certain conceptual frameworks in coordinating practice.
I conclude that CEI not only succeeds in reconstructing the conceptual history of "sex hormones," but also offers a broader lesson for philosophers of science. By attending to small, incremental, practice-embedded iterations, CEI reveals how scientific concepts become empirically robust and socially authoritative at once. This case thereby enriches our understanding of conceptual development in the life sciences and highlights how modes of inquiry shape, and are shaped by, the conceptual tools through which scientists investigate the world.

Making evidence of mechanisms more reliable

Saúl Pérez-González
University of Valencia
saul.perez@uv.es

Over the last years, evidence of mechanisms has achieved great importance both in philosophical debate and in scientific practice. This kind of evidence is considered crucial for establishing and extrapolating causal relationships. Moreover, the growing relevance of evidence of mechanisms has extended beyond the biomedical sciences to the social sciences. Nonetheless, important concerns have been raised regarding the variable quality of evidence of mechanisms and the reliability of mechanism-based reasoning. Evidence of mechanisms is not always accurate and does not always adequately account for the operating causal mechanisms. Its assessment, consequently, is crucial to overcome those problems and use the evidence reliably. The aim of this paper is to address the evaluation of evidence of mechanisms, with a particular focus on the social sciences.
The evaluation of evidence of mechanism, despite its relevance, has received limited attention in the methodological and philosophical literature. Moreover, previous contributions on this topic, although valuable, suffer from significant limitations. Howick, Glasziou, and Aronson (2010, 2013) identify problematic aspects of evidence of mechanisms and highlight the relevance of careful assessment. Nevertheless, they do not provide a concrete and operational evaluation framework to guide scientific practice. Advocates of evidential pluralism (Parkkinen et al., 2018), on the other hand, offer a comprehensive evaluative framework and provide several quality indicators. However, their proposal encounters difficulties relating to case-specific adaptation, context dependence, and generality.
Building on previous proposals, we develop the Support–Absence Framework (hereafter, the SA framework). The SA framework sets out a range of criteria for assessing both the quality of evidence of mechanisms and the degree of support it provides for a given mechanistic hypothesis. Two main types of criteria are distinguished: support criteria and absence-of-error criteria. Support criteria assess whether the evidence concerns real mechanisms and whether these mechanisms are relevant to the hypothesis under examination. They aim to distinguish genuine mechanism-based reasoning from mere storytelling. Absence-of-error criteria, in turn, evaluate whether the evidence is undermined by problems typically associated with mechanistic studies. They aim to distinguish reliable from unreliable mechanism-based reasoning. Evidence of mechanisms that scores highly across several

criteria should be regarded as higher quality—and as providing stronger support—than evidence that performs well on fewer.
To explore and assess the application of the SA framework in the social sciences, we consider both process tracing and agent-based modelling. We argue that the framework can be successfully applied to evaluate process tracing studies as well as agent-based models. To support this, we discuss real and representative mechanistic studies. In the case of process tracing, we consider and assess Winward's (2021) study on mass categorical violence. Regarding agent-based modelling, we discuss the well-known Bearman et al.'s (2004) study on adolescent romantic and sexual networks.

References

Bearman, P. S., Moody, J., & Stovel, K. (2004). Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. American Journal of Sociology, 110(1), 44–91.
Howick, J., Glasziou, P., & Aronson, J. K. (2010). Evidence-based mechanistic reasoning. Journal of the Royal Society of Medicine, 103(11), 433–441.
Howick, J., Glasziou, P., & Aronson, J. K. (2013). Can understanding mechanisms solve the problem of extrapolating from study to target populations (the problem of 'external validity')? Journal of the Royal Society of Medicine, 106(3), 81–86.
Parkkinen, V.-P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M. P., Norell, C., Russo, F., Shaw, B., & Williamson, J. (2018). Evaluating evidence of mechanisms in medicine: principles and procedures. Springer.
Winward, M. (2021). Intelligence Capacity and Mass Violence: Evidence From Indonesia. Comparative Political Studies, 54(3–4), 553–584.

Between Thickness and Comparability: Value Pluralism and Measurement Drivers in Global Well-Being Metrics

Charlotte Constanze Poller
Wuppertal University
poller@uni-wuppertal.de

Global measurements of well-being shape how individuals, institutions, and governments understand and pursue the "good life". The World Happiness Report (WHR) is one of the most influential: it publishes annual cross-national rankings based on survey data and links "happiness" to social and economic indicators, positioning itself as an empirical basis for public-policy debate (Helliwell et al. 2025). The WHR thus provides a paradigmatic case for examining how complex, value-laden concepts are rendered measurable globally. To analyze such measurement practices, Basso & Alexandrova (2025) conceptualize measurement "drivers" as competing epistemic, ethical, pragmatic, and metrological considerations that guide design choices and inevitably force compromises. I argue that in the WHR pragmatic and metrological drivers, especially global comparability and communicability, systematically dominate epistemic and ethical drivers, yielding a deliberately "thin" operationalization of well-being.
The WHR's core measurement of well-being relies on a single life-evaluation item asking respondents to place their current life on a 0–10 ladder. Although this measurement presents itself as neutral and universally applicable, its normative underdetermination leaves unspecified how people arrive at their evaluations and which aspects of life they take to

matter for their well-being. In Basso & Alexandrova's (2025) terms, the WHR embodies a compromise that prioritizes operability, cross-national comparability, and communicability over making the normative content of well-being explicit and contestable. This design choice is not without problems. As Thoma (2024) argues, policy-relevant measures must, in principle, allow for the inclusion of citizens' value commitments; otherwise, they lack normative legitimacy. This requirement becomes especially demanding in the case of well-being. As Alexandrova (2017) emphasizes, well-being is a thick, context-dependent concept: different societies endorse different considerations as constitutive of a good life. Since such considerations vary across contexts, global well-being measurement confronts persistent disagreement over which aspects are constitutive of well-being. This generates a conceptual dilemma: adequate measurement requires some degree of thickness in order to avoid arbitrariness and to legitimize policy guidance, yet substantial value-disagreement undermines the abstraction and standardization on which global comparison depends. How thick, then, can global well-being measures be while remaining their comparability? I argue that the WHR resolves this tension by prioritizing pragmatic and metrological drivers over epistemic and ethical ones, effectively standardizing away substantively relevant differences in what well-being is taken to be. However, this solution comes at the cost of normative adequacy. Drawing on Thoma's (2024) argument that policy-relevant measures must allow for the inclusion of plural value commitments, I therefore defend a minimum-thickness requirement for global well-being measurement: at least some explicit, contestable constituents of well-being should be included.

In light of this dilemma, I outline three complementary strategies for implementing minimum thickness by rebalancing competing measurement drivers at different levels of the measurement enterprise. (1) Instrument level: enriching the WHR ladder with a small set of explicitly normative dimensions shifts priority toward epistemic and ethical drivers by increasing substantive guidance and transparency, but does so at the expense of metrological drivers such as universality and cross-national comparability. (2) Governance level: a functional separation of roles reassigns drivers across levels, reserving pragmatic and metrological drivers for global coordination metrics while relocating epistemic and ethical drivers to context-sensitive local measures. (3) Reporting level: replacing the master-number logic with pluralistic dashboards reweights pragmatic and rhetorical drivers in favor of epistemic plurality and ethical contestability, increasing deliberative visibility while reducing communicative simplicity.

References:

Alexandrova, A. (2017). A philosophy for the science of well-being. Oxford University Press.

Basso, A., & Alexandrova, A. (2025). Measurement requires compromises: the case of economic inequality. Studies in History and Philosophy of Science, 113, 88-97.

Helliwell, J. F., Layard, R., Sachs, J. D., De Neve, J. E., Aknin, L. B., & Wang, S. (2025). World happiness report 2025.

Thoma, J. (2024). Social science, policy and democracy. Philosophy & Public Affairs, 52(1), 5-41.

Examining epistemic and ethical practices in voice data collection among voice biomarker industry researchers

Anaïs Rameau
Weill Cornell Medical College
anr2783@med.cornell.edu

The human voice contains features linked to various diseases such as Parkinson's, Alzheimer's, depression, COPD, and heart failure. Advancements in voice technology, including audio signal analysis and natural language processing, have led to growing commercial interest in leveraging voice biomarkers for noninvasive disease screening, monitoring and diagnosis. Voice data collection is relatively inexpensive and easily carried out via ubiquitous smartphone technology, making the growing field of voice biomarkers scalable and enticing for innovators.

However, creating voice datasets in an epistemically sound and ethical manner is difficult. It has historically been limited by heterogeneous collection methods, complexities related to linguistic and cultural diversity among source populations, and non-standardized data storage and processing practices. These persistent inconsistencies can undermine the therapeutic relevance of resulting voice biomarker products and may have negative downstream consequences on clinical and therapeutic decisions and patient care. This study aims to provide an epistemic and ethical perspective on current practices of voice biomarker researchers in private industry to highlight current methodological and ethical limitations in voice biomarkers scalability.

Our analysis is based on a qualitative interview study using an interpretive framework rooted in abductive reasoning. We conducted semi-structured, one-on-one virtual interviews with researchers in voice biomarker startups worldwide. The interview questions focused on the identification of target patient populations, guiding ethical principles, consent procedures, data collection methods, financial incentives, and demographic and geographic information on participant populations. The qualitative data obtained from these interviews were coded and analyzed using MaxQDA software to identify common themes.

We interviewed researchers from eleven startups, with each interview lasting 30-60 minutes. Qualitative analysis revealed several key epistemic and ethical themes: 1) variability in audio capture methods with uncertain impact on scientific validity of data, 2) ambiguity surrounding privacy protection measures in voice data collection, 3) unresolved ethical questions around consent for cognitively impaired patients or those who do not fully understand how their voice data may be used, 4) recognition of the importance of partnering with patients and their advocates in aligning startup research goals and ethical data use and 5) tendency to source voice data from low- and middle-income countries for voice biomarker applications primarily targeting high-income markets despite concerns about generalizability of findings across populations and risks of epistemic exploitation.

Our findings highlight considerable heterogeneity in epistemic and ethical practices governing voice data collection among voice biomarker startups, including in variability in consent procedures, transparency about data usage, privacy protection and population sampling. This suggests a lack of industry-wide standard and lack of regulatory guidance. The pattern of sourcing voice data from low-and middle-income countries for commercial product development may be driven by regulatory asymmetries in privacy protection, cost-efficiency considerations and weaker consent standards. This raises obvious ethical concerns with epistemic repercussions, such as limited generalizability of findings across linguistically and culturally different populations. This study illustrates the ethical and epistemic importance of transparency in voice data collection, and how commercial pressures,

regulatory gaps, and global inequities jointly undermine the emerging field of voice biomarkers.

Testing the Disciplinary Assumption: Philosophical Commitments Across Scientific Disciplines

Brian Robinson
Texas A&M University-Kingsville
brian.robinson@tamuk.edu

Chad Gonnerman
University of Southern Indiana
cgonnerman@usi.edu

Troy Hall
Oregon State University
Troy.Hall@oregonstate.edu

Aaron McCright
Michigan State University
mccright@msu.edu

Gina M. Pizzo
Rice University
gp70@rice.edu

Katie Plaisance
University of Waterloo)
kplaisan@uwaterloo.ca

Cross-disciplinary research (CDR) has become increasingly common for addressing complex socio-scientific problems, such as climate change, global pandemics, and ethical challenges with the rapid use of AI An important class of challenges relates to the ways different scientists conceptualize the world and its scientific study (Palghat et al., 2017; Phillipson et al., 2009). These conceptualizations frame how scientists understand and pursue their research practice, reflecting epistemological, ontological, and value dimensions of scientific engagement with the world. We refer to these conceptualizations as philosophical commitments. The prevailing assumption—what we term the disciplinary assumption—is that philosophical commitments are relatively uniform within disciplines but differ substantially between them (Eigenbrode et al., 2007; Miller et al., 2008; Osbeck & Nersessian, 2017). Previously, no study has examined such philosophical commitments with a large representative sample of scientists across disciplines. This paper presents results from such a study. We surveyed 4,113 research-active scientists at Carnegie R1 institutions across the United States, drawn from a sampling frame of 23,642 tenure-system science faculty across 20 disciplines representing the life sciences, social and behavioral sciences, and physical and earth sciences.
Our study addresses a critical measurement challenge that has limited prior research, specifically assessing disciplinary affiliation. To measure disciplinary membership, we used a

multidimensional set of indicators: earned degree, departmental affiliation, professional association membership, publication venues, peer perceptions, and self-identity. This approach captures the complexity of scientists' disciplinary positioning beyond simple self-categorization.
The survey instrument assessed a wide range of philosophical commitments. We assessed five dimensions of ontological commitments (theory, causation, laws, realism, and truth), six of epistemological commitments (explanation/prediction, hypothesis generation/testing, confirmation, replication, confidence, and generalizability/transferability), and six of axiological commitments (acceptability of values guiding the research process, acceptability of advocacy, objectivity, defense of the value-free ideal, the relationship between diversity and objectivity, and the social costs of error).
In this presentation, we will report findings on the extent to which distinct clusters of philosophical commitments correspond with distinct scientific disciplines. Our analysis employs cluster analysis and structural equation modeling with confirmatory factor analysis to identify patterns of philosophical commitments within and across the 20 specific disciplines represented in our sample.
This work should be of interest to those in the SPSP community, both in terms of the methodological approach and content. Our work fits within—and helps to advance—the empirical turn in philosophy of science (Verbeek 2021). Moreover, while previous studies have examined scientists' philosophical commitments in one domain, such as metaphysics (e.g., Beebe and Dellsén 2020; Henne et al. 2024) or values (e.g., Steel et al. 2018), we investigated several types of philosophical commitments, allowing us to conduct richer analyses about the relationships among them. Our analyses also look across a wide range of disciplines, as opposed to one or a few disciplines (as is more common, including ins some of the sources already mentioned).

The Ethics of Measurement: Measurement Based Care in Youth Mental Health

Sebastian Rodriguez Duque
University of Cambridge
sr2203@cam.ac.uk

While there is a strong tradition exploring the consequences of measurement in education there has been little work on the ethical aspects of measurement in health care, and mental health specifically (Rodriguez Duque et al., 2024; Tal & Truijens, 2025). Meanwhile, the promise of measurement in mental health has fueled enormous funding in countries such as Canada to gather clinical data and build 'learning health systems' based on measurement-based care (MBC), often especially to serve young people. That is because 12-24 is the age range where the first signs of future mental health trouble will emerge (McGorry et al., 2024). MBC refers to the establishment of a baseline and the use of measurement questionnaires to track changes over time (Sarvet, 2020; Van Cleave et al., 2012). In mental health MBC specifically refers to the use of patient-reported outcome measures (Kroenke & Unutzer, 2017). These psychometric instruments are a rich ground for exploring the role of ethics in measurement for at least two reasons. One, because of the values that are implicated in the design and use of measurement questionnaires (e.g. in the use of DSM based constructs). Two, because of the value-laden consequences of using measurement questionnaires. In the philosophy of measurement, there has been a recent interest in the role of non-epistemic values in measurement (Alexandrova & Fabian, 2021; Basso & Alexandrova, 2025; Larroulet Philippi, 2021; Ohnesorge, 2024; Tal & Truijens, 2025).

PROMs have themselves been the focus of recent philosophical attention because they purport to represent what matters to patients from the patient's own point of view (McClimans, 2024). Yet the wholesale adoption of these instruments remains under-explored even if they are instruments of dubitable validity (Alexandrova & Haybron, 2016; Borsboom, 2006) built around DSM categories of dubitable validity (Andreasen, 2006). To my knowledge only two studies have explored what-it-is-like for patients to be measured in mental health (Truijens et al., 2019, 2022). In this paper, I add to this literature in two ways. First, I report my preliminary findings from qualitative interviews undertaken in partnership with Foundry – an Integrated Youth Service in BC, Canada in the spring of 2026. I will explore the experiences of youth using measurement, as well as their own interpretation of standardized measures which are now being put to use in their context. Second, I argue that the ethical dimensions of measurement encompass both the consequences of measurement at a population level, and the direct impact that the act of being measured has on a person. I explore the ethical and philosophical implications of the evidence uncovered in the qualitative interviews and point to one key difficulty for the ethics of measurement: the interaction between the values of a measurement instrument and a measurement practice more broadly (i.e. MBC), and the values of a patient (e.g. a young person) and their community.

The Method of Methods: C. S. Peirce on Statistics and Scientific Practice

Niall Roe
University of Cambridge
nrr32@cam.ac.uk

Peirce's philosophy of statistics is based on his experimental practice and designed to be practically useful for scientists. In this paper, I provide a new frame for understanding and motivating Peirce's philosophy of statistics – one based on his experimental work and situated in his larger philosophy of science – and use it to respond to two suggestions that Peirce harboured Bayesian sympathies.

Peirce understood "logic" broadly, not merely as formal logic, but as the "art of devising methods of research,—the method of methods" (W4, 378). The "logic of science" was the study of scientific practice, and Peirce gave statistics pride of place therein, referring to it as "the logic of science propter excellentiam" (R244, 66). Statistics earns its privileged place because it can diagnose and improve the logic of science itself. Such that "it is not merely an instrument of the reasoning, but is furthermore an organon of the logic of modern science", an organ which "might almost be likened to its heart" (R 244, 66-67; R763, 1).

Peirce developed his conception of statistics to be fit for scientific purpose. This was a pressing task. Scientific advancements precede the lessons we learn from them, as "practice must, to some extent, precede theory" (R692, 9). Nineteenth century statistics was dominated by one such novel method – the method of inverse probabilities, or Bayesian inference. The method was debated, but endured for want of an alternative. It was "accepted with a sigh" but persisted as "there [had] not yet been philosophy enough to expel it" (Zabell (1991, 214, 3)). The formalized alternative emerged with R.A. Fisher (Fisher 1922). But it has been recognized that the philosophical alternative was already well underway in Peirce's work (Hacking 1980; Mayo 1993). Peirce's answer is frequentist, and justified by experimental practice.

Against recent interpretations that read Peirce as harbouring Bayesian sympathies, I show that his frequentism runs deeper than such accounts allow. Burch (2010) portrays Peirce as

endorsing iterative subjective updating, while Atkins (2023), following Kasser (2016b; 2016a), reads Peirce's "Method of Balancing Reasons" (MBR) as a model of Bayesian belief revision. In response to Burch, I show how the iterative Bayesian approach cannot get off the ground on Peirce's picture: if probabilities are construed psychologically, they are of no interest to science; if they are construed via Peircean frequentism, then the required notion of a prior probability is incoherent. In response to Atkins, I show that the MBR is not a scale for weighing subjective credence, but a logical procedure for measuring the relative significance of arguments whose frequentist probabilities are already empirically, objectively established.

By situating Peirce's views on statistics within his broader philosophy of science, the paper highlights his distinctive effort to take probability out of the mathematical clouds and place it on grounds that could support experimental practice. I try to do the same, drawing from Peirce's own experimental data and examples as much as possible. In doing so, I aim to reaffirm the importance of statistics within Peirce's philosophy of science.

Works Cited:

Atkins, Richard Kenneth. 2023. Peirce on Inference: Validity, Strength, and the Community of Inquirers. Oxford Scholarship Online. Oxford University Press. https://doi.org/10.1093/oso/9780197689066.001.0001.

Burch, Robert. 2010. "If Universes Were as Plenty as Blackberries: Peirce on Induction and Verisimilitude." Transactions of the Charles S. Peirce Society 46 (3): 423. https://doi.org/10.2979/tra.2010.46.3.423.

Fisher, R. A. 1922. "On the Mathematical Foundations of Theoretical Statistics." Philos. Trans. Roy. Soc. London Ser. A 222 309-368.

Hacking, Ian. 1980. "The Theory of Probable Inference: Neyman, Peirce and Braithwaite." In Science, Belief and Behaviour: Essays in Honour of R. B. Braithwaite, edited by David H. Mellor and Richard B. Braithwaite. Cambridge University Press.

Kasser, Jeff. 2016a. "Confidence, Evidential Weight, and the Theory-Practice Divide in Peirce." Transactions of the Charles S. Peirce Society 52 (2): 285. https://doi.org/10.2979/trancharpeirsoc.52.2.09.

Kasser, Jeff. 2016b. "Two Conceptions of Weight of Evidence in Peirce's Illustrations of the Logic of Science." Erkenntnis 81 (3): 629–48. https://doi.org/10.1007/s10670-015-9759-5.

Mayo, Deborah G. 1993. "The Test of Experiment: C. S. Peirce and E. S. Pearson." In Charles S. Peirce and the Philosophy of Science: Papers from the Harvard Sesquicentennial Congress, edited by Edward Carter Moore. University of Alabama press.

Peirce, Charles S. 1966. "The Charles S. Peirce Papers, Microfilm Edition, Thirty Reels with Two Supplementary Reels Later Added. (Cited as 'R [Manuscript #], [Page#]')." Photographic Service. With Richard S Robin. Harvard Library.

Peirce, Charles S. 1982. Writings of Charles S. Peirce, A Chronological Edition. (Cited as "W[Volume #], [Page #]"). Edited by Peirce Edition Project. 8 vols. Indiana University Press.

Stigler, Stephen. 2005. "Fisher in 1921." Statistical Science 20 (1). https://doi.org/10.1214/088342305000000025.

Zabell, S. L. 1991. "Ramsey, Truth, and Probability." Theoria 57 (3): 211–38. https://doi.org/10.1111/j.1755-2567.1991.tb00838.x.

Zabell, Sandy. 1989. "R. A. Fisher on the History of Inverse Probability." Statistical Science 4 (3): 247–56.

Socially Responsible Science and Engineering in Practice: Promise and Perils of Institutional Policies

Kristina Rolin
Philosophy, Tampere University
kristina.rolin@tuni.fi

Erik Inkinen
Software Engineering, Tampere University
erik.inkinen@tuni.fi

In recent years, calls for socially responsible science and engineering (SRSE) have emerged alongside traditional epistemological and ethical issues in philosophy of science (de Melo-Martín & Intemann 2023). Philosophers argue that to be socially responsible, scientists must follow appropriate moral and social values in different stages of scientific inquiry and respond to citizens' knowledge needs (Alexandrova 2018; Brown 2020; Kourany 2010). When scientists make moral and social value judgments, they should respect democratic principles by aligning their judgments with democratically endorsed values (Intemann 2015; Kitcher 2001; Schroeder 2021), or the values of those social groups who are impacted by their expert judgment (Lusk 2021; Rolin 2024). Social responsibility is seen as especially important when scientists function as experts who offer research-based advice to various publics. Members of the public and policymakers do not want to rely on scientific research and technological solutions shaped by moral and social values they have good reasons to reject (Rolin 2021).

Philosophers' calls for SRSE are often followed by recommendations for institutional reforms. In our presentation, we discuss the pros and cons of institutional efforts to implement SRSE in practice. While institutions can do many things to promote SRSE, institutional policies are not without costs and risks which need to be weighed against advantages when such measures are proposed. Philosophers argue that for scientists to be able to conduct their research and communicate their results in socially responsible ways, they need support from scientific institutions (Douglas 2021; Resnik & Elliott 2015). Institutional support can mean many things: opportunities for education and professional development, access to advisors, resources for engaging stakeholders, protecting scientific experts from harassment and intimidation, and rewarding excellence in SRSE. Institutional policies can also involve mechanisms for compliance, such as requiring scientists to prepare a plan for broader impact work or integrating issues of social responsibility into mandatory review by institutional review boards (de Melo-Martín & Intemann 2023; Kourany 2010). We argue that despite their costs, support policies are promising avenues for implementing socially responsible science in practice, whereas mechanisms for compliance involve risks. We analyze two risks, one, the risk that mechanisms for compliance are used to curtail academic freedom, and two, the risk that they lead to empty proceduralism which reduces social responsibility into box-ticking exercise. The threat to academic freedom arises when mechanisms for compliance become tools for anti-pluralist politics in science (Kinzelbach et al. 2025). Anti-pluralist politics threatens to undermine SRSE because it dictates responsibilities in undemocratic ways. In our view, Douglas's (2024) concern with the improper politization of science does not fully capture what is at stake in threats to academic freedom. The risk of empty proceduralism arises when engagement with the beneficiaries of research becomes an administrative requirement rather than a relationship guided by moral and social values. Empty proceduralism threatens to undermine SRSE because social

responsibility is seen primarily as an obligation to institutions rather than as a responsibility to citizens and communities.

References:

Alexandrova, A. (2018). Can the science of well-being be objective? The British Journal for the Philosophy of Science, 69(2), 421–445.
Brown, M. J. (2020). Science and Moral Imagination: A new ideal for values in science. University of Pittsburgh Press.
de Melo-Martín, I., and Intemann, K. (2023). Socially responsible science: Exploring the complexities. European Journal for Philosophy of Science 13: 33.
Douglas, H. (2021). Scientific freedom and social responsibility (pp. 68-87). In Science, freedom, democracy, edited by P. Hartl & A. T. Tuboly. Routledge.
Douglas, H. (2024). Structuring institutions for responsible and accountable research. Philosophy of Science 91(5): 1149-1158.
Intemann, K. (2015). Distinguishing between legitimate and illegitimate values in climate modeling. European Journal for Philosophy of Science, 5, 217–232.
Kinzelbach, K., Lindberg, S. I., Lott, L., & Panaro, A. V. (2025). Academic Freedom Index 2025 Update. FAU Erlangen-Nürnberg and V-Dem Institute.
Kitcher, P. (2001). Science, truth, and democracy. Oxford University Press.
Kourany, J. (2010). Philosophy of Science after Feminism. Oxford University Press.
Lusk, G. (2021). Does democracy require value-neutral science? Analyzing the legitimacy of scientific information in the political sphere. Studies in History and Philosophy of Science Part A, 90, 102-110.
Resnik, D. B., & Elliott, K. C. (2015). The Ethical Challenges of Socially Responsible Science. Accountability in Research, 23(1), 31–46.
Rolin, K. (2021). Objectivity, trust and social responsibility. Synthese 199: 513–533.
Rolin, K. (2024). Earning epistemic trustworthiness: An impact assessment model. Synthese, 203: 39.
Schroeder, S. A. (2021). Democratic values: A better foundation for public trust in science. The British Journal for the Philosophy of Science, 72(2), 545–562.

Outsourcing of inference in clinical medical research and its practical implications

Tatsuyuki Sato
Hokkaido University
sato.tatsuyuki.i2@elms.hokudai.ac.jp

Clinical medical research and basic biomedical research are two pillars of medical science. Clinical research typically investigates whether interventions produce detectable effects in patients, while basic research investigates the mechanisms of those effects. Although this division is familiar in practice, the epistemic difference between the two has not been clearly articulated. Here, I argue that the two differ not merely in subject matter but also in the location of the core inferential step. In clinical research, inference is situated at the collective level, whereas in basic research it lies at the individual level. This structural difference implies that preregistration, which has shown effectiveness in clinical research, may not be suitable for basic research.

Clinical research, including randomized controlled trials and observational studies, aims to establish statistical associations between an exposure and an outcome. A typical study reports the number of participants, the number of outcomes, and statistical measures such as effect sizes and p-values, but refrains from making general claims, such as "drug D is effective for condition C." The inferential step from data to generalized claims and clinical recommendations is intentionally deferred to collectives such as guideline committees, a structure I refer to as an outsourcing of inference. Individual researchers are responsible for producing high-quality statistical evidence, while general causal claims are proposed by collective bodies in the form of guidelines and statements.

Basic research, by contrast, investigates mechanisms of disease and bodily processes. It aims to understand mechanisms by integrating multiple experiments and constructing models of underlying biological processes, such as "molecule X phosphorylates molecule Y, thereby upregulating transcription factor Z and improving cardiac angiogenic signaling." In this context, the crucial inferential work involving model construction and causal interpretation occurs at the level of individual researchers. Theoretical commitments are introduced at the individual level rather than the collective level.

In philosophy of science, there have been extensive discussions of evidential pluralism in asserting causal claims, primarily focusing on the balance between statistical and mechanistic evidence. My account adds a novel layer of argument to the issue from the perspective of social epistemology, showing that the two types of evidence are not only different in their subject matter, but also in the social structure in which they are produced.

This account has practical implications for efforts to improve scientific practice. Preregistration has become the norm in clinical research, and there is now a movement to extend it to basic biomedical research to address the reproducibility crisis. However, these efforts have not advanced as much as expected. The argument proposed here may suggest that preregistration is well-suited to clinical research, where individual researchers can focus on reporting what was done and what was observed, but not suitable for basic research, where data curation and inference proceed in an entangled manner within individual researchers. Improving medical science, therefore, may require moving beyond a one-size-fits-all approach and instead tailoring strategies depending on the epistemic architectures of the target type of research.

Microbial Entangled Lives: Exploring the Interconnections of Metabolism through Ecofeminism and Indigenous Soil Wisdom

Tamar Schneider
The Open University of Israel
tamarsc@openu.ac.il

The challenges in microbiome studies are understanding the causal relations between the microbiome and health or disease (Lynch et al., 2019). Population studies aim to link microbial composition (i.e., who is present and in what quantities?) to the individual's physiological condition (Ma et al. 2024; Maghini et al. 2025). Similarly, soil microbiome examines the correlation between microbial diversity and soil fertility (Mganga et al. 2024). These correlation studies seek to identify the core microbial taxa that are either beneficial or harmful and to treat diseases or improve land fertility. However, a shift towards investigating microbial metabolism and biogeochemical interactions offers a more effective framework for understanding microbiomes as ecological communities. This approach emphasizes

interdependence and the importance of interactions rather than focusing on composition and biodiversity. (Schneider 2023, 2025). By exploring microbial interactions and their connections with the land, plants, insects, animals, and humans, I offer a contextual framework to redefine the goals of microbiome health studies.
Drawing from ecofeminist scholars, I discuss the African eco-feminist and indigenous soil knowledge (ISK) perspective on the relationships between humans and soil, highlighting the principle of mutual care (Siwila 2022; Randriamaro, Z. 2024; Matsapa and Thelma 2025). These relationships are characterized by temporal and spatial ecological cycles. I explore these knowledge systems and their understanding of soil metabolic relationships and the lessons they provide for understanding and interacting with our microbiomes. I then discuss ways this view can be applied to microbial and host-microbial metabolic relationships that nurture both the land and gut microbiomes.
My discussion connects methodologies and scientific practices with historical, socio-political, and epistemic background beliefs using feminist perspectives of situating knowledges and Critical Contextual Empiricism. First, I discuss the dominant techno-science rationale in both human and soil microbiome studies and their often ignorant approach to Indigenous knowledge systems, while aiming to improve health or soil fertility. This pattern reflects a broader history of exploitation, perpetuating the abuse of land, resources, and communities. Furthermore, I discuss the epistemic consistency between the Western microbial subordination by classification and the germ theory with this colonial notion of subordination and control (Adera et al. 2021; Buthelezi-Dube et al. 2022). Such a view is also consistent with colonial practices in the past and present of controlling, engineering, and reconstructing the environment to use and abuse it, in mining and cheap labor, such as the Congo cobalt and copper mining (Lang, Manahan, and Bringel, 2024).
Then, I examine perspectives on the soil microbiome through the lens of African eco-feminism. This approach emphasizes the importance of relationships rather than the subordination of microbiomes, humans, and soil. These interdependent relationships involve cycles of food production, including planting, growing, harvesting, preparing, eating, decomposing, and ultimately returning nutrients to the land. They reflect the cycles of life, growth, death, decay, and recycling. Integrating indigenous soil knowledge with African ecofeminism, I discuss possible approaches to the study of the human microbiome, incorporating indigenous healing practices and food knowledge that connect humans, land, and microbes.

Understanding animal behaviour: What's good for the goose...?

Ludo Schoenmakers
Konrad Lorenz Institute for Evolution and Cognition Research (KLI)
l.l.j.schoenmakers@gmail.com

Scientific understanding has become a central topic in philosophy and epistemology of science. So far, most work has been focused on the physical sciences (e.g. de Regt, 2017; Lawler et al., 2022), while investigations of the role of scientific understanding in the biological sciences have been limited to three studies covering only a small part of the full spectrum (Lopez-Orellana & Cortés-García, 2019; Poliseli, 2020; Varga, 2024, for evolutionary-developmental biology, ecology, and biomedical research, respectively). In order to properly assess the role of scientific understanding in the biological sciences, it is key that a wider spectrum is investigated. In this talk, I contribute to widening the spectrum

by presenting the results of a philosophy of science in practice project aimed at the determining the role of scientific understanding in animal behaviour science.
Animal behaviour science is a unique and comparatively young discipline that brings together physiology, evolutionary biology, developmental biology, phylogeny, and psychology in an effort to understand animal behaviour (Tinbergen, 1963). Crucially, research into animal behaviour is conducted in both laboratory and field environments, using both experimental and observational approaches, and combining both qualitative with quantitative methods. The result is a highly differentiated scientific field in which a wide variety of different methodologies, approaches, and theories are employed. This raises a large number of philosophical questions in and off itself, many of which have been understudied. Here, however, I focus on a specific question: How do animal behaviour scientists working in a field environment come to understand animal behaviour?
In order to answer this question, I combine classical philosophy of science (i.e., conceptual and literature analysis) with participant observation and semi-structured interviews with animal behaviour scientists. Observational work and interview questions are based in part on de Regt's (2017) contextual account of scientific understanding, according to which a phenomenon is understood scientifically when it can be explained based on an intelligible theory T, where the intelligibility of theory T is dependent on the use-facilitating qualities that scientists attribute to it. Observations and interviews have been conducted during and after a one-month research stay at the Konrad Lorenz Forschungsstelle (KLF) in Grünau im Almtal, at the foot of the Austrian Alps, which is one of the foremost research institutes investigating animal behaviour in a field environment in Europe. It is particularly famous for its multigenerational work with geese. Based on my findings, I argue that scientific theories as understood by de Regt (2017) play both a diminished and a different role in animal behaviour science than they do in the physical sciences. This can be explained in part by animal behaviour scientists increased sensitivity to local conditions, and hence reduced emphasis on the generality of theories, and in part by different pragmatic and psychological aspects playing a role in animal behaviour research. Finally, I present an account of scientific understanding in animal behaviour science, and I discuss some of its implications.

References:

de Regt, H. W. (2017). Understanding Scientific Understanding. Oxford University Press.
Lawler, I., Khalifa, K., & Shech, E. (Eds.). (2022). Scientific Understanding and Representation: Modeling in the Physical Sciences. Routledge.
Lopez-Orellana, R., & Cortés-García, D. (2019). On Understanding and Modeling in Evo-Devo. In M. Fontaine, C. Barés-Gómez, F. Salguero-Lamillar, L. Magnani, & √Å. Nepomuceno-Fernández (Eds.), Model-Based Reasoning in Science and Technology: Inferential Models for Logic, Language, Cognition and Computation. Springer Verlag.
Poliseli, L. (2020). Emergence of scientific understanding in real-time ecological research practice. History and Philosophy of the Life Sciences, 42(4), 51.
Tinbergen, N. (1963). On aims and methods of Ethology. Zeitschrift Für Tierpsychologie, 20, 410–433.
Varga, S. (2024). Scientific understanding in biomedical research. Synthese, 204(2), 66.

Transitioning South Africa: an examination of the 'transition' framework in South African gut microbiome research

Elian Schure
Utrecht University
e.c.m.schure@uu.nl

Over the last few decades, the human gut microbiome has received significant scientific attention, particularly in relation to medicine and public health. Within gut microbiome research, there are multiple ways of categorising communities under study. There are racial and ethnic categories, divisions based on 'Western' and 'non-Western', and divisions based on subsistence strategy (e.g. 'hunter-gatherer', 'pastoralist, and 'industrial').
In South Africa, gut microbiome researchers often describe populations under study as 'in transition'. Although this 'transition' is not debated in the research, the exact nature of this transition remains unclear. The transitions are based on shifts in human health states that arise during 'industrialisation' and/or 'urbanisation'. Yet, they focus on various supposed health outcomes, such as the rise of non-communicable diseases ('epidemiological transition'), 'westernisation' ('lifestyle transition'), or health issues caused by processed foods ('diet transition'). The various uses of 'transition' create epistemological confusion; what do these transitions refer to, and where are South Africans transitioning to? I rely on integrated history and philosophy of science, as well as semi-structured interviews, to uncover the origins of the transition narrative and the role this framework plays in current South African gut microbiome research.
I traced the transition framework back to its origins as the 'epidemiological transition', introduced by epidemiologist Abdel Omran in 1971. The epidemiological transition states that a country's industrialisation coincides with a shift from infectious diseases to non-communicable diseases. This theory is widely used in biomedical and public health research, and researchers have developed other theories of transition, as seen in South African gut microbiome studies. The theory faced criticism from historians and public health experts, as Omran sometimes misrepresents historical evidence to support his theory. His theory also claims to be universally applicable, a claim that has been debated for decades. In South Africa, where infectious diseases still coexist with non-communicable diseases, the epidemiological transition does not precisely describe the country's public health situation. Despite these criticisms, South African gut microbiome researchers continue to apply the epidemiological transition and related frameworks.
I also conducted semi-structured interviews with public health researchers and gut microbiome scientists in South Africa. Based on interviews and historical research, I argue that researchers in South Africa employ the transition framework to demonstrate that the gut microbial diversity of the South African population falls between the diversity of 'Western' and 'non-Western' populations. However, 'microbial diversity' is epistemologically unclear, as there are various ways to measure diversity. On a normative level, placing South Africa between the 'non-Western' and 'Western' populations also implies that South African communities are necessarily moving towards 'Western' lifestyles, suggesting that the move towards the West is irreversible. I conclude that the 'transition' framework, which remains prevalent in public health and certain areas of biomedical research in South Africa and the Global South, should be re-examined, as it reinforces ideas of linear progress in populations and is epistemologically confusing.

Methodological Tractability and Explanatory Adequacy: Reconceptualising the Individual-Population Debate in Neuroscience

Maria Serban
University of East Anglia
m.serban@uea.ac.uk

When investigating neural function, researchers can focus on individual neurons—analysing single-unit responses and receptive fields—or study neural populations through simultaneous recordings and collective dynamics. This choice appears increasingly polarised as new technologies make both empirically tractable, while theoretical arguments suggest "the fundamental computational unit of the brain is the population" (Ebitz & Hayden, 2021). Yet evidence from motor neuroscience reveals that both approaches remain productive: multi-electrode arrays record hundreds of neurons simultaneously while refined single-unit techniques continue yielding insights. More puzzling still, different approaches sometimes produce competing explanations—as in motor preparation, where individual neuron persistence competes with population-level null space dynamics (Kaufman et al., 2014)—while elsewhere providing complementary perspectives.

Framing this as a simple methodological choice between individual versus population approaches risks distorting the complex landscape of contemporary neuroscience. When Ebitz and Hayden argue that "there is something new about population-level thinking that simply cannot be understood as a composition of single-neuron concepts," they highlight important emergent properties. Yet this insight coexists with continued productive individual-neuron research across sensory coding, motor control, and cognitive neuroscience, suggesting the choice cannot be resolved through appeals to methodological superiority alone.

This paper develops a two-dimensional analysis that reconceptualises this debate. The apparent methodological opposition conflates two independent dimensions: how we study neural systems (methodological approach) and what explanatory questions we ask (theoretical focus). The first dimension concerns recording techniques—single-unit versus population methods. The second concerns explanatory targets—whether we ask what information is encoded or how computations are implemented. Recognising these as independent rather than coupled dimensions reveals four distinct research approaches, each with different explanatory strengths and limitations. This reframing makes a philosophical contribution: what appears as a technical dispute about recording methods actually reflects deeper questions about explanatory targets and the relationship between empirical tractability and theoretical adequacy.

Rather than develop this framework abstractly, the paper demonstrates these distinctions through Edgar Adrian's methodological trajectory from the 1920s through 1930s. Adrian's progression from single sensory fibre recordings to motor unit coordination studies to population-level EEG investigations illustrates how neural phenomena resistant to individual neuron-level analysis drive methodological innovation. This case allows concrete analysis of tradeoffs between empirical tractability—whether particular phenomena can be measured using available methods—and explanatory adequacy—whether particular approaches can account for observed phenomena.

Adrian's trajectory anticipates contemporary debate structure while operating under severe technological constraints. His single-unit breakthroughs established key neural coding principles—the all-or-none principle, rate coding, sensory adaptation—that remain central to individual-neuron approaches. Yet his work on motor unit coordination and population synchronisation revealed explanatory challenges that individual-level analysis could not resolve, challenges he articulated through an emerging informational framework—messages,

codes, signals—that reconceptualised neural communication. This drove him towards population-level methods, though technological limitations meant population approaches remained empirically constrained, creating persistent gaps between explanatory aspirations and methodological capabilities.
Adrian's case thus raises questions about the relationship between technological capability and theoretical understanding in neuroscience: Did his informational framework emerge from genuine explanatory gaps at the individual neuron level, or from technological limitations in measuring population phenomena? Does this historical case illuminate how contemporary population approaches derive their theoretical frameworks—from novel computational insights or from unprecedented empirical access to population data? The paper concludes by drawing on contemporary motor neuroscience to demonstrate the framework's continued relevance for understanding apparently competing research programs.

Active or Organized Matter: When Biological Experiments Bridge Theory and Governance

Sergei Shevchenko
Ruhr University Bochum
sergei.shevchenko@ruhr-uni-bochum.de

This paper examines how research practices in the early Soviet life sciences forged alliances between scientific theories and their political functions. These alliances were not merely ideological: they were materially enacted and stabilized in experimental systems, where ontological assumptions embedded in epistemic things intertwined with the capacity of these systems to generate and sustain the activities of those epistemic things (Rheinberger 1997; Chang 2011; Rheinberger 2020). It was precisely through the scaling-up of such systems that science's governmental role became visible.
The Early Soviet context is exemplary because, in the 1920s, materialist dialectics, experimental practice, and the political role of science were being shaped simultaneously. Life scientists actively debated non-reductionist conceptions of matter and its capacity for self-organization – ideas that also resonated with international movements in theoretical biology seeking a path between vitalism and mechanism. Within this broader trend, two figures illustrate how experimental systems could anchor distinct ontological and political commitments: Alexander Bogdanov, who emphasized structure and organization, and Alexander Oparin, who emphasized instability and dynamics. Both shaped later Soviet systemic approaches and were entangled in transnational debates on origin of life and biological organization – involving Haldane and Bertalanffy respectively (Hyman and Brangwynne 2012; Şenalp and Midgley 2023).
A trained physician, Bogdanov developed Tectology, a "universal science of organization" that identified parallels in the architecture of natural and social systems. He understood biological change primarily as a restructuring influence exerted by one system on another. Blood, for him, carried the organism's structural profile—its immunological and hormonal configuration, its formal cause. His hemotransfusion experiments, and the institutionalization of a transfusion network, sought to produce predictable restructuring across individuals. The resulting experimental system exchanged structuring elements of blood for mutual benefit, reshaping capacities and labor activities and creating a physiological unity among participants. In this organized circulation, Bogdanov located the contribution of science to a

planned economy: scaling up the transfusion system was to serve as a scientific contribution to socialist organization.
Oparin, by contrast, foregrounded the unstable dynamics of matter. He proposed that life emerged through complex, non-mechanistic movements arising from simpler forms, driven by efficient causes internal to matter itself. His experimental "primordial soup" system, unlike the later Urey experiment, did not rely on external energy inputs: it was intended to exhibit matter's endogenous generativity. This view aligned with a political conception of science as activating nature's productive forces and mobilizing citizens' collective energy. Oparin merged his scientific and agitational roles, popularizing colloid chemistry as a scalable demonstration of matter's intrinsic dynamism and as an instrument for accelerating biochemical development.
Taken together, these cases exemplify broader trends to rethink biological order without reductionism. They reveal two distinct alignments between ontology and governance: Bogdanov's structural matter with science as organizer, and Oparin's dynamic matter with science as mobilizer. In both, experimental practices served as the medium that materially linked research, ontological assumptions, and political purpose, showing how biological experiments could bridge theory and governance.

Integration in Psychotherapy: From Ideals of Unification Towards Assimilative Pluralism

Johanna Silva-Stüger
Department of History and Philosophy of Science, University of Cambridge
js2918@cam.ac.uk

There is a pervasive use of the term integration in the recent history of psychotherapy (research). The field has long been characterised by a plurality of modalities, clustered into psychodynamic, cognitive-behavioural, humanistic and systemic approaches. In the last three decades, "Integration" of these modalities has become a dominant trend, with dedicated journals, handbooks, professional associations, and training programs. Today, around 40 percent of therapists in the US and Canada, and nearly half in the United Kingdom and Australia identify as "integrative" (Norcross and Alexander 2019). This trend warrants a philosophical investigation. Descriptively, we can ask what is meant by integration, and whether different actors mean the same thing by it. On the normative side, what kind of integration is desirable? These are the issues I am discussing in this talk.
The first part of my talk is diagnostic: Integration in the psychotherapy research literature can mean vastly different things. It is instructive to model integration as a two-place predicate ("integrate x with y"), and differentiate four different meanings depending on how one occupies the placeholders: modality integration, integration of therapy with other healing practices, integration of insights from therapy research into therapeutic practice, and integration with neighbouring disciplines such as neuroscience. I then suggest that what unifies the integration talk is a rejection of dogmatism, the latter of which integrationists take as impeding patient-responsive practice. My analysis shows that this view is motivated by Kuhnian ideas of scientific maturity as unification, and shaped by the history of rivalry among therapy modalities.
Focusing on modality integration in the second part, I critically evaluate a canonical four-tier taxonomy of integration attempts: (1) full integration (synthesising theories into one overarching framework); (2) assimilative integration (retaining one primary theory while

incorporating interventions from other modalities); (3) practical eclecticism ("theory-free" selection of interventions originally developed within different modalities); and (4) common-factors approaches (identifying shared change principles across modalities). My analysis of exemplars for each of these approaches reveals that actually just one "kind" of modality integration, namely assimilative integration, lives up to its promises, whereas the others are either instances of assimilative integration in disguise or not integrative at all. Two examples are the so-called Transtheoretical Approach and Integrative Relational Psychotherapy, which I interpret as assimilative behavioural and assimilative psychodynamic therapy, respectively. Based on this diagnosis of the relevant psychotherapy research literature, the last part of my talk focuses on normative issues. Unification ideals motivating integration approaches (1), (3) and (4) as outlined above, are untenable not just due to the complexity of the subject matter (as argued by Mitchell 2002), in our case psychopathology, but fundamentally because of incommensurable aims and values across modalities. Nevertheless, fruitful interactions among modalities remain possible through selective assimilations of elements into coherent systems of practice, especially with respect to aims and values. Building on Chang's (2012) active pluralism and my insights from the study of psychotherapy integration, I will suggest assimilative pluralism as a new normative position within discussions of pluralism and unification in philosophy of science.

Inventing Fluctuations – Philosophical Practices in Looking Through the Microscope

Michael Stoeltzner
University of South Carolina
stoeltzn@sc.edu

For decades the zig-zag motions of pollen particles suspended in a liquid, first observed under the microscope by the botanist Robert Brown in 1827, evaded explanation. These included increasingly sophisticated attempts to invoke the kinetic theory of gases to ascribe them to impacts of the liquid's molecules. Being largely studied by chemists and city hygienists, they were an unlikely candidate to convince skeptical physicists of the existence of atoms. But precisely this happened in the years 1905-9 through theoretical work by Albert Einstein and Marian von Smoluchowski, and the experiments of Jean Perrin. The goal of my paper is to show at the path of von Smoluchowski that the history of Brownian motion, often quoted in the debates between realists and constructive empiricists, did not happen in a philosophical vacuum and teaches some lessons for the philosophy of experiment and scientific realism.

Explaining Brownian motion, I argue, involved the combination of bottom-up experimental research with a theoretically sophisticated and philosophically informed understanding of statistical mechanics and probability. Only then was it possible to accept a quantity of a new kind as lawful: the fluctuations of particle positions. And to recognize that the velocities of the pollen particles that generations had observed under the microscope were merely an artefact of the observation scale. Moreover, Brownian motion won the acceptance for atomism where Boltzmann had failed; despite increased mathematical sophistication his investigations never left the theoretical foundations.

This episode of joint experimental-theoretical-philosophical practice teaches important lessons for present philosophical debates. Experimental traditions may become an epistemic liability despite increased precision. Explorative experimentation sometimes needs conceptual or mathematical directions about the quantities to look for, among them the

frequentist interpretation of probability, not necessarily explicit theories or their predictions. On the other hand, the invention of a new, highly abstract quantity, such as fluctuations of position, also required an accessible domain of experimentation. Both permitted the clear identification of a phenomenon even before Bohr in 1913 provided a sufficiently specific model of the atom. Brownian motion was thus not an instance of theory-mediated measurement – as Smith and Seth have claimed.
Scientific realists often miss that Perrin's Les Atomes justified atomism through a plurality of phenomena that yielded the Loschmidt-Avogadro number. Rather than showing the longevity of theoretical terms, the history of fluctuations thus seems a case in point for Hasok Chang's definition of realism qua operational coherence. Yet while there are clear similarities, the plurality of fluctuation phenomena is of a different kind than the pluralist progress outlined in Inventing Temperature where the operational coherence concerned what seemed an obvious observable. Instead, the bottom-up establishment of a new type of observable, fluctuations, in different measurement processes was only made possible by previously establishing a conceptual framework that allowed fluctuations to emerge before the macroscopic limit of statistical mechanics was reached. Thus, when looking through the microscope at the pollen particles Brown had observed, scientists in 1905-6 not only experienced a gestalt switch from particle velocities to fluctuations of particle positions, but a change of philosophical perspective.

Exposomics' Functional Profile

Predrag Šustar
Department of Philosophy, Faculty of Humanities and Social Sciences, University of Rijeka
psustar@uniri.hr

In this article, I examine the methodological and conceptual assumptions underlying the emerging field of exposomics, a major biomedical initiative that uses big data approaches in investigating relationships between a combination of different environmental exposures (i.e., the exposome) and diseases inflicting an individual organism (see Wild 2005, 2012; Fayet et al. 2024; Wan et al. 2025). Such an endeavor presupposes a clear and defensible notion of disease, malfunction, and dysfunction. Additionally, any naturalistic account of disease should in turn rely on a coherent understanding of a biological function. Nevertheless, even in genomics and postgenomics, the fields on which exposomic research has been conceptually modelled, accounting for function ascriptions has proved highly debatable, as shown by the "ENCODE controversy" (see, e.g., Doolittle 2013). The controversy does not consist simply in an extremely diverging estimate with regard to the functionality of the human genome. That is, the corresponding scientific and philosophical debate revealed radical methodological disagreements about how to detect functions. Moreover, it also emphasized a tension between an overly permissive notion of so-called 'biochemical functions' and a far stricter evolutionary notion tied basically to natural selection processes. Yet, many consider the latter being too narrow for medical purposes, as repeatedly claimed by the ENCODE and upgraded follow-up postgenomics projects.
Now, I argue here that postgenomic and, specifically, exposomic research imply a more combinatorial and, accordingly, middle-ground account of functions than championed by the "ENCODE controversy" itself. In other words, on one side, the one according to which a component-part's function is understood as its regular and robust causal contribution to the operation of a larger system. On the other, in evolutionary contexts, that contribution is

typically cashed out in terms of fitness-enhancement; but, for biomedical purposes, it should be expanded to include contributions to human wellbeing, more generally. The shift under consideration acknowledges that some features beneficial in medical settings may be mere by-products from an evolutionary perspective. But, they still count as functional for explanatory and therapeutic aims. In sum, on my account of functions, the scientific practice of exposomics quite directly suggests not only a rapprochement between an evolutionarily neutral approach and different evolutionarily-laden accounts, but also how to carry out that rapprochement in the functions debate, more strictly. Finally, that might also help in dissipating some misunderstandings concerned with the ways in which the functional assumptions in the ENCODE project and its detractors have been dealt with.

References

1. Doolittle, W. F. 2013. "Is Junk DNA Bunk? A Critique of ENCODE." Proceedings of the National Academy of Sciences of the United States of America 110 (14): 5294–5300. https://doi.org/10.1073/pnas.1221376110
2. Fayet, Yohan, Thomas Bonnin, Stefano Canali, and Elodie Giroux. 2024. "Putting the Exposome into Practice: An Analysis of the Promises, Methods and Outcomes of the European Human Exposome Network." Social Science & Medicine 354: 117056. https://doi.org/10.1016/j.socscimed.2024.117056
3. Wan, Melissa, Elisabeth M. Simonin, Mary Margaret Johnson, Xinyue Zhang, Xiangping Lin, Peng Gao, Chirag J. Patel, Aroub Yousuf, Michael P. Snyder, Xiumei Hong, Xiaobin Wang, and Vanitha Sampath. 2025. "Exposomics: A Review of Methodologies, Applications, and Future Directions in Molecular Medicine." EMBO Molecular Medicine 17 (4): 599–608. https://doi.org/10.1038/s44321-025-00191-w
4. Wild, C. P. 2005. "Complementing the Genome with an 'Exposome': The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology." Cancer Epidemiology, Biomarkers & Prevention 14 (8): 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456
5. Wild, Christopher P. 2012. "The Exposome: From Concept to Utility." International Journal of Epidemiology 41 (1): 24–32. https://doi.org/10.1093/ije/dyr236

Mapping the epistemological, operational, and ethical challenges of genomic risk stratification in oncology: insights from a narrative literature review

Julia Tinland
Aix-Marseille University
julia.TINLAND@univ-amu.fr

The newfound capacity to distinguish between lower, intermediate and higher genomic risk groups in oncology is motivated by the wish to modulate screening frequency accordingly. It is often presented as a way to address issues of overdiagnosis through the de-escalation of screening for low-risk persons. The literature examining the complex challenges generated by these fast-paced scientific developments is rapidly growing, multifaceted and disparate. Indeed, these issues may concern various dimensions, spanning the production of reliable knowledge, its translation into clinical practice and diverse implications for those affected. Moreover, they may fall within the scope of different disciplines, such as public health and

medical research, but also philosophy, biomedical ethics, law, economics, psychology, the social sciences, etc.
Our aim in this presentation is to explore and map out how different scientific fields have appraised the emergence of genomic risk stratification in oncology and the challenges it raises. This, in turn, allows us to offer a philosophically informed interpretation of how risk stratification may transform screening practices.
We opted for a narrative literature review in order to map out as comprehensively as possible how these issues emerge from this wide-ranging, rich and still growing body of work, as well as to provide a philosophically reflective, discursive and critical interpretation of our findings. Narrative approaches provide in-depth qualitative analysis of complex, interdisciplinary literature and are particularly well suited to explore innovative and recent scientific phenomena.
We highlight how (1) epistemological uncertainties—such as the limited predictive validity of polygenic risk scores, the arbitrariness of risk thresholds or the lack of ancestry population diversity—reverberate into (2) operational domains, necessitating adequate integration into existing screening regulations and policies, planning for the disclosure of (sometimes incidental) results and informing genetically-related people, dealing with long-term follow-up and implementing sufficient data protection. These operational challenges, in turn, give rise to (3) significant ethical concerns regarding over-medicalisation, lack of acceptability, consent, inequalities of access and potential discrimination, as well as the implications of differentiated risk communication for the individuals concerned. We argue that addressing such complexity requires both evidence-based and philosophical inquiries that do not isolate scientific validity from its social and normative consequences and are sensitive to the situatedness of these practices within diverse institutional and healthcare contexts.
In doing so, we aim to show how the philosophy of science in practice benefits from recognising the deep interconnectedness of epistemic and ethical dimensions, emphasising that knowledge production and justification are intrinsically bound to their normative implications, and from engaging qualitative approaches such as narrative literature reviews.

Evidence-Based Policy and Distributive Epistemic Injustice

Kato Van Roey
University of Antwerp
kato.vanroey@uantwerpen.be

I argue that orthodox Evidence-Based Policy (EBP), as an institutional epistemic practice privileging evidence from randomized controlled trials (RCTs) and quasi-experimental methods (QEDs), can produce distributive epistemic injustice. I develop this claim through the case of small community-based non-profit organisations (NPOs), showing how orthodox EBP's evidence hierarchies suppress valuable situated forms of knowledge. I then suggest that an EBP+ approach may alleviate these problems.
EBP's privileging of RCTs as the gold standard method to gauge programme effectiveness has, in practice, caused law- and policymakers to assume that RCTs and QEDs are the only permissible methods for obtaining evidence to decide which social programmes work and merit funding (see e.g. Family First Prevention Services Act 2018). As a result, NPOs that do not provide policymakers with evidence from RCTs or QEDs risk missing out on funding in favour of those who can. However, as Marwell and Mosley (2025) argue in Mismeasuring Impact, there appears to be a deep methodological mismatch between the demands of RCTs and the realities of the non-profit sector. Because many social programmes operate at a small

scale in highly contextualized settings, experimental evaluations often result in either null effects, e.g. due to underpowered studies (p. 36), or what they term 'dumb effects': unreplicable positive effects or findings that are irrelevant to real-world practice (pp. 43-51). Moreover, conducting experimental studies disproportionately burdens smaller NPOs — often led by people from, and serving, marginalized communities — who rarely have the resources to run them without compromising their other operational activities (e.g. p. 60, pp. 99-102; 120-121). Consequently, many programmes "will never have the chance to be labelled 'effective' or 'evidence-based'" (Marwell and Mosley 2025, p. 134).
These structural obstacles to producing relevant knowledge raise broader questions of distributive epistemic injustice. Drawing on Coady's (2010) right-to-know account, I argue that the requirement to conduct RCTs or QEDs to secure programme funding constitutes not merely a methodological mismatch but a violation of these NPOs' right to produce knowledge that is in their interest to know. Indeed, EBP facilitates the production of knowledge that policymakers assume they need while hampering the locally situated knowledge production of these NPOs. More specifically,
(a) it obliges these NPOs to collect knowledge that is often epistemically unreliable;
(b) when reliable, the knowledge is frequently irrelevant or inactionable for improving their social programmes;
(c) and the demands of experimental studies constrain their capacity to generate relevant and actionable situated knowledge, resulting in "undone science", unjust ignorance and contributory injustices.
Thus, advancing RCTs and QEDs as the only valid assessment methods in this sector seems to be epistemically and ethically undesirable.
Finally, I suggest that an EBP+ methodology (Shan & Williamson 2023) — which abandons evidence hierarchies by emphasizing the integration of association studies with mechanistic evidence to assess programme effectiveness — might be epistemically more reliable and might help mitigate some distributive epistemic injustices, because it is better attuned to the types of evidence these NPOs already gather today.

Why "Predictive Fertility" Fails as a Criterion of Theory Choice in Cosmology?

Andrés E. Vázquez-Quijano
Faculty of Philosophy and Letters, National Autonomous University of Mexico
andresemm.vazquez@gmail.com

Here, I examine the methodological soundness of "predictive fertility" as a criterion for resolving theory choice problems, with a particular focus on contemporary cosmology. Merritt (2020) and McGaugh (2015) argue that competing cosmological theories should be assessed in terms of their "predictive fertility", that is, the extent to which they generate successful novel predictions. On this view, Modified Newtonian Dynamics (MOND) should be preferred over the standard model of cosmology, since it has allegedly produced a series of novel predictions later confirmed by observation (Merrit 2021). If this assessment is correct, however, it leads to a methodologically troubling consequence: both the Big Bang theory and the standard cosmological model ought to have been rejected, insofar as they have failed to yield genuine novel predictions.
The Cosmic Microwave Background, traditionally regarded as a paradigmatic case of successful novel prediction (Durrer 2015), illustrates the difficulties involved in applying this criterion to cosmological practice. Its complex epistemic history complicates its classification

as a genuine prediction: the initial detection was not recognized as such, and subsequent theoretical developments required substantial modifications in order to account for its observed features (Partridge 2019). This, together with the conclusions drawn above, stands in tension with actual scientific practice in cosmology and raises serious doubts about the methodological soundness of "predictive fertility" when applied to this field.
In what follows, I argue that predictive fertility is a methodologically defective criterion for evaluating cosmological theories, given the distinctive epistemic conditions of cosmological research—namely indirect epistemic access, unrepeatability, lack of experimental manipulability, and uncontrollable environmental disturbances (Massimi and Peacock 2015). In light of these constraints, I propose alternative criteria aimed at capturing scientific practice in this domain.
I proceed in four steps. First, I introduce the predictive fertility criterion and examine its application to two central theory choice problems in cosmology: MOND versus the standard model, and the Big Bang versus the Steady-State theory. Second, I challenge the methodological soundness of this criterion by appealing to the specific epistemic conditions under which cosmological inquiry is conducted. Third, I propose alternative criteria for assessing theoretical alternatives, grounded in the historical character of cosmological science (Hull 1975). Finally, I offer some concluding remarks.

References:

Durrer, Ruth. 2015. “The cosmic microwave background: the history of its experimental investigation and its significance for cosmology.” Classical and Quantum Gravity 32 (12): 124007.
Hull, David L. 1975. “Central subjects and historical narratives.” History and theory 14 (3): 253–274.
Massimi, Michela, and John Peacock. 2015. “The origins of our universe: laws, testability and observability in cosmology.” In Philosophy and the Sciences for Everyone. Routledge.
McGaugh, Stacy S. 2015. “A tale of two paradigms: the mutual incommensurability of ΛCDM and MOND.” Canadian Journal of Physics 93 (2): 250–259.
Merrit, David. 2021. “MOND and methodology.” In Karl Popper’s Science and Philosophy. Springer.
Merritt, David. 2020. A philosophical approach to MOND: Assessing the Milgromian research program in cosmology. Cambridge University Press.
Partridge, R. B. 2019. “The cosmic microwave background: from discovery to precision cosmology.” In The Oxford Handbook of the History of Modern Cosmology. Oxford University Press.

What Makes Big Biology Worth Pursuing? New Dimensions of Scientific Value

Parvathi Vijay
Université Bordeaux Montaigne
parvathi.vijay@etu.u-bordeaux-montaigne.fr

Ignazio Arrabito
Université Bordeaux Montaigne
ignazio257.ia@gmail.com

The emergence of large-scale collaborative initiatives in biology – such as the Human Genome Project (HGP), ENCODE, and the Human Cell Atlas (HCA) – has reshaped how biological knowledge is produced, validated, and utilised. These projects transcend traditional experimental paradigms by blending hypothesis-driven inquiry, infrastructure development, and technological innovation. However, existing philosophical frameworks for evaluating scientific "pursuitworthiness", i.e., the criteria determining whether a research endeavour merits investigation, struggle to assess such initiatives. Developed primarily for small-scale hypothesis-testing or physics-inspired "big science", traditional models inadequately capture the hybrid nature of big biology, which functions simultaneously as knowledge generator, technological incubator, and community resource.
The paper examines the limitations of current pursuitworthiness frameworks through comparative case studies of three landmark biological projects – The Human Genome Project, The Encyclopedia of DNA Elements, and the Human Cell Atlas. We argue that their value stems not just from answering specific research questions but from enabling diverse downstream investigations, democratising technologies, and establishing shared standards. The analysis reveals tensions between big biology's organisational realities and philosophical assumptions. Unlike physics projects with relatively centralised facilities and stable theories, biological initiatives navigate irreducible complexity, decentralised collaboration, and evolving societal expectations. Funding cycles prioritising rapid deliverables further complicate evaluations, as projects must balance immediate outputs with long-term infrastructural goals. To address these challenges, the paper recommends the integration of a multidimensional system for assessing pursuitworthiness in large-scale biology; recognising that a project's worth evolves across phases – from feasibility testing to community adoption. To illustrate this, we propose a tripartite framework considering epistemic, infrastructural, and technological dimensions of value. The analysis highlights the need for dynamic, context-sensitive evaluation that balances accountability with epistemic humility, and recognises the delayed and distributed impacts of such initiatives. We call for models that align with biology's evolving practices while inviting further refinement of criteria to better capture the multifaceted contributions of large-scale projects in the life sciences.

More-than-Human Explanation: Epistemic Payoffs of non-Sociocentric Historiography

Tvrtko Vrdoljak
University of Cambridge
tv328@cam.ac.uk

Recent decades have seen the emergence of "more-than-human" (MtH) historiography: accounts that integrate nonhuman dynamics (e.g., climate, disease, ecology) into the explanation of historical phenomena. Yet the philosophical significance of this development remains underexamined. This paper offers a rational reconstruction of Kyle Harper's The Fate of Rome (2017) as a paradigmatic instance of MtH historical explanation and uses it to develop a general argument for the epistemic advantages of non-sociocentric approaches to social inquiry. By "non-sociocentric," I mean explanatory practices that do not privilege human social variables a priori but treat both human and nonhuman processes as potentially difference-making, depending on the explanandum at hand.
I argue that Harper's distinctiveness can be reconstructed through two concepts from the philosophy of science: causal selection and triangulation. First, a Hart and Honoré-style abnormality-based analysis of causal selection shows that Harper reorganises explanatory

salience in the long-debated collapse of the Roman Empire. Whereas standard sociocentric accounts foreground political overreach, social fragmentation, and external military pressures, Harper treats two nonhuman shocks, the Late Antique Little Ice Age (536–660 CE) and the Justinianic plague, as the decisive deviations from the "normal course of events" that rendered these familiar intra-social stresses fatal. This shift is not merely additive, it transforms the explanatory architecture by repositioning human and nonhuman factors within a trigger–stressor causal structure.

Second, Harper's evidential practice is best understood as triangulation, the coordinated use of methodologically independent lines of evidence. His reconstruction draws together paleoclimate proxies (ice-core sulfate layers, tree-ring chronologies), ancient pathogen DNA, bioarchaeology, settlement archaeology, and textual sources. I show that this triangulation advances three epistemic aims (robustness, epistemic expansion, and consilience) and thereby enabling Harper to detect and justify causal claims that remain inaccessible within the evidential constraints of sociocentric historiography.

Having reconstructed Harper's methodology in these terms, the paper evaluates the rationale and limits of sociocentrism itself. I identify two of its strongest justifications—its concern with avoiding determinism and its reliance on a disciplinary division of labour—and argue that both lose force once we recognise (a) the deep contingency of nonhuman dynamics and (b) the historical rather than epistemologically necessary character of disciplinary boundaries. Because sociocentrism narrows admissible causes and evidential sources upstream (prior to investigating the explanandum) it structurally risks under-detecting difference-makers and thereby generating incomplete or distorting explanations. In contrast, the non-sociocentric stance enlarges the space of admissible causal candidates and broadens the evidential base, increasing the likelihood of tracking historically operative factors wherever they arise.

Finally, I argue that these epistemic payoffs generalise beyond Harper's case. Drawing briefly on other MtH scholarship in historiography (e.g., Crosby, McNeur, Black), I show that non-sociocentric explanation enhances inquiry across diverse historiographic modes (colligatory synthesis, urban political history, and institutional history) by enabling causal, interpretative, and synthetic insights foreclosed under sociocentric constraints.

Overall, the paper demonstrates that MtH historiography is not a niche environmental subfield but a philosophically significant reorientation of social explanation, one that clarifies how causal selection and evidential reasoning ought to respond to the entangled human–nonhuman (more-than-human) structure of the past.

From Inductive Risk to Policy-Warrant Risk: Rethinking the Justificatory Role of Systematic Reviews

Qianru Wang
Ghent University
qianru.wang@ugent.be

In contemporary evidence-based policymaking, it is widely treated as axiomatic that trustworthy guidance should rest on systematic reviews (SRs) of the relevant research. Guideline handbooks from bodies such as the Institute of Medicine/National Academy of Medicine and the WHO explicitly state this. In parallel, the wider evidence-based policy literature describes SRs as the "industry standard" and "best source of information for making clinical and policy decisions". These formulations function not merely as methodological recommendations, but as a governance device that elevates SRs into a

primary source of institutional authority. Acting in accordance with an SR-based guideline is presumptively taken to be responsible, accountable, and hence legitimate.
The paper reconstructs this justificatory pattern as the doctrine of SR-based Authorization (SRA): methodological rigor in the form of systematic review confers (at least prima facie) legitimacy on policy decisions. On this picture, by minimizing scientific uncertainty about "what works" (i.e., by successfully governing inductive risk), one simultaneously supplies sufficient public justification for the policy decision of "what ought to be done" (i.e., the risk of failed public justification is treated as resolved). Part of the attraction of SRA lies in a tempting shortcut in handling of moral and political responsibility: once we have done the best that can be done with the evidence, we are inclined to treat our broader responsibilities as if they had thereby been discharged. I shall refer to this as a justificatory short-circuit.
Empirically, however, evidence-policy and research-utilisation literatures document a stubborn disconnect between the epistemic prominence of SRs and the justificatory work they are supposed to do in complex policy arenas. SRs are designed to produce generalised, relatively context-insensitive claims about efficacy, whereas legitimate policy decisions require public reasons that speak to concrete contexts, value trade-offs, burden distribution, and political accountability. In response to this disconnect, much of the research-utilisation tradition treats it as an implementation or communication problem, focusing on how to improve the uptake of SRs through better communication, incentives, or procedures, rather than on whether SRs are, in principle, capable of supplying the kind of justification public policy demands. This paper argues that the root of the disconnect lies deeper. It lies in a categorical mistake that conflates the logic of research with the logic of policy.
Conceptually, the paper distinguishes inductive risk (the risk of factual error) from policy-warrant risk (the risk of public justification failure in policy decisions). Based on that, the paper proposes an alternative orientation, the view of SR as License (SRL). On this view, SRs are indispensable technologies for governing inductive risk. They provide an epistemic license to rely on certain factual premises. But they do not, even prima facie, authorise policy decisions, in that policy-warrant risk must be addressed through independent procedures of public justification. The contribution of the framework is both diagnostic and reconstructive. It clarifies what SRs can and cannot do in the justificatory structure of public decision-making, and it sketches a more defensible way of situating evidence within democratic governance.

Is the Solution to the Replicability Crisis in Science Scientific Audit?

Henry Weatherburn
University of Aberdeen
henry.weatherburn76@alumni.imperial.ac.uk

Replicability is widely regarded as a cornerstone of scientific reliability and key to science's epistemic authority. Scientific findings are trusted, both within expert communities and by the public, because results are expected to be independently repeatable under relevantly similar conditions. In recent years, however, large-scale replication failures across multiple disciplines have given rise to what is commonly described as a "replicability crisis". This study examines the nature of that crisis, conceptual ambiguities surrounding it, and to what extent existing philosophical and methodological proposals have succeeded in addressing it. I begin by analysing the emergence of the replicability crisis and the philosophical issues it raises, particularly epistemic trust, scientific self-correction, and the social organisation of

knowledge production. A central obstacle to progress, I argue, lies in inconsistent and overlapping use of key terms such as replicability, reproducibility, and repeatability across disciplines. After reviewing major frameworks, including the International Vocabulary of Metrology (VIM), the National Academies of Sciences, Engineering, and Medicine Reproducibility and Replicability in Science Report and Nature Human Behaviour's. I propose a functional taxonomy that distinguishes empirical replicability, computational reproducibility, metrological precision conditions, and generalisability. This taxonomy is defended as pragmatically necessary for audit, governance, and cross-disciplinary evaluation, while remaining sensitive to epistemic pluralism.

I then survey philosophical diagnoses of the causes of the replicability crisis, in four overlapping domains: publication practices, statistical methodology, experimental design, and social-epistemic structures. While numerous reforms, including open science practices, pre-registration, statistical reform, and changes to incentive structure, have been proposed, I argue that these measures, though valuable, have had limited success in producing systemic change.

Against this background, the core contribution of this work is the development and defence of scientific audit, embedded within quality management systems (QMSs), as a generalisable solution to the replicability crisis in the quantitative sciences. Drawing on detailed radiation dosimetry, medical laboratory science, and national metrology institute case studies, I show that structured audit regimes, combining independent measurement (where feasible), data verification, procedural oversight, and documentation, have already resolved replicability crises in safety critical scientific domains. These examples demonstrate the transformation of replicability failures from systemic threats into managed risks through institutionalised audit cycles.

I propose a tiered model of audit. Where independent re-measurement is possible, scientific audit provides a strong form of epistemic assurance grounded in replicability. In cases involving unique events, longitudinal studies, or clinical trial, where re-measurement is impracticable, computational reproducibility audits offer a weaker but still meaningful form of confirmation. I further argue for the extension of this framework to quantitative social science, while recognising that qualitative social science lies outside the practical and philosophical scope of replicability audit due to its interpretive epistemic aims.

My overall conclusion is that scientific audit, when independently governed and transparently implemented within QMS frameworks, offers a robust, philosophically grounded, and operationally feasible resolution to the replicability crisis in the quantitative sciences. In doing so, it reconfigures scientific self-correction from an aspirational ideal into a measurable and enforceable practice, thereby strengthening epistemic trust and institutional credibility of science.

Transferring cells from zoos to laboratories: The issue of accessibility in "cellular anthropology" and its epistemic consequences

Yoshinari Yoshida
University of Copenhagen
yoshinari.yoshida@ind.ku.dk

In an emerging area called "cellular anthropology," pluripotent stem cells are artificially generated from different primate species, in particular apes (Prescott et al. 2015). Such induced pluripotent stem cells (iPSCs) of primates are used as models for investigating

species-specific patterns of embryonic development, and compared with each other in order to acquire insights into human evolution. How well do primate iPSCs serve as model systems in such research? While there are various criteria for evaluating experimental organisms and model systems (Dietrich et al 2020), this presentation focuses on the criterion of accessibility. By accessibility, I mean ease of supply of the biological system in question. For example, intensive experimental research on a specific organism relies on stable supply of the organism. This is often dependent on specific institutional settings, such as stock centers and community ethos that prioritizes collaboration. Such institutional settings have played key roles for establishment and maintenance of model organisms and had epistemic consequences (Ankeny and Leonelli 2020).

To characterize the issue of accessibility concerning primate iPSCs, this presentation discusses the results of qualitative research conducted in Japan, UK, and Germany. In vitro technologies allow stem cell researchers to bypass common challenges in accessing experimental organisms. However, the uniquely interdisciplinary status of primate iPSCs leads to a distinct set of challenges concerning accessibility. To explain this, I emphasize that primate iPSCs are both specimens of exotic animals and living cells capable of indefinite proliferation. As specimens of exotic animals, primate cells are often acquired from zoos. This occurs typically when a captive primate dies, which brings significant contingency and unpredictability to primate iPSC research. It also means that the moments that provide scientists with opportunities to obtain samples coincide with moments of tragedy for the zoo. Thus, stem cell researchers have to invest time to establish and maintain good relationships with zookeepers to secure their cooperation. An information platform originally aimed at promoting primate conservation plays a role in establishing such relationships in the case in Japan. As living cells capable of indefinite proliferation, primate iPSCs raise issues concerning responsible use and distribution. I describe a case in which stem cell researchers themselves arrange a special material transfer agreement that imposes unusual constraints on distribution of primate iPSCs in order to secure long-term accessibility.

These features have epistemic consequences. The contingency and unpredictability involved in acquisition of primate cells influence what questions scientists pursue by using primate iPSCs. The necessity of maintaining good relationships with zookeepers also affects orientations and scales of research. The limitations on use and distribution of primate iPSCs put constraints on specific types of collaborations while favoring others. I combine these observations and discuss how accessibility conditions shape knowledge production.

References

Ankeny RA, Leonelli S (2020) Model organisms. Cambridge University Press.
Dietrich MR, et al (2020) How to choose your research organism. Studies in History and Philosophy of Science Part C, 80, 101227.
Prescott SL, et al (2015) Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. Cell, 163(1), 68–83.

Disorders of Consciousness and Varieties of Dysfunction

Michael Young
Massachusetts General Hospital and Harvard Medical School
michael.young@mgh.harvard.edu

Patients who sustain severe brain injury may experience a range of profound disorders of consciousness (DoC), including coma, vegetative state / unresponsive wakefulness syndrome (VS/UWS), and minimally conscious state (MCS). Classification and clinical ascription of these disorders has historically been bounded by assessment of patient behaviors, which have been taken to be our best surrogate marker of underlying conscious experience. However, we argue that this prevailing approach to classifying and diagnosing disorders of consciousness rests on a misleadingly narrow philosophical conception of the function of consciousness and of the possible range of dysfunctions that might affect it. We explore these themes through a range of clinical cases drawn from our real-world clinical experience in neurology. Each case involves apparent disruption of consciousness with differential preservation or deterioration of distinct function(s) subserved thereby, including wakefulness, awareness, attention, memory, sensorimotor activity, executive function, emotion, and language. Through this clinical lens, we shed light on how prevailing DoC nosology and diagnostic criteria are contaminated by a misleadingly narrow view of the function of consciousness that prioritizes a single function (viz., production of intentional, responsive behavior) to the exclusion of other functions. We also explore how recent advances in neurotechnology, including functional neuroimaging and electrophysiology, have problematized prevailing orthodoxy in the clinical neurology of consciousness by enabling detection of conscious states without observable behavioral responsiveness at the bedside. Finally, we explore epistemological challenges of assessing consciousness through neuroimaging and electroencephalography-based signatures of volitional brain responsiveness. Our exploration begins to bridge the gap between philosophy and clinical practice, shedding light on the complex nature of consciousness, function, dysfunction, and its clinically relevant implications for diagnosing and classifying DoC following brain injury, illustrating how philosophical clarity surrounding notions of function and dysfunction may facilitate more nuanced diagnoses, prognoses, and tailored treatment approaches.

References:

Young, M. J., Nagel, T., Mudrik, L., Roskies, A., Diringer, M., & Fischer, D. (2025). Covert consciousness: epistemic limits of its definition and detection. Neurocritical Care, 1-12.
Kriegel, Uriah. "The value of consciousness." Analysis 79.3 (2019): 503-520.
Young, M. J., and Edlow, B.L. "The quest for covert consciousness: bringing neuroethics to the bedside." Neurology 96.19 (2021): 893-896.
Edlow, B.L., Fecchio, M., Bodien, Y.G., Comanducci, A., Rosanova, M., Casarotto, S., Young, M.J., Li, J., Dougherty, D.D., Koch, C. and Tononi, G., 2023. Measuring consciousness in the intensive care unit. Neurocritical Care, pp.1-7.
Diserens, Karin, et al. "A focus on subtle signs and motor behavior to unveil awareness in unresponsive brain-impaired patients: the importance of being clinical." Neurology 100.24 (2023): 1144-1150.
Schnakers, Caroline, et al. "What names for covert awareness? A systematic review." Frontiers in Human Neuroscience 16 (2022).
Curley, William H., et al. "Electrophysiological correlates of thalamocortical function in acute severe traumatic brain injury." Cortex 152 (2022): 136-152.
Edlow, Brian L., et al. "Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies." Nature Reviews Neurology 17.3 (2021): 135-156.

Improving theory building in psychology by contrastive explanation

Luiza Yuan
Institute for Logic, Language and Computation, University of Amsterdam
l.s.b.yuan@uva.nl

Dingmar van Eck
Institute for Logic, Language and Computation, University of Amsterdam
d.vaneck@uva.nl

The paper applies philosophical work on explanation-seeking questions and the contrastive nature of explanations to address the theory crisis in psychology, signaled by scientists as a lack of progress in psychological theory development. One core critique is that psychologists show a one-sided focus on identifying and describing effects rather than procuring explanations for the occurrence of these effects (Cummins, 2000; Fried, 2020; Van Rooij & Baggio, 2021).

We argue that philosophical analyses on the contrastive structure of the explanandum and the 'interest relativity' of explanations (Lipton, 1990; van Fraassen, 1980; Ylikoski & Kuorikoski, 2010) provide useful insights for: (1) developing methodological recommendations for how to move from effects to (deeper) explanations and (2) informing choices about what sort of evidence bases are relevant for validating explanations targeting specific aspects of explananda phenomena.

First, although identified effects in psychology can often be taken as 'bare causal' explanations (Glennan, 2017), the key issue is a lack of practical methodology for how to increase the explanatory power of psychological explanations towards the construction of theories. Contemporary textbooks on psychological research methods (e.g., Howitt & Cramer, 2007; Morling, 2017) show that researchers are trained to ask descriptive ("Are university students anxious?"), correlational ("Do university students who are anxious tend to smoke more?"), or causal ("Is anxiety in university students caused by low self-confidence in their abilities?") questions.

While these research questions are suited for detecting possible causal relata and effects, we argue that additional methodological recommendations are needed for advancing explanations beyond such bare causal forms. Specifically, clarification of the contrastive nature of explanations is needed: Existing methods insufficiently recognize that explanations are procured not simply to answer, "Why X?" but "Why X rather than Y?" questions. The philosophical insight here is that, because the same fact can have multiple foils, these foils can be used to point to different structures of the target phenomena and to acquire a more complete set of explanations. We demonstrate that this strategy is useful in the case of contemporary psychological theorizing on resilience from a dynamic network perspective. Explicating different contrastive explanation-seeking questions ("Why fact X rather than [foil Y]?"), enables researchers to improve explanations, such as by specifying conditions under which the causal relations obtain rather than not and to uncover assumptions underlying its conceptualizations.

Second, given that foil(s) determine which cause(s) are relevant for an explanation-seeking context, improved specification of foil(s) increases the precision of the explanandum. This increased precision, in addition to increasing explanatory power, helps determine which empirical evidence is relevant given the aspect of the phenomenon the explanation is targeting. We show this in the case of theories on suicide. While many theories and models rely on empirical evidence on non-suicidal self-harm or suicide-ideation, specifying the aims

of the theories in fact[foil] form clarifies why such evidence can be in fact irrelevant for evaluating explanations of suicide.
Overall, we demonstrate how insights from philosophy of science can be directly applied to improve existing methodologies in theory-building in psychology. (496 words)

References

Cummins, R. (2000). " How does it work?" vs. “What are the laws?" Two conceptions of psychological explanation. In F. Keil and R. Wilson (eds), Explanation and Cognition (pp. 117-145). MIT Press. https://doi.org/10.1093/acprof:osobl/9780199548033.003.0016
Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. Psychological Inquiry, 31(4), 271-288. https://doi.org/10.1080/1047840X.2020.1853461
Glennan, S. (2017). The new mechanical philosophy. Oxford University Press.
Howitt, D., & Cramer, D. (2007). Introduction to research methods in psychology. Pearson Education.
Lipton, P. (1990). Contrastive explanation. Royal Institute of Philosophy Supplements, 27, 247-266. https://doi.org/10.1017/S1358246100005130
Morling, B. (2017). Research methods in psychology. WW Norton & Company.
van Fraassen, B. (1980). The scientific image. Oxford University Press.
van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high- verisimilitude explanatory theories in psychological science. Perspectives on Psychological Science, 16(4), 682-697. https://doi.org/10.1177/1745691620970604
Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. Philosophical Studies, 148(2), 201-219. https://doi.org/10.1007/s11098-008-9324-z

---

From glaciers to journals: on the "epistemic gap" in field-based geoscience

Jaiden Zak
Department of Geoscience, University of Wisconsin--Madison
jazak@wisc.edu

Andrew Zeppa
Department of Philosophy, University of Wisconsin--Madison
zeppa@wisc.edu

A central insight of the philosophy of science in practice is that philosophical analysis benefits significantly from closer engagement with the activities of practicing scientists (Ankeny at al. 2011). In particular, given the considerable limitations of scientific publications, the reconstruction and evaluation of scientific reasoning must take into account what actually happens in the laboratory, in the field, and in the broader context of science and society. Though the epistemic “gap” or “mismatch” between what scientists do and what they communicate is widely acknowledged (Medawar 1963; Knorr-Cetina 1981; Rouse 1990; Suppe 1998; Schickore 2008), developing a general account of this gap and its significance is complicated by the diversity of practices and communication norms across cultures, disciplines, research teams, and individuals.
When research involves fieldwork, the gap between practice and publication may be even more pronounced, in part because the practical constraints of the field tend to be magnified

and are liable to change rapidly and unexpectedly. As a result, scientists in the field often must adapt their methods, research questions, and research targets on the fly—arguably giving non-epistemic values an outsized role to play in fieldwork relative to laboratory and theoretical work. To bring these differences to light, we draw on the firsthand experiences of researchers doing fieldwork in glaciology to develop a conceptual framework for analyzing the "epistemic gap" in field-based geosciences. We identify several issues that are particularly salient in the communication of field-based research and situate these within a higher-level taxonomy of the details of scientific practice that might be miscommunicated in or absent from scientific publications.
We then evaluate two potential responses to this informational mismatch. First, scientists could incorporate more details of their practice into their reports. This could be accomplished, for example, by loosening the conventional structures of reports or by adding supplemental narrative sections. Second, recognizing that practical and epistemic disincentives may limit such changes, another response is to broaden the recognition of scientists and non-scientists alike that the written record alone cannot adequately answer many questions about science. For example, we might adapt science pedagogy and doctoral training to increase awareness of and responsiveness to these issues in scientific research. Similarly, in humanities disciplines that study science (e.g., philosophy of science and science studies), we might require firsthand laboratory or fieldwork experience, as well as training in qualitative empirical methods, as a standard component of training. We argue that both strategies should be pursued in tandem.

References

Ankeny, R., H. Chang, M. Boumans, and M. Boon, 2011. "Introduction: Philosophy of Science in Practice." European Journal for Philosophy of Science 1, no. 3: 303–7. https://doi.org/10.1007/s13194-011-0036-4.
Knorr-Cetina, K., 1981. The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science. Pergamon International Library of Science, Technology, Engineering, and Social Studies. Pergamon Press.
Medawar, P. B., 1963. "Is the Scientific Paper a Fraud?", Listener, 70, pp. 377–78.
Rouse, J., 1990. "The Narrative Reconstruction of Science 1." Inquiry 33, no. 2: 179–96. https://doi.org/10.1080/00201749008602217.
Schickore, J., 2008. "Doing Science, Writing Science*." Philosophy of Science 75, no. 3: 323–43. https://doi.org/10.1086/592951.
Suppe, F., 1998. "The Structure of a Scientific Paper." Philosophy of Science 65, no. 3: 381–405.

Classifying change and changing classifications: ecosystem kinding in an age of climate change

Andrew Zeppa
Department of Philosophy, University of Wisconsin--Madison
zeppa@wisc.edu

Ecosystems operate simultaneously across multiple temporal scales—from daily population fluctuations to centuries-long climate trends—yet ecosystem classification frameworks have traditionally left temporal dimensions implicit and underspecified. This paper examines how

climate change has transformed this theoretical oversight into an urgent practical crisis, exposing fundamental tensions between the dynamism of ecosystems and the relative stability of taxonomic frameworks. Through philosophical analysis of contemporary ecological classification, particularly the IUCN's Global Ecosystem Typology (GET), I argue that ecologists face an unavoidable trilemma: maintain fixed classifications that track changing ecosystems (risking category obsolescence), allow classifications to evolve with ecosystems (losing temporal continuity), or embrace scale-indexed pluralistic frameworks (multiplying ontological commitments).

This paper traces how ecosystem classification evolved from Clements' equilibrium-based succession theory through the non-equilibrium revolution to the GET's functional-compositional synthesis (Clements 1916; Holling 2001; Keith et al. 2022). While the GET represents a major advance by integrating what ecosystems contain (composition) with what they do (function), it leaves temporal scale largely implicit—a lacuna that climate change makes untenable. As anthropogenic forcing disrupts the slow variables (e.g., climate, soil) that traditionally grounded ecosystem identity, novel species assemblages emerge with no historical analogues, forcing questions about whether these represent genuinely new ecosystem types requiring new taxonomic categories.

Ecological practice reveals that ecologists have adopted an untheorized pragmatic pluralism, matching classification grain to research purposes without explicit temporal specification. This works tolerably well for stable systems but breaks down as transformation accelerates. Different applications—conservation planning (decades to centuries), carbon accounting (annual to decadal), biodiversity monitoring (taxon-specific turnover rates)—implicitly assume different temporal scales, yet the GET provides no formal mechanism for coordinating these perspectives.

This case supports Dupré's (1993) promiscuous realism. Different temporal scales legitimately generate different but equally real ecosystem kinds, each capturing patterns appropriate for particular purposes. I argue that the path forward requires temporally explicit classification. Ecosystem typologies should specify temporal grain (observation frequency), temporal extent (the period of classification validity), relevant process timescales, and expected stability horizons. The GET's spatial hierarchy should be matched by explicit temporal hierarchy, with upper levels corresponding to slow processes, middle levels to intermediate dynamics, and lower levels to fast variables. This approach would make implicit temporal assumptions explicit, enable principled pluralism with clear coordination protocols, and acknowledge the value-laden nature of temporal framing decisions.

References

Clements, F. E. (1916). Plant Succession: Analysis of the Development of Vegetation. Carnegie Institution of Washington Publication Sciences, 242, 1-512. http://dx.doi.org/10.5962/bhl.title.56234.

Dupré, J. (1993). The Disorder of Things: Metaphysical Foundations of the Disunity of Science. Harvard University Press.

Holling, C. (2001). Understanding the Complexity of Economic, Ecological, and Social Systems. Ecosystems 4, 390–405. https://doi.org/10.1007/s10021-001-0101-5.

Keith, D.A., Ferrer-Paris, J.R., Nicholson, E. et al. (2022). A function-based typology for Earth’s ecosystems. Nature 610, 513–518. https://doi.org/10.1038/s41586-022-05318-4.

Abstracts of Symposia

(alphabetical by last name of corresponding organizer)

Frameworks for Reparative Science

Hannah Allen
University of Texas at Rio Grande Valley
hannah.allen@utrgv.edu

Ian Peebles
Arizona State University
ipeebles@asu.edu

Jada Wiggleton-Little
Ohio State University
wiggleton-little.1@osu.edu

This symposium introduces reparative science—a normative approach to the study of science that acknowledges science's role in perpetuating socio-political, epistemic, and personal harms and redirects scientific practice toward repair. The practice of designing reparative science, which is modeled in four talks in this symposium, both acknowledges these past harms, their cumulative effect on the world, and places a distinct eye on a future for scientific practice oriented towards identifying, correcting, and mitigating these harms.
Reparative science in practice, therefore, is descriptive and historical; we begin our philosophical inquiry by applying nuanced, historically embedded investigations into the failures of science. The project is also goal-oriented and aspirational as we turn our attention toward correcting harms produced by a diverse array of scientific activities, including concept formation and utilization, theory adoption, and the development and deployment of technologies. By drawing on historical episodes in diverse domains of science, we develop a framework that incorporates justice, repair, and mitigation in discussions of philosophy of science in practice.
Far from being a mere criticism of science, the reparative science framework identifies the sources of error with a mind to improve science in practice, acknowledging the positive role that science continues to play in many parts of our society.

The work presented in this symposium is actively engaged in reparative science, an approach in the philosophy of science in practice oriented towards interrogating the ways in which scientists and philosophers of science might apply some of the tenets of reparations. Broadly understood, reparations aim to make amends for past wrongs. Construed in the legal sense, reparations are often discussed or instantiated as cash payments made to individuals or states who have been harmed by those states or individuals who have been actively engaged in doing harm. A more robust understanding of reparations, however, as Ta-Nehisi Coates famously outlines, involves not just mere disbursement of payment, but the ability to understand our collective history, its consequences, and to imagine an entirely new state-building enterprise (Coates 2014). Olúfẹ́mi O. Táíwò similarly conceives of reparations as a broad, forward-looking methodology: "reparation is a construction project" (4).
To this end, reparative science in practice has four main components: (1) it recognizes the importance of science in the practice of world-making. Insofar as science is concerned with the acquisition and validation of knowledge, this knowledge is instantiated in theoretical and practical ways in our world: scientific findings help us understand ancient people and the

history of the earth, biological and chemical processes that underpin our bodily and social functions, and pave the way for technological developments. Although such findings often improve quality of life, the history of scientific enterprise has been marked with notable instances of harm.

To that end, reparative science (2) imposes a historical approach as a way of understanding current scientific practice. As Táíwò (2021) argues, a proper understanding of current systems requires an understanding of their history. Reparative science begins with this historical approach to discuss current-day affairs, and also (3) advocates for a process-based approach that acknowledges the social embeddedness of scientific knowledge and practice. Many of the discussions of this nature have taken place in the domain of science and values, where the consequences of scientific activities bear on discussions of the nature of science. Notably, discussions surrounding inductive risk offer a new depiction of science that necessitates the involvement of non-epistemic value judgements in scientific activity (Douglas 2000, Havstad 2022, Steel 2010). However, there are other consequences to scientific work that must be interrogated and resolved. This symposium presents those consequences with a mind to repairing the harms done by an array of scientific activities and proposes a new reparative framework for understanding science.

Finally, reparative science in practice (4) is constructive, forward-looking, and oriented toward a more equitable distribution of scientific resources, knowledge and beneficiaries. Scientific reparation in practice trades in a comfortable view of science for an accurate one, correcting for past epistemic injustices, and suggesting paths forward that improve the trajectory of science. Reparative science offers an integrative framework for discerning how the distribution of resources, influence of values, assessments of error, and evaluations of scientific success, among other facets, have contributed to harm caused by scientists and scientific findings. Articulating a framework of reparative science in practice allows an opportunity to offer loci for intervention, do mitigating work, and amelioration.

Composed of work by early career scholars in philosophy, each of the papers included in this symposium centers on different scientific fields and highlights aspects of reparative science in practice. The breadth of content demonstrates the wide applicability for reparative science in practice, which in turn, has implications for the relationship between history and philosophy of science, the intersection of values in science, political philosophy, and bioethics, and characterizations of scientific success.

Individual paper abstracts

Hannah Allen: “Practicing Pursuitworthiness: Promotoras and Community-Based Science in the Colonias”

The nearly 1,200 mile border in South Texas between the U.S. and Mexico is home to thousands of regulatory-free communities of over half a million people in total, many of whom lack access to most basic infrastructure, including a lack of potable water, sewage systems, electricity, and internet. Referred to as colonias, the Spanish word for neighborhood, these communities were first established in the 1950s when developers purchased large tracts of land that were otherwise not desirable because they were often hilly, difficult to drain or irrigate, and prone to flooding. Colonia developers parceled the land into smaller tracts, reselling it at a low value, but high interest rate, to individuals who otherwise wouldn’t be able to purchase land on their own. To this day, colonias are made up of young, predominantly Hispanic individuals, who are considered low to very low income.

In this paper, I examine one of the approaches to public health interventions taken by the Texas A+M Colonias Program, namely, the training and implementation of promotoras who

are Community Health Workers from colonias across the Borderlands region. I argue that the use of promotoras in colonias in the South Texas Borderlands region is a pursuitworthy public health intervention, one which contributes to scientific success on an epistemic and normative level. Questions of pursuitworthiness — that is, questions as to which scientific endeavors are worth our time and money— often involve consideration of projects which, though they may not be the most promising individually, contribute to a diverse array of approaches scientifically (Santana 2022). Although the approach taken by the Colonias Program might traditionally be considered outreach or an intervention, I argue that this approach be folded into scientific research due to its proven capacity to generate situated knowledge about a community's health and well-being, including individual's environment, chronic disease and barriers to care.
My argument points out that pursuitworthiness judgments are already being made implicitly in funding and policy decisions that marginalize community-based programs and that recognizing initiatives like promotoras as pursuitworthy reframes scientific responsibility in contexts of inequality. Further, I connect pursuitworthiness in this example and reparative science, pointing out the reparative aims baked into the implementation of promotoras in colonias. This approach demonstrates how beginning with normative, justice-based aims can result in fulfilling epistemic aims later in pipeline.

Ian Peeble: "Disentangling Race-Based Discriminations"

The past decade has brought with it increased research and commentary on the creation, persistence, and remediation of race-based discrimination. While this scholarship has produced valuable insights, it also reveals a persistent challenge: the lack of clear conceptual frameworks to differentiate between distinct forms of race-based discrimination. This talk addresses current conflations and ambiguities in the constructs currently used by clarifying four forms of race-based discrimination that shape health outcomes and health care: racism, racial implicit bias, racial anxiety, and xenophobia. Disambiguating and disentangling these constructs allow for more accurate, causally explanatorily powerful, practically useful, and appropriately complex theorizing of these phenomena. This, in turn, can aid in the design of more effective interventions at the micro-, meso-, and macro-level of health care. Additionally, introducing greater conceptual nuance will enable more precise normative evaluations. Ultimately, I aim to illuminate the nature and etiology of these forms of race-based discrimination, as well as how they operate, interact with each other, and can best be remediated. The resulting constructs will offer researchers, policymakers, and practitioners more precise tools for studying and redressing race-based discriminations and their implication in racial and ethnic health disparities.
This comports with the aims of reparative science by targeting the conceptual frameworks used to institute corrective policies in healthcare contexts in the pursuit of reducing racial and ethnic health disparities and promoting better health care for all.

TJ Perkins: "NAGPRA and the Reorientation of Archaeological Practice"

The Native American Graves Protection and Repatriation Act (NAGPRA), first enacted in 1990, is often portrayed as a negative legal constraint on free scientific inquiry (Weiss and Springer 2020). This framing reflects a narrow and outdated conception of scientific practice that prioritizes a small suite of methodologies. The 2024 updates to NAGPRA—which, among other changes, require museums to obtain tribal consent before displaying or researching Indigenous human remains—have already prompted institutional shifts, such as the Chicago Field Museum's decision to cover contested exhibits and initiate repatriation of

thousands of artifacts. At first glance, these developments resemble familiar elements of political reparations, particularly the redistribution of goods to historically marginalized communities. The scientific practices in line with NAGRPRA's guidelines emerge not as an impediment to science but as an instance of institutional change worthy of attention for understanding scientific practice oriented toward ethical responsibility, epistemic humility, and methodological pluralism. Drawing on how practicing archaeologists have engaged with NAGPRA, this paper argues that claims about its negative epistemic effects are overstated and that it exemplifies a reparative science in practice.

First, NAGPRA reorients science toward repair and justice. Scientific institutions have long benefited from the extraction and hoarding of Indigenous remains obtained through violent means (Krupa 2025). Treating such collections merely as research objects obscures the other reasons they may be valuable and perpetuates harm. By centering tribal sovereignty and recognizing the moral status of Indigenous ancestors and the need for repair, NAGPRA integrates justice into the aims of scientific practice. Taken together, these features show NAGPRA to be an exemplar for better science—one that is robust, accountable, and reparative.

Second, NAGPRA cultivates epistemic virtues that enhance scientific rigor. Consultation challenges researchers to reassess assumptions, question entrenched interpretive habits, and remain open to alternative explanations. The law also promotes transparency and reflexivity: scientists must justify claims about affiliation and origin, not only to disciplinary peers but to descendant communities with lived stakes in the outcomes. These practices align with widely endorsed scientific norms such as peer scrutiny and responsiveness to criticism (Longino 1990).

Finally, the fulfillment of NAGPRA guidelines broadens the evidential base of scientific inquiry. Traditional scientific epistemologies have often privileged physical remains and material culture while discounting Indigenous knowledge systems. NAGPRA's consultation requirements bring oral histories, place-based knowledge, cultural expertise, and alternative explanatory frameworks into scientific deliberation. These sources frequently provide information unavailable from archaeological or osteological records alone, enabling more nuanced interpretations of cultural continuity, migration, and identity. Rather than restricting data, NAGPRA expands what counts as evidence and makes for a more pluralistic, robust, and accountable science.

Jada Wiggleton-Little: "Missed in a Missed Disease: Nonwhite Women with Endometriosis and the Right to be Known"

Endometriosis is a chronic disease associated with severe, life-impacting pelvic pain. While it is estimated that endometriosis affects roughly 10% (190 million) of women of reproductive age globally (WHO, 2025). Endometriosis is also considered a missed disease due to its unknown etiology and long diagnostic delays. However, Black women are 50% less likely to receive an endometriosis diagnosis, and this is due in large part to the historic association between endometriosis and whiteness. This paper considers how Black and Brown women with endometriosis have suffered epistemic wrongs due to their being missing within an already missed disease. Appealing to the concepts of the 'right to be known' and 'epistemic reparations,' I argue that the intentional inclusion of Black and Brown voices in endometriosis research is needed to repair the epistemic status of Black and Brown women with endometriosis and to help ensure the recognition of endometriosis in Black and Brown women in the future.

Endometriosis is a gynecological condition heavily influenced by class, and gender politics. In the 18th century, clinicians depicted the typical endometriosis patient as a white, middle-

class woman who had delayed childbearing (Bougie, Healey, Singh, 2019). Fueling this depiction was the assumption that women of color were less civilized and less likely to experience the stress of modern life. Thus, endometriosis was considered ‘rare’ in this patient population. Such ideas remained prominent in gynecology textbooks and popular culture, leading to differential diagnoses based on race. For example, Black women were less likely to receive an endometriosis diagnosis, and instead, were likely to be misdiagnosed with pelvic inflammatory disease—a disease typically associated with sexually transmitted infection and improper gynecological care. This glossing over the Black or Brown woman as typical endometriosis patient not only leads to incomplete knowledge of the etiology of endometriosis, but it also allows medicine to ignore how psychosocial stressors like structural racism influence the expression of endometriosis in marginalized communities.
Nonwhite women with endometriosis not only have the right to know what is potentially happening in their bodies, they also have a right to be known in the knowledge produced about conditions that also impact them. As Lackey and others have argued, the right to be known is a distinct epistemic right that arises out of injustices often perpetrated by institutions. The right to be known is violated when subjects are prevented from giving knowledge about their own experiences because they have been rendered invisible or are systemically distorted in knowledge production. Epistemic reparations are then the intentionally reparative actions in the form of epistemic goods given to those epistemically wronged, reparative actions given to those who have systemically failed to be known (Lackey, 2022). To repair their epistemic status with regards to gynecological pain, the medical community must ‘bear witness’ to how Black and Brown women and their lived experiences have been repeatedly left out of this missed disease

---

Roundtable: Aesthetic Directions in Philosophy of Science in Practice: Charting the Field, and Looking at the Future

Chiara Ambrosio
Department of Science and Technology Studies, UCL
c.ambrosio@ucl.ac.uk

Milena Ivanova
University of Cambridge Professional and Continuing Education/ Leverhulme Centre for the Future of Intelligence/ Newnham College
mi342@cam.ac.uk

Gemma Anderson-Tempini
Artist/Research Fellow, University of Exeter
g.e.anderson@exeter.ac.uk

Alice Murphy
Department of Philosophy - University of St Andrews
amlm2@st-andrews.ac.uk

The aesthetic dimension of science has intrigued historians, philosophers, and scientists alike for a very long time, and yet until very recently it has remained at best on the periphery of the history and philosophy of science community. Despite notable exceptions (see for example Farmelo 2002; McAllister 1996), aesthetic values, judgments, and practices in the sciences seem to have followed a trajectory analogous to that of models, analogies, and representations: relegated to the context of discovery, they were sidelined as dispensable in philosophical justifications of scientific theories and explanations, or forgotten altogether. Fortunately, philosophy of science has moved on. The establishment of integrated HPS as a field in its own right has substantively opened up areas of investigation where a full consideration of aesthetics could acquire a progressively more central role within historically-sensitive philosophical inquiry. Closer collaborations with colleagues in the philosophical fields of aesthetics and philosophy of art, prompted by research on models, representations, and more recently imagination and creativity in science, opened new pathways of research and productive 'trading zones' across these two fields of philosophical investigation. Last but not least, the philosophy of science in practice community has substantively shifted the focus of philosophical investigation precisely to aspects of scientific inquiry where aesthetic practices, values, and judgments have found their natural home, and are fully integrated into accounts of how scientists approach various aspects of their research – from modelling and representing to experimenting, understanding, classifying, and much more.
This roundtable does not have a single organiser. It is a collaborative proposal from all of us four presenters, and it is partly a celebration of work we have done with colleagues over the years to bring aesthetics, philosophy of science, artistic research, and scientific practice in a closer dialogue. We could not think of a more apt community than philosophy of science in practice to take stock of the work we have done thus far, and more importantly to think of new ways for this work to move forward. The panel will start from four collections we published recently: The Aesthetics of Science (Ivanova and French, 2020), Drawing Processes of Life (Anderson-Tempini, Dupré and Wakefield, 2022), The Aesthetics of

Scientific Experiments (Ivanova and Murphy, 2024), and Abstraction in Science and Art: Philosophical Perspectives (Ambrosio and Sánchez-Dorado, 2025). We also want to use this panel as a tribute to the late Helen De Cruz (1978-2025), whose recent Philosophy Illustrated (2021) is a unique example of aesthetic investigation in which the very practice of drawing gives new concreteness and accessibility to the otherwise abstract and elusive genre of thought experiment in philosophy. Aesthetic practice, as seen through the lens of De Cruz's work, can be a mode of philosophical inquiry in its own right, which fully complements, and indeed extends, the approach to aesthetics we aim to highlight in this roundtable.

Taken together, these collections form a coherent body of work that shows the pervasiveness of aesthetics across a range of scientific and philosophical practices. Judgments of beauty or elegance, for instance, play out not just in theory-building and theory choice, but are pervasive in practices of experimenting and intervening upon the world (Ivanova, Murphy). Aesthetic practices such as abstracting remain poorly investigated in philosophy of science, and yet there is a large literature in aesthetics as well as in past and current artistic research that offers a repository of investigative and conceptual tools to understand what is it that scientists do when they selectively disregard or omit information (Ambrosio). And while the close interaction between art and science is highly visible in the history of science where we see scientists with artistic training producing scientific artefacts, such as images of newly discovered entities and processes through newly invented instruments, such as the microscope, today we tend to treat these areas of activity as entirely separate, with their own institutional goals and methods. Yet, recent collaborations in which aesthetic practices are the research show that this separation is not just unjustified, but actually detrimental in that it may preclude the development of novel methodologies and avenues of inquiry that can be beneficial to scientific, artistic, and philosophical practice alike. Gemma Anderson-Tempini's work, featured as a practitioner's contribution in this panel, is a case in point. From her collaborations with philosophers and biologists, to her work with mathematicians, to her residencies in natural history collections around the UK, Anderson-Tempini has used drawing as an exploratory mode of inquiry in which aesthetics functions – perhaps in the most fundamental sense of the term – as a way of extending and reconfiguring experience: moving from substances to processes in cellular research, visualising and exploring four dimensional space and, more recently, investigating the topology of life processes such as the human mycobiome. This is analogous to how De Cruz deploys drawing as a mode of philosophical inquiry in her liberatory reimagining of the function of thought experiments. Indeed, the common thread between De Cruz and Anderson-Tempini's works is precisely that drawing does not merely serve as illustration (in its naïve and still far too common conception of 'accessory' to scientific/philosophical work) but is constitutive of the research. It is an aesthetic and genuinely experimental practice in its own right, which reconfigures and expands the remit of experience.

Thus, this roundtable has three main goals. First, we would like to highlight emerging commonalities in our works, reflecting on recent developments in this area of philosophy of science in practice, and identifying emerging questions for further research. Secondly, we would like to reflect methodologically on the different approaches in this area of research, the ways in which these different methods may afford new ways to collaborate with practitioners, and crucially the spaces and infrastructures (institutional, financial) that this varied and interdisciplinary approach to aesthetic research demands. Thirdly, by highlighting the highly interdisciplinary and collaborative spirit of this growing field of study, we would like to facilitate a conversation between practitioners, philosophers, historians, and sociologists of science, and invite reflections on the future of this field.

The format of the roundtable will align to these goals: we will begin with 10-minute presentations on promising directions in the aesthetics of science emerging from our recent

edited collections (aesthetic values, the aesthetics of scientific experiment, abstracting and drawing as aesthetic practices), and spend the remaining time of the session in a facilitated discussion with the roundtable audience. We hope that a concrete output of the session will be a 'wish list' of future directions of research in the field, which will be collectively drafted in dialogue with the SPSP community.

References:

Ambrosio, C. and Sánchez-Dorado, J. 2025. Abstraction in Science and Art: Philosophical Perspectives. London: Routledge.
Anderson-Tempini, G., Dupré, J. and Wakefield, J. 2022. Drawing Processes of Life. Chicago: Intellect Press.
De Cruz, H. 2022. Philosophy Illustrated. Oxford: Oxford University Press.
Farmelo, G. 2002. It Must Be Beautiful. Great Equations of Modern Science. London: Granta Books.
Ivanova, M. and French, S. 2020 The Aesthetics of Science: Beauty Imagination and Understanding. London: Routledge.
Ivanova. M. and Murphy, A. 2024. The Aesthetics of Scientific Experiment. London: Routledge.
McAllister, J. 1996. Beauty and Revolution in Science. New York: Cornell University Press.

Institutions in Practice: Rethinking Scientific Organizations

Justin Biddle
Jimmy and Rosalynn Carter School of Public Policy, Georgia Institute of Technology
justin.biddle@pubpolicy.gatech.edu

Haixin Dang
Department of Philosophy, Case Western Reserve University
haixin.dang@case.edu

Joyce Havstad
Department of Philosophy, Clinical and Translational Science Institute, University of Utah; SOCRATES Centre for Advanced Studies, Hannover
joyce.havstad@utah.edu

Symposium Description

This symposium explores the relationship between institutional values and organizational arrangements, providing normative reasons why particular arrangements might (or might not) be beneficial for scientific practice. Scientific researchers work in organizations – such as university research labs, regulatory agencies, and firms – with particular organizational structures that influence scientific decision making, including decisions about which projects to pursue, which methods to adopt, and how to manage epistemic and inductive risks. Organizations, in turn, operate within broader institutional environments, or systems of norms that guide behavior and structure social practice. The organizations within which scientists work, as well as their broader institutional environments, affect the ability of science to contribute to individual and social goods – including human flourishing, democratic participation, and the growth of knowledge. The current worldwide threats to these goods cannot be understood without identifying the institutional and organizational arrangements that give rise to them, and these threats cannot be countered without rethinking the norms that guide such arrangements.

Philosophers of science and values have recently turned their attention to institutions and organizations. Much work on institutions and values in science addresses fundamental theoretical questions about the nature and role of institutions. That work remains important and complementary to our aims. This symposium, however, takes a different, practice-oriented tack: we focus on specific normative recommendations for scientific organizations: Torsten Wilholt examines the institutional structures that allow science to promote democratic accountability and legitimacy. He distinguishes between two harmful types of politicization of science – production politicization and authority politicization – and argues that efforts to democratize science must guard against both, if they are to be consistent with democratic legitimacy and warranted trust in science.

Justin Biddle examines different organizational structures for facilitating responsible research and development that promotes human wellbeing. He distinguishes between different types of "ethics boards" – including external advisory committees, internal ethics offices such as Research Ethics Committees (RECs) and Institutional Review Boards (IRBs) and others. On the basis of case studies from organizational practice, he argues that there are fundamental limits to what these boards can achieve.

Haixin Dang explores transparency as an institutional value and examines what it means for an institution or organization to be transparent. She argues that while there might be moral and political reasons for organizations to be transparent, full transparency might impede the epistemic goals of science.

Joyce C. Havstad asks, from an institutional standpoint, what is wrong with "publish or perish"? She argues that it is an instance of value consolidation distortional to proper method choice in science. While drawing on work from social epistemology on value capture, swamping, and collapse (Nguyen 2024), she distinguishes this episode of value consolidation in virtue of its emphasis on a highly constitutive or internal (to science) rather than externally sourced value. Combating the distortional effects of "publish or perish" on method choice will not be as simple as excising an invasive value-species. Instead, that effort requires re-establishing the other competing values integral to proper method choice in science.

Science and the Government: Two Forms of Politicization
Torsten Wilholt

The question of how to manage values in science often leads to calls for greater democratic accountability. However, this requirement must be carefully distinguished from detrimental forms of politicization that threaten the integrity and public role of science. This talk analyzes two distinct, harmful forms of politicization that arise specifically from the relationship between government and scientific institutions, arguing that a core institutional task is to guard against these dangers.

I first distinguish between production and authority politicization. Epistemic production politicization occurs when political powers directly interfere with scientific research to steer it toward certain results or, more commonly, to prevent the generation of politically inconvenient knowledge. Recent examples include the freezing of federally funded projects on topics related to diversity, equity and inclusion. I will argue that the main reason why epistemic production politicization always constitutes a direct institutional failure is best explained on the basis of democratic theory: it undermines democratic legitimacy. Even a democratically legitimate government does not have the right to poison the well from which it draws its legitimacy. Democratic decisions are only legitimate if they are based on a democratic process in which citizens can assert their values and interests. And this ability of the citizens presupposes knowledge. The crucial harm is not the mere existence of restrictions, but the suppression of knowledge that is essential for public debate on ongoing political issues.

In contrast, epistemic authority politicization is a more diffuse, systemic phenomenon. It occurs when scientific issues become entangled in salient socio-political identity conflicts, such that acceptance of facts becomes a marker of political allegiance. The most familiar example is climate change denial, where acceptance of evidence correlates strongly with political affiliation. Unlike epistemic production politicization, this form is often not orchestrated by a single actor but emerges from the interplay of many factors: from the dynamics of social media, from the strategic communication of populists, from our own psychological tendencies towards identity protection. It erodes warranted trust by causing the public to selectively judge scientific facts based on group loyalties.

Politicization of science need not always be problematic, but the two forms I single out have demonstrable negative consequences. The analysis has direct implications for how institutions may incorporate democratic accountability. It demonstrates that democratizing science must not be conflated with subjecting it to governmental control or allowing it to be subsumed by political conflicts. Combating production politicization requires institutional defenses: demanding scientific autonomy, protecting whistleblowers, and resisting censorship. Addressing authority politicization, however, presents a more complex institutional challenge. It calls for scientific communication strategies that clearly delineate the roles of science and politics—providing facts and scenarios without short-circuiting

democratic deliberation by presenting political choices as scientifically predetermined. By clearly delineating these two forms of harmful politicization, this talk provides a framework for institutions to safeguard the indispensable, yet vulnerable, role of independent science in a democratic society.

Managing Values through Ethics Boards: Limitations and Tradeoffs
Justin B. Biddle

Science and technology are value-laden enterprises. Decisions about which projects to pursue; how inductive and epistemic risks should be managed, how research results should be applied, and many others require the application and management of values – including, in some cases, ethical and political values. While individual researchers can and should consider the moral and political consequences of their decisions, they can also create organizational structures – or "ethics boards" – that facilitate the management of values in science and technology. This paper has three primary aims: (1) distinguishing different types of ethics boards; (2) articulating an ideal for how ethics boards should operate; (3) arguing that, in practice, ethics boards face tradeoffs that make this ideal virtually impossible to sustain.
Based on empirical studies of existing organizations, I distinguish three categories of ethics boards: external advisory committees, internal ethics offices, and compliance enforcement (or auditing) bodies. External advisory committees are composed of outside experts who offer guidance on complex ethical questions, typically in the form of non-binding recommendations. Internal ethics offices are units within organizations that work collaboratively with other units to facilitate responsible value management. Research Ethics Committees (RECs) and Institutional Review Boards (IRBs) are examples of internal ethics offices. Finally, compliance enforcement bodies conduct assessments to determine whether research applications meet specified criteria – including ethical criteria – and they can level sanctions in case of noncompliance. These compliance enforcement or "auditing" bodies play increasingly important roles in efforts to govern artificial intelligence (AI) systems.
My empirical study of existing organizations coupled with conceptual analysis reveals three necessary characteristics that must be present for ethics boards to optimally facilitate responsible management of values: independence, information, and empowerment. An ethics board should provide advice that is independent. That is, its advice should not be unduly influenced by other organizational units – especially those that might have conflicting interests, such as marketing departments. An ethics board should have sufficient information to provide advice that is relevant and impactful. And an ethics board should be empowered. As the organizational unit that has the most expertise related to ethics, it should have some ability to influence other units to adjust their practices based on ethical considerations.
And yet, in the real world, it is virtually impossible for an ethics board to sustain all three characteristics. In practice, these three characteristics trade off one another. It is hence feasible for ethics boards to have any two of these characteristics; but having any two will make it next to impossible to sustain the third. Due to these tradeoffs, ethics boards that have two of these characteristics can facilitate responsible management of values, although not optimally. Ethics boards that have fewer than two are unlikely to facilitate responsible management of values – and they might be used to engage in ethics washing.
While I have developed both conceptual and empirical arguments for these conclusions, this talk will examine case studies of ethics boards in practice to illustrate and confirm the limitations and tradeoffs that ethics boards face.

Transparency and Its Discontents: Rethinking Openness in Scientific Institutions
Haixin Dang

In recent years, there have been increasingly forceful calls for greater transparency in science. Within the Open Science Movement, transparency is heralded as a central value—one that promises to make science more accountable, more trustworthy, and more consonant with democratic ideals. At a moment when scientific skepticism and denialism exert growing influence, transparency is often presented as a corrective: a way for science to regain public confidence by opening its practices to broader scrutiny. However, this is too quick; it is far from obvious that transparency can result in the social and epistemic goods its defenders herald. In this talk, I examine what it means for an institution or system to be transparent and whether transparency is well-suited as an epistemic goal for epistemic institutions.

On a first pass, transparency appears straightforward: when institutions are transparent, outsiders have access to information about how agents inside them act; when institutions are secret, that information is withheld. But despite the Open Science Movement's tendency to present transparency as an unqualified good, it is in fact a costly and complex value to operationalize. A growing literature in political science and among scholars of institutional trust challenge the assumption that transparency reliably produces accountability or improves institutional performance. In some cases, secrecy may actually serve democratic ends more effectively than openness. Political philosophers have argued that secrecy and transparency are not opposites but complementary values within institutional design. Transparency reshapes the behavior of agents and the flow of information within an organization, and these changes are not always beneficial. Calls for greater transparency can backfire: knowing that their actions will be publicly examined, agents may engage in strategic behavior, shade the truth, or withhold information, thereby impeding rather than enhancing the proper functioning of the institution. Secrecy allows agents more freedom to act and for agents to be more honest without fear of public scrutiny.

These worries apply directly to scientific institutions. Whether we take science as a whole to be an institution or take scientific institutions (such as journals, societies, funding agencies, etc.) as the target of analysis, philosophers of science should take the downstream effects of transparency on scientific practice seriously. Full transparency can impede scientific research rather than support it. The demand that scientists record every step and dataset "for transparency's sake" can slow down their workflow. Requirements to publicize raw data may lead scientists to avoid recording data they would prefer not to disclose. In short, people change their behavior when they know their actions will be made public, and not all of those changes are ones we should want to encourage. Calls for transparency must be accompanied by serious theorizing about which aspects of scientific practice should be made visible and which ought to remain deliberately secret.

Ultimately, I argue against treating transparency as an epistemic goal for scientific institutions. Instead, we need to examine how both transparency and secrecy might serve the epistemic aims of science and how these values can be balanced. Transparency may be valuable for our social or moral goals, but it may not automatically align with the epistemic goals of science. In some cases, these goals may even be in tension.

What's Wrong with Publish or Perish?
Joyce C. Havstad

Imagine you are an administrator at a scientific institution or funding agency who cares little about the quality of life of individual scientists. What is in your remit is the production of

high-quality science, at quantity—whether valued for itself or the financial, reputational, and other goods you expect that product to generate. Now consider "publish or perish," that high-pressure climate experienced by so many scientists today. From the point of view of an administrator such as this, what if anything is wrong with "publish or perish"?

Here I argue that even if we entirely disregard the detrimental cost individual scientists pay in terms of quality of life under the existing "publish or perish" regime, there is still reason to object to said regime. I draw on work from social epistemology to explicate why. Thi Nguyen (2024) has lately introduced the phenomenon of value capture, along with related phenomena of value swamping and value collapse. Value capture and swamping are both characterized by the transition, via specification, from an initially complex value-state towards an alternative, simplified value-state. I posit that one way to move from a complex value-state to a simplified value-state is to go from being committed to a cluster of values, all of which need to be put into conversation with one another, to either a single value or a unified value-state which only needs to be specified and realized by itself.

Let us call the move from a complex, clustered value-state to a simpler, more streamlined value-state the phenomenon of value consolidation. Externally sourced values are not themselves a signature characteristic of value consolidation. Rather, value consolidation is the process of winnowing down a set of values in a complexity-reducing way that often confers both internally and externally realized benefits of enhanced communicability, confirmability, and standardization. To relate this concept to those already extant in Nguyen's literature, value consolidation is potentially a species belonging to the genus of value swamping, and one that can lead towards value collapse.

Elsewhere I argue that competing-method choice in science is ideally characterized by treating five value-criteria of fruitfulness, salience, useability, potency, and reliability in irreducibly judgment-based consideration with one another. But "publish or perish" puts pressure on scientists to develop and adopt methods which are fruitful (generative of new research findings) regardless of whether those methods are salient (track a phenomenon of interest), potent (confer a lot), and reliable (perform similarly in distinct instances within a shared context of application). Fruitfulness is not a value-criterion external to science, so this is not value capture. Rather, it is value consolidation. What is wrong with publish or perish even from this hypothetical standpoint is that fruitfulness needs to be ordered, weighed, and balanced in accord with salience, usability, potency, and reliability.

Resistant to Reduction: Mindfulness, Hallucination, and Empathy in the Lab and in the Clinic

Agnes Bolinska
University of South Carolina
bolinska@mailbox.sc.edu

Riana Betzler
San Jose State University
riana.betzler@sjsu.edu

Michael Dickson
University of South Carolina
dickson@sc.edu

Symposium

This symposium investigates challenges that arise when trying to study multidimensional, dynamic, and otherwise hard-to-pin-down phenomena using the standard tools and techniques of the biomedical sciences. While all of these phenomena are of great theoretical and practical interest, including in clinical contexts, they are, in one way or another, resistant to reduction. Through analysis of three important cases–mindfulness, hallucination, and empathy–this symposium untangles the reasons for this resistance to reduction and opens up conversation about how scientific investigation of these phenomena might proceed more fruitfully in the future.

Riana Betzler's paper critiques the idea that empathy can be delivered in "doses." In doing so, it demonstrates some of the pitfalls associated with attempts to study dynamic, interpersonal phenomena using the tools of evidence-based medicine–especially the randomized controlled trial (RCT). While these efforts appear to demonstrate the effectiveness of empathy, they do not disambiguate the multiple roles that different dimensions of empathy might play within the clinical encounter, and therefore are of limited use to practicing clinicians.

In a similar vein, Agnes Bolinska's paper raises questions about the suitability of RCTs for researching mindfulness as a treatment for illness. Using an RCT of a program called Mindfulness-Based Stress Reduction as a case study, it challenges the idea that mindfulness can be considered an "active ingredient," which can be isolated from the programs within which it is "delivered" and accurately measured.

Michael Dickson's paper addresses recently developed experimental protocols for the study of hallucination. While these protocols might enable the study of brain activity associated with 'perception without an associated external cause', they fall short of enabling the laboratory study of the experience of hallucination, because that experience has features that are not (at least not at present) reproducible by the protocols. Moreover, in its clinically relevant sense, 'hallucination' is not correctly characterized, reductively, as something like 'false perception plus context'.

Together, our papers show how a reductive approach to studying complex phenomena may misrepresent and distort those phenomena. This problem is especially pressing when conclusions from such studies have important clinical applications.

Paper Abstracts:

1) A Dose of Empathy?
Riana Betzler

The title of a recent general-audience article by the philosopher of medicine Jeremy Howick reads, "A dose of empathy is all some patients need. "Howick, a champion of evidence-based medicine, describes the outcome of a recent meta-analysis of the effectiveness of empathy in medicine (Howick et al., 2018). The results are impressive. Empathy appears to have several physiological benefits, such as reducing the duration of the common cold, as well as interpersonal benefits, such as improving communication and decision-making in the medical encounter. Howick also points out that empathy is much less risky than other medical interventions. Thus, we should all be clamoring for our dose of empathy.
Howick's work is representative of a recent trend that attempts to make traditionally humanistic features of medicine (such as empathy) evidence-based. Such views consider empathy to be an intervention, to be delivered in doses, much as a pill might be. We can also, according to these views, analyze the effects of empathy using the gold standard of evidence-based medicine: the randomized controlled trial. By bringing humanistic medicine under the framework of evidence-based medicine, authors working within this trend aim to provide hard evidence for the intuition that "soft skills" such as empathy are genuinely important parts of medical practice and not just frills on top of the real work of diagnosis and treatment. This paper examines the challenges associated with bringing humanistic medicine under the framework of evidence-based medicine by focusing on the research on empathy. I begin by raising conceptual issues about empathy that make integrating randomized controlled trials from various sources into meta-analyses especially difficult. I also consider whether it makes sense to think of empathy as an intervention—as something that we can give "doses" of as we might pharmaceuticals. I draw a parallel with research on the effectiveness of psychotherapy to illustrate the types of methodological challenges associated with using randomized controlled trials to assess the effectiveness for empathy. I conclude by arguing that empathy is not the sort of thing that comes in "doses" or that functions as a discrete medical intervention and that the randomized controlled trial is not the right way to assess the effectiveness of empathic medical care. Empathy is, instead, a skill comprised of multiple competencies that must be deployed flexibly and sensitively in response to the demands of individual situations and needs. Highlighting the skill-based nature of empathy has several consequences for medical training and evaluation. Finally, I suggest that we ought to be open to forms of evidence beyond the randomized controlled trial when thinking about the importance of empathy and other humanistic features of medicine.
References:
Howick, J. (25 April 2018). A dose of empathy is all some patients need. https://www.jeremyhowick.com/dose-empathy-patients-need/
Howick, J., et al. (2018). Effects of empathic and positive communication in healthcare consultation: a systematic review and meta-analysis. Journal of the Royal Society of Medicine, 11(7), 240-52.

2) Mindfulness as an "Active Ingredient"?
Agnes Bolinska

In recent years, there has been an exponential growth of research supporting mindfulness—a non-judgmental awareness of moment-to-moment experience—as a treatment for conditions such as chronic pain and depression. Yet mindfulness is often "delivered" to patients within a

more complex package. For example, in the eight-week mindfulness-based stress reduction (MBSR) program, participants meet weekly in groups to learn meditation and yoga, practicing them daily as homework. Researchers have sought to identify what is responsible for the beneficial changes that participants experience: Is mindfulness MBSR's "active ingredient," or do these changes result from other components of the program, such as its supportive structure or participants' expectations that it will help?

To answer this question, researchers have designed randomized controlled trials comparing MBSR to a control condition called the Health Enhancement Program (HEP). HEP mirrors MBSR's structure but replaces mindfulness training with music therapy, functional movement, and nutrition education (MacCoon et al. 2012). Mindfulness can be considered the active ingredient of MBSR if MBSR but not HEP increases mindfulness, and if MBSR is more effective than HEP.

In this paper, I raise several problems for this approach to mindfulness research. First, I argue that HEP is an inadequate control condition for MBSR. Despite researchers' intentions, it plausibly also trains mindfulness. HEP's music therapy component, for instance, begins by instructing participants to still their minds and focus on a piece of music, and its functional movement component includes training in proprioception, a kind of bodily awareness.

Second, I raise doubts about the possibility of measuring mindfulness. Self-report measures of mindfulness have been criticized for being biased (Grossman and van Dam 2011). Measuring mindfulness indirectly using a thermal pain task, in which lower perceived intensity and unpleasantness of hot stimuli indicate higher levels of mindfulness, is thought to be less prone to the biases than self-report measures (MacCoon et al. 2012). I argue, however, that the thermal pain task does not measure mindfulness better than self-report measures do. In order for the thermal pain task to be a better measure of mindfulness, (i) it must be a valid measure of mindfulness and (ii) it must measure mindfulness more accurately than self-report measures. It's far from obvious that the thermal pain task meets the first condition: although decreases in the intensity and unpleasantness of a hot stimulus may result from increases in mindfulness, they need not do so, and mindfulness is not defined by intensity and unpleasantness of hot stimuli. Moreover, even if the first condition were met, I show that the thermal pain task is susceptible to many of the biases of self-report measures. Its role as an adequate measure of mindfulness is thereby seriously undermined.

I conclude by questioning the value of attempting to isolate mindfulness as an "active ingredient" and the reductionistic approach to mindfulness research to which it belongs.

References:

Grossman, Paul, and Nicholas T. Van Dam. 2011. "Mindfulness, by Any Other Name…: Trials and Tribulations of sati in Western Psychology and Science." Contemporary Buddhism 12 (1): 219–39.

MacCoon, Donal G., Zac E. Imel, Melissa A. Rosenkranz, Jenna G. Sheftel, Helen Y. Weng, Jude C.
Sullivan, Katherine A. Bonus, Catherine M. Stoney, Tim V. Salomons, Richard J. Davidson, and Antoine Lutz. 2012. "The Validation of an Active Control Intervention for Mindfulness-Based Stress Reduction (MBSR)." Behaviour Research and Therapy 50 (1): 3–12.

3) Chasing Shadows: The Laboratory Study of Hallucination
Michael Dickson

Research into hallucination faces two obstacles. First, hallucination (even restricted to a single underlying condition such as schizophrenia or dementia) is not a single phenomenon.

In addition to occurring in various sense modalities, its qualitative features—prevalence, persistence, content, complexity, emotional valence, etc.—vary across contexts. It is possible that all hallucination involves some single underlying neural process, expressed differently in different contexts, but it is also possible that it does not. Either way, the phenomenon as experienced is radically context-dependent.

Second, hallucination in the wild can be difficult to predict and control. While there have been (occasionally successful) attempts to capture spontaneous hallucination in a laboratory, and there are many phenomenologically-oriented interviews of persons who hallucinate spontaneously, the difficulty of prediction and control makes standard laboratory-based techniques very challenging.

Some researchers propose to tackle the second obstacle by developing reliable procedures for inducing hallucination, procedures involving drugs, sensory deprivation, and (carefully designed) sensory stimulation. The main point of this paper is that these approaches to the second obstacle exacerbate the challenge of contextuality posed by the first. More specifically, while studying brain activity in persons with induced hallucinations could reveal something about the neurological processes involved (for example in the visual or auditory cortex), the clinical relevance of such studies is likely to be extremely limited.

This conclusion runs counter to the ambitions of some otherwise exciting and valuable programs of research. For example, Rogers et al. (2021) advocate for "a new laboratory-based approach to hallucination research" (p. 2), using techniques such as those mentioned above, refined with an aim "to manipulate and predict characteristics of experimentally induced hallucinations … for systematic investigation of hallucination phenomenology using more rigorous behavioural and neural methods" (p. 6). Ultimately, they hope that this research might be "leveraged to help those suffering with hallucinatory conditions" (p. 7). However, as the first obstacle highlights, repeatable laboratory protocols are likely to miss out on phenomenologically and clinically relevant context. The main example will be auditory verbal hallucination (AVH). There are several reasons (beyond the obviously relevant aspect of content) that the experience of AVH is not (easily) reproducible by the kinds of exogenous inducement of hallucination mentioned above. AVH often occur in a context in which the person is also experiencing fundamental doubts about (what is taken to be) 'common sense' (such as that one is surrounded by causally potent objects in the first place), feelings of expectancy or intrinsic meaningfulness that can make small details feel very important, and so on (Fusar-Popli, et al. 2022). It will be argued that these contextual features are part of the experience of hallucination, not 'add-ons' to be understood only after 'hallucination itself' has been studied.

References

Fusar-Popli, P. et al. (2022) "The lived experience of psychosis," World Psychiatry 21:168-188.

Rogers, S., et al. (2021) "Hallucinations on demand: the utility of experimentally induced phenomena in hallucination research," Philosophical Transactions of the Royal Society B 376:20200233.

Towards a Comparative Philosophy of Science

Karim Bschir
University of St. Gallen
karim.bschir@unisg.ch

Simon Lohse
Radboud University Nijmegen
simon.lohse@ru.nl

Henk de Regt
Radboud University Nijmegen
henk.deregt@ru.nl

Federica Russo
Utrecht University
f.russo@uu.nl

Raphael Scholl
University of Geneva
raphael.scholl@unige.ch

Organizers:
Karim Bschir, University of St. Gallen, Switzerland,
Simon Lohse, Radboud University Nijmegen, Netherlands

In recent decades, philosophy of science has diversified considerably. Whereas the first half of the 20th century was characterized by a strong focus on physics and later also on biology, the past decades have witnessed the rise of various philosophies of the special sciences (Allhoff 2010; Lohse & Reydon 2017). As a consequence, we are increasingly dealing with a fragmented field of philosophies of the special sciences. This diversity not only carries the threat of fragmentation but also offers a rich reservoir of opportunities for comparison between the different (philosophies of) the special sciences.

Although there already exist several contributions in the current literature that provide philosophical analyses across disciplines (e.g. Hoyningen-Huene 2013; du Crest et al. 2023), there is as of yet no integrative debate. The proposed symposium wishes to change this by exploring a new research program that aims at a systematic comparison of scientific disciplines and practices from a philosophical perspective. This approach will generate important insights and promote progress in several dimensions, in particular:

Comparison of epistemic practices: A comparative approach provides us with a rich and nuanced picture of epistemic practices (such as explanation, prediction) in different disciplines. It enables the identification of relevant similarities and differences of these practices and the factors responsible for these. This can also provide useful insights for classical questions of general philosophy of science, for example with regards to scientific change, on which the comparative approach allows a fresh perspective.

Comparative conceptual analysis: A comparative approach enables a better understanding of scientific core concepts such as “objectivity” or “reliability” by highlighting contrasts and analogies across fields (e.g.: To what extent and why are concepts like “reliability” used differently in scientific practice?).

Comparative meta-philosophy of science: A comparative approach can create insights into philosophical practices. This may facilitate a better understanding of differences in doing philosophy of science by exploring why philosophy of science is sometimes considered a reflective meta-discipline and sometimes on a continuum with science.
In order to facilitate a discussion on comparative philosophy of science, the proposed symposium brings together philosophers with different specializations: After an introductory talk on the idea of comparative philosophy of science (Bschir & Lohse), de Regt analyzes differences and similarities with respect to understanding and intelligibility standards in different fields; Russo investigates practices of data generation and analysis across disciplinary contexts suggesting that these involve a spectrum of integrated qualitative and quantitative practices; and Scholl explores Crombie's notion of "styles of scientific thinking" suggesting that it can be helpful in highlighting parallels between disciplines, but still allowing appreciation of their particularities.
The proposed symposium can serve as a starting point for an innovative research avenue within philosophy science with a strong focus on the study of scientific practice. At the same time, it will make an important integrative contribution to the philosophies of the special sciences and shed new light on core question of general philosophy of science about the nature of science and the special epistemic status of scientific knowledge.

Abstracts:
1. Karim Bschir, University of St. Gallen & Simon Lohse, Radboud University Nijmegen:
Introduction – From Fragmentation to Comparative Philosophy of Science

This introductory talk presents the motivation and the groundwork for of a comparative philosophy of science. We start by giving a historical overview of the differentiation of philosophy of science into various sub-disciplines. We will briefly highlight how institutional developments after World War II and the so-called historical turn in the first decades of the second half of the 20th century created an environment in which received notions of science's development and its rationality were increasingly challenged. Descriptive sociological and historical approaches to scientific knowledge production, such as HPS or SSK, moved the focus away from classical general debates on explanation, confirmation, falsification, induction etc. Philosophers of science became increasingly interested in local aspects and problems within specific scientific fields or disciplines. In combination with an accelerating disciplinary differentiation within science itself, these developments led to a fragmentation of the field which itself led to the emergence of different philosophies of the special sciences with different objectives, projects and debates. If we were to identify a grand arch within the history of philosophy of science in the 20th century, it can be found in the shift away from search for universal methodological standards and debates about the possibility of an Einheitswissenschaft towards a highly differentiated view of science that highlights heterogeneity and fragmentation rather than unity and similarities.
This pluralization within philosophy of science has resulted in a field with a high-resolution that is better able to do justice to the heterogeneity of its subject areas. At the same time, however, the field is threatened with a loss of cohesion. To address this challenge, we propose to develop a systematic program in comparative philosophy of science.
The second part of the talk will sketch the broad lines of such a program. After briefly discussing historical predecessors of our proposal, we will describe various goals of a comparative philosophy of science (e.g. contributing to a general theory of science and "myth busting"). We will then distinguish between five dimensions of comparison, namely epistemic aims, scientific core concepts (see talk #2), epistemic practices (see talks #2 and #3), styles of thinking (see talk #4), and meta-philosophical comparison – some of these

dimensions will be explored by the other talks of the symposium as indicated. Finally, we will discuss methodological and institutional challenges for the outlined program and highlight some implications for teaching philosophy of science.

2. Henk de Regt, Radboud University Nijmegen: Scientific understanding across disciplines: in varietate concordia?

Understanding is a generally acknowledged aim of science, but study of scientific practice teaches us that strategies for producing understanding, and criteria for when it is achieved, differ widely across scientific disciplines. We present a comparative analysis of the ways in which understanding is obtained in different disciplines: what is shared and which differences can be found? Given the diversification of modes of explanation and criteria for understanding, can we still identify a common core? Starting point is the contextual theory of scientific understanding developed in De Regt (2017), which allows for contextual variation in criteria for understanding. While this work focuses mainly on historical variation, our paper explores disciplinary variation in by comparing scientific disciplines ranging from physics to biology and psychology. We evaluate our findings in light of the contextual theory of scientific understanding.
Scientific understanding of phenomena is typically generated by constructing explanations. Therefore, we first analyze the modes of explanation employed in the different fields. Across all disciplines, some form of causal explanation is ubiquitous. This aligns with the current consensus in philosophy of science, but our comparative perspective shows how 'causal' is instantiated in multiple, context-sensitive guises. Mechanistic explanation appears as a salient subclass of causal explanation, especially in higher-level sciences such as biology, neuroscience, and psychology. Overall, higher-level sciences dealing with complex systems are more pluralistic in their explanatory strategies than physics and chemistry. By contrast, covering-law explanation appears marginal in contemporary practice and is largely restricted to theoretical physics.
Second, we investigate the role of theories and models in acquiring understanding. The contextual theory assumes that explanations are typically model-based, where model construction requires intelligible theories. Intelligibility standards are context dependent and may exhibit disciplinary variation. Our comparative analysis supports this general picture but reveals substantial variation in the nature and function of theory–model relations. Thus, in physics fundamental, mathematical theories are essential for understanding. In chemistry and the geosciences, such theories still figure in the background, but understanding is usually acquired via qualitative domain-specific theories and associated models. In the higher-level sciences, by contrast, modeling frequently takes precedence over theorizing, and models tend to be less tightly linked to fundamental theories.
Finally, we compare intelligibility standards across disciplines. Despite variation, several common preferences emerge: visualizability, qualitative interpretability, and the use of analogies are widely valued, reflecting a general preference for causal and mechanistic accounts that afford intuitive insight. We conclude that, although explanatory strategies and criteria for understanding and intelligibility differ significantly across disciplines, there is a robust common core centered on modeling and intelligibility. A suitably generalized contextual theory would thus be able to accommodate both the unity and the diversity of scientific understanding. We will briefly discuss a generalization of the contextual theory, proposed by Kostic and Khalifa (2025), that accommodates this disciplinary variation.

3. Federica Russo, Utrecht University: Data generation and analysis across the QUALitative-QUANtitative spectrum. A reassessment and a proposal

Data generation and analysis are common practices in the sciences, whether natural, life, social, or biomedical ones. This makes questions of data generation an excellent candidate for exploring comparative approaches in philosophy of science. Often, questions of data generation are associated with the alleged superior status of fields that adopt quantitative methods for data generation and analysis that would confer objectivity to the results. Instead, areas that employ qualitative or ethnographic methods, such as the social sciences, hold an irreducible element of subjectivity and idiosyncrasy, because the researcher is too close to the subjects of study. In the spirit of a comparative approach, one should dismantle this hierarchy of methods, already by making the simple observation that some social science areas do use quantitative approaches. I will not, however, tackle this problem from the perspective of objectivity, as it is usually done, but from the (alleged) distinction between qualitative and quantitative methods and data. I do so by running comparison not directly among natural and social contexts but moving from a foundational debate in social science methodology: the long-standing quarrel about qualitative (QUAL) and quantitative (QUAN) methods, and the contribution given in Mixed Methods Research.
In social science methodology, it has long been debated about the priority, validity, and preferability of quantitative over qualitative methods. It is a debate that traces back to the original development of quantitative methods in social science, in part inspired by the use of quantitative methods in natural science. The debate is still ongoing, with significant contributions by theorizers and practitioners of Mixed Methods Research. While a large part of the debate has been framed in terms of a choice between quantitative and qualitative methods, in this contribution I shift focus from methods (for data collection and analysis) to practices for data generation and analysis. In the spirit of a comparative approach in philosophy of science, I mobilize the conceptualization of data as semantic information as is developed in the philosophy of information, and explain the sense in which generating quantitative or qualitative data involves a semanticisation that happens through a spectrum of integrated – quantitative and qualitative – practices.
Through a comparison of these practices, I conclude that while a clear distinction between qualitative and quantitative data and methods is in order, generation and analysis of data in fact involve a spectrum of integrated qualitative and quantitative practices. I identify at least three reasons why this philosophical result is of relevance: (i) shedding light on scientific method, also for teaching purposes, (ii) reconsidering the politics of research funding, and (iii) for its policy implications.
More generally, and again in the spirit of a comparative approach in philosophy of science, this result should help us rethink the alleged superiority of quantitative practices such as measurement in natural science experiments or causal inference in randomized controlled trials.

4. Raphael Scholl, University of Geneva: Revisiting Styles of Scientific Thinking

The historian of science A. C. Crombie proposed that we can recognize a number of distinct "styles of scientific thinking". Here I will argue that this notion should be revisited and developed by philosophers of science. It is not just that the notion of styles of thinking is descriptively adequate in so far as it captures strategies of epistemic problem solving that can be found across time and across disciplines. It is also that the styles of thinking promise to bridge the disciplinary silos in which much philosophy of science is currently conducted. By

finding and understanding styles of scientific thinking over time and across disciplines, we can do philosophy of science in a comparative mode.

Crombie's notion of styles of thinking was previously received within the philosophy of science by Ian Hacking, who used his own term of "styles of reasoning". While Hacking's treatment of the notion is stimulating, I will argue that it has loaded the styles with problematic philosophical baggage. Hacking took styles of reasoning to create new candidates for truth and falsehood, arguing that questions cannot even be asked until there is a type of reasoning by which they can be answered. Hacking also took styles of reasoning to be in a certain sense insular. He suggested that styles of reasoning are "self-authenticating" and cannot be justified by facts or arguments external to the style itself. These are strong theses worthy of discussion, but Crombie's original notion of styles of scientific thinking is leaner and more narrowly descriptive. It is that descriptive notion of styles on which I want to build.

My argument will proceed by example. I will focus on three case studies from different disciplines in which two styles of scientific thinking come into tension: in these instances, they are the "hypothetical modeling" style and the "experimental" style. In the early 20th century, the younger members of T. H. Morgan's "fly room" believed that they had sufficient evidence for the chromosomal theory of inheritance from the model-mediated interpretation of crossing experiments, while Morgan himself required more direct evidence that chromosomes cause the expression of traits. In mid-20th-century population biology, mathematical modelers and field experimenters disagreed about how best to study interspecific competition. Finally, in the later 20th century, the traditional theoretical practices of economists were challenged by econometrics. I will show, first, that far from creating new candidates for truth or falsehood, the different styles of scientific thinking were alternative attempts to answer the same questions. Second, the reasoning involved in the styles is not insular, but connected. The different styles can support each other mutually. We will see that comparing styles of scientific thinking across disciplines highlights parallels but still allows us to appreciate the rich particulars of each discipline. While each discipline faced similar methodological tensions, their solutions were instructively different. Thus, a focus on styles of thinking enables a fruitful new approach of comparative philosophy of science.

References:

Allhoff, F. (Ed.) (2010). Philosophies of the Sciences. Wiley-Blackwell.

Bschir, K., Lohse, S., & Chang, H. (2019). Introduction: Systematicity, the Nature of Science? Synthese, 196(3), 761–773. https://doi.org/10.1007/s11229-018-1685-z

De Regt, Henk W. 2017. Understanding scientific understanding. New York: Oxford University Press.

du Crest, A., Valkovi*ƒ*á, M., Ariew, A., Desmond, H., Huneman, P. & Reydon, T. A. C. (Eds.) (2023). Evolutionary Thinking Across Disciplines: Problems and Perspectives in Generalized Darwinism Springer International Publishing.

Hoyningen-Huene, P. (2013). Systematicity: The Nature of Science. Oxford University Press.

Kosti*ƒ*á, Daniel, and Kareem Khalifa. 2025. Does functional connectivity explain? Synthese 206: article 207. https://doi.org/10.1007/s11229-025-05296-w

Lohse, S., & Reydon, T. (Eds.) (2017). Grundriss Wissenschaftsphilosophie: Die Philosophien der Einzelwissenschaften. Meiner.

Equitable Open Science: A Practice-Based Philosophical Approach

Emma Cavazzoni
Technical University of Munich
emma.cavazzoni@tum.de

Nathanael Sheehan
University of Exeter/ Technical University of Munich
nathanaelsheehan@gmail.com

Rachel A. Ankeny
WUR/Adelaide
rachel.ankeny@wur.nl

Cassandra Sedran-Price
University of Sydney
cassandra.sedran-price@sydney.edu.au

Sabina Leonelli
Technical University of Munich
sabina.leonelli@tum.de

There are significant epistemic risks in implementing open science across widely different research settings, chief among them the marginalisation of contributions from low-resourced environments. This raises questions about the relation between open and good science. To address these concerns, this session combines philosophical analysis of the epistemic significance of research environments with ethnographic investigation on how researchers in different environments enact and conceptualise OS, as well as collaboration with scientists and policymakers towards better understanding the conditions for scientific work. Building on empirical fieldwork carried out within the ERC Project "A Philosophy of Open Science for Diverse Research Environments" (2021-2027), session participants develop a philosophy of openness that emphasises the diversity of research environments around the world and articulates the conditions under which open science can leverage such diversity to promote good research practice. Emma Cavazzoni opens the session by reflecting on how the push for shareable data and technologies creates tensions in pest–plant interaction research in Northern Italy; Nathanael Sheehan details conditions of closed-openness with global health data, as managed by research facilities in Cambridge (UK) and Salvador (Brazil). Rachel Ankeny and Cas Sedran-Price then examine the use of Indigenous data in ecological and environmental modelling in Australia and New Zealand, highlighting the concerns this creates for open science; and Sabina Leonelli closes the session discussing ways to enhance epistemic and structural justice within agronomic research carried out in Ghana and in Greece, to support improved research outcomes and recognition for contributions coming from low-resources research environments.

Papers:
Tensions in Preparing and Maintaining the Feld for Data-Intensive Plant Science Research: When Standardized AI Meets Local Complexity (Emma Cavazzoni)

In this presentation, I investigate the tension that the push for shareable data and technologies creates in pest–plant interaction research, and its implications for research practice. I focus

specifically on how it shapes the crucial work of preparing and maintaining the field and technologies for data-intensive AI research, and on the dynamic of mutual adjustment involved.

Conducting pest-plant interaction research in the field means that the objects of study are not isolated entities but exist in interaction with other organisms (Latour, 1983), embedded within a natural environment that is difficult to control (Knorr-Cetina, 1992). This situation requires researchers and their technologies, originally designed for precision work in predictable lab environments, to adapt to nature's conditions: they must be modified to withstand the unexpected and harsh circumstances of the specific field under investigation (Kohler, 2002). Conversely, the field itself has to be reshaped to accommodate instruments and research protocols. Yet not all aspects of it can be altered – soil fertility, soil composition, and tree vigor, for instance, frequently remain beyond control. In such cases, the cycle reverses again, forcing researchers to further adapt their tools to these intractable environmental conditions. The result is a continuous interplay in which field and technologies reshape one another in the pursuit of a workable alignment.

This situation, however, generates a tension in contemporary pest-plant interaction research, where data-intensive technologies like software and hardware with integrated AI have become increasingly central (Sheard et al., 2024; Sánchez Herrera et al., 2024). These AI systems are designed to be standardized for sharing and reuse across diverse contexts, whilst also being highly complex, with many stages of their computation remaining opaque to human understanding (von Eschenbach, 2021). As a consequence, scrutinizing and tailoring them to suit specific research needs becomes difficult: researchers cannot fully trace their internal operations and thus struggle to adapt them for particular aims. This produces a clear tension: fieldwork requires that technologies be adapted to local conditions, yet the data-intensive AI systems increasingly adopted are themselves resistant to such adaptation.

In this talk, I investigate this tension and consider the dynamics that arise when preparing and maintaining both field and technologies involved in data-intensive AI research on pest–plant interactions. Following Wylie (2021), I focus on preparation and maintenance because, despite often remaining invisible (Shapin, 1989), adaptation lies at their core, making the tension especially salient and deeply influential in scientific knowledge production. My reflections draw on six months of ethnographic work and collaborations with Haly.Id, a research project based in Northern Italy which develops large datasets and AI technologies for a targeted monitoring of the presence in crop fields of the highly invasive pest Halyomorpha Halys (Ferrari et al., 2023; Giannetti et al., 2024). The discussion is organized around three key dimensions that strongly influence how fields and technologies are adapted to one another and that highlight the tensions arising from this process: social relations, environmental conditions, and methodological choices.

Bibliography

von Eschenbach, W.J. (2021) "Transparency and the Black Box Problem: Why We Do Not Trust Ai," Philosophy and Technology, 34(4), pp. 1607–1622. Available at: https://doi.org/10.1007/s13347-021-00477-0.

Ferrari, V. et al. (2023) "Evaluation of the potential of Near Infrared Hyperspectral Imaging for monitoring the invasive brown marmorated stink bug," Chemometrics and Intelligent Laboratory Systems, 234, p. 104751.

Giannetti, D. et al. (2024) "First use of unmanned aerial vehicles to monitor Halyomorpha halys and recognize it using artificial intelligence," Pest Management Science [Preprint]. Available at: https://doi.org/10.1002/ps.8115.

Knorr-Cetina, K. (1992) “The Couch, the Cathedral, and the Laboratory: On the Relationship Between Experiment and Laboratory in Science,” in A. Pickering (ed.) Science as Practice and Culture. Chicago, IL: University of Chicago Press. Available at: https://philarchive.org/rec/CETTCT (Accessed: January 29, 2024).
Kohler, R.E. (2002) Landscapes and Labscapes: Exploring the Lab-Field Border in Biology. Chicago, IL: University of Chicago Press. Available at: https://press.uchicago.edu/ucp/books/book/chicago/L/bo3640043.html (Accessed: January 18, 2024).
Latour, B. (1983) Visualisation and Cognition: Drawing Things Together. Boston: Harvard University Press. Available at: https://doi.org/10.1002/9780470979587.ch9.
Sánchez Herrera, M. et al. (2024) “Systematic challenges and opportunities in insect monitoring: a Global South perspective,” Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 379(1904), p. 20230102. Available at: https://doi.org/10.1098/rstb.2023.0102.
Shapin, S. (1989) “The invisible technician,” American scientist, 77(6), pp. 554–563.
Sheard, J.K. et al. (2024) “Emerging technologies in citizen science and potential for insect monitoring,” Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 379(1904), p. 20230106. Available at: https://doi.org/10.1098/rstb.2023.0106.
Wylie, C.D. (2021) Preparing Dinosaurs: The Work behind the Scenes. Cambridge, Massachusetts: The MIT Press. Available at: https://doi.org/10.7551/mitpress/12643.001.0001.

Closed-Openness in Practice: A Comparative Analysis of Data Sharing Across Diverse Research Environments (Nathanael Sheehan)

Knowledge-sharing norms (KSNs) refer to the agreements, expectations, and accountability structures that govern how communities and relationships ought to share and withhold knowledge. As Grasswick notes, KSNs specify not only when sharing is appropriate, but also when it is epistemically or morally improper, and how communities hold one another responsible for these decisions (Grasswick, 2011). Within Open Science (OS), debates around KSNs tend to take on a universalistic tenor, as though transparency, sharing, and access inherently possess epistemic value irrespective of context. Yet policy formulations and international OS initiatives are typically guided by the principle of making research outputs “as open as possible, and as closed as necessary” (European Research Executive Agency, 2025).

Although the terms “open”, “closed”, “possible” and “necessary” are all legally ambiguous (Méndez and Cueva, 2021), this KSN expresses a broadly alludes to the recommendation that scientific outputs such as data and software ought to be made accessible via the relevant open repositories for a given output. However, the outputs may be kept closed given plausible reasons - such as risk of commercial exploitation or obligations in grant-agreements - where making the output open may run counter to a researcher’s interests (European Research Executive Agency., 2025). While there has been significant work in detailing the process in which researchers can make their outputs as open as possible vis-à-vis the development of research data management principles such as FAIR, there has been significantly little philosophical work detailing the conditions for acting as closed as necessary in OS across diverse research environments.

This talk addresses that gap by advancing a comparative, practice-based account of what I term closed openness: the situated processes through which partial, conditional, or strategically modulated openness is enacted. Drawing on ethnographic fieldwork and semi-

structured interviews conducted in two contrasting health data infrastructures: the European Bioinformatics Institute (EBI) in Cambridge, UK, and the Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS) in Salvador, Brazil; I argue that the mechanisms of "closed openness" arise from, and are justified by, the material, institutional, and sociopolitical conditions under which researchers work. While both institutions share a commitment to managing large-scale health data for public benefit, their capacities to enact "closed-openness" diverge sharply. At the EBI, closed-openness practices such as data clearance houses and predefined embargo periods are designed primarily to secure trust with data-submitting researchers, mitigating concerns about "scooping" and ensuring proper attribution. At CIDACS, by contrast, these mechanisms operate largely as safeguards against external control, particularly from powerful international foundations and research centres, thereby protecting the institution's autonomy and stewardship over sensitive national data. Taken together, these cases support other sociological and philosophical observations that obstacles to equitable participation in OS cannot be reduced to questions of digital connectivity or the formal provision of access (Leonelli, 2023; Bezuidenhout, 2025). Rather, they are embedded instead in relational, organisational, and epistemic arrangements that shape who can meaningfully engage with data, for what purposes, and with what forms of recognition and authority. The concept of "closed openness" thus draws attention to the situated negotiations through which not completely open forms of openness are enacted and to the diverse conditions that make openness both desirable and, at times, necessarily constrained.

References

Bezuidenhout, L. (2025). Critically Unpacking the Concept of Equity for Open Science. Social Epistemology, 0(0), 1–18. https://doi.org/10.1080/02691728.2025.2475460
De la Cueva, J., & Méndez, E. (2022). Open science and intellectual property rights. How can they better interact.
Grasswick, H. E. (2011). Liberatory Epistemology and the Sharing of Knowledge: Querying the Norms. In H. E. Grasswick (Ed.), Feminist Epistemology and Philosophy of Science (pp. 241–262). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6835-5_12
Leonelli, S. (2023). Philosophy of Open Science. Elements in the Philosophy of Science. https://doi.org/10.1017/9781009416368
Open science. European Research Executive Agency. (n.d.). https://rea.ec.europa.eu/open-science_en

Strategies for Open Science with Indigenous Data in the Life Sciences in Australia
(Rachel A. Ankeny and Cassandra Sedran-Price)

There is an increasing focus on open science practices around the globe, including in locales such as Australia where there also is considerable attention to developing appropriate practices for upholding Indigenous Peoples' rights including data sovereignty. This paper explores proposed protocols associated with use of Indigenous data in the Australian life sciences, and articulates points of tension in these protocols which provide best practices with current-day practices and these fields. We propose intervention points for improving these practices, focusing on those points in the data lifecycle which contribute most significantly to epistemic injustice, drawing on the recent CARE Data Maturity Model. We emphasize that open science can be actively cultured in this domain when deep attention is paid to these intervention points and to more contextualized definitions of open science.

Towards Equitable Open Science: Addressing Structural Injustice (Sabina Leonelli)

This paper argues for diversity and justice as key philosophical tenets for equitable as well as robust open science practices. Building on extensive ethnographic research on everyday challenges within life science sites in Italy, Ghana and Greece (Leonelli 2024, Leonelli and Trappes 2025), I first focus on three forms of structural injustice that underpin disparities in the extent to which researchers have access to resources to conduct their work, and the opportunity to use such resources to further their scientific goals. These are: (1) resource injustice linked to inequity in resourcing; (2) focus injustice relating to misalignments between resourcing and goals; and (3) labor injustice caused by misalignments between scientific goals and labor conditions. I then argue that these types of structural injustice have significant epistemic implications, resulting in: shifts in research content; the exclusion of substantive portions of scientific community from contributing to knowledge production to the best of their abilities; the exclusion of specific methods and data sources from global scientific discourse, on grounds other than their scientific quality and relevance; an overarching loss of quality and reliability in processes and outcomes of inquiry; and the inability to implement openness and related forms of constructive scrutiny. I conclude that addressing such injustice requires strategies to promote diversity within the scientific workforce, geographical location of research centres, and types of goals and methods supported by open science infrastructures (Leonelli under review 1). I exemplify such strategies under the label of 'humane openness' (Leonelli under review 2) focused on engagement as the condition of possibility for the sharing of research insights and materials.

References

Leonelli, S and Trappes, R. (2025) Research in the Multiplex: Navigating Tensions and Opportunities in Transdisciplinary Environments. European Journal for the Philosophy of Science 15:62 https://doi.org/10.1007/s13194-025-00684-y
Leonelli, S. (2024) Globalising plant knowledge beyond bioprospecting? History of Anthropology Review 48https://histanthro.org/notes/globalizing-plant-knowledge/
Leonelli, S. (manuscript under review) Not all research environments are created equal: Structural epistemic injustice in low-resourced scientific settings.
Leonelli, S. (manuscript under review) Contrasting visions of inquiry for an Open Society: Henri Bergson and Karl Popper between humane and rational openness.

Methods of Integrated History and Philosophy of Science

Hasok Chang
University of Cambridge
hc372@cam.ac.uk

Mauricio Suarez
Complutense University of Madrid
msuareza@ucm.es

Hauke Riesch
Brunel University
Hauke.Riesch@brunel.ac.uk

Thomas Bonnin
Politecnico di Milano
thomas.bonnin@polimi.it

Jan Baedke
Ruhr University Bochum
Jan.Baedke@ruhr-uni-bochum.de

Rachel Ankeny
Wageningen University
rachel.ankeny@wur.nl

One important strand of the philosophy of science in practice (PSP) is philosophical engagement with the history of science with a focus on the development of scientific practices. However, there has been surprisingly little interaction between PSP and the "integrated history and philosophy of science (HPS)" movements. This symposium seeks to encourage fresh methodological thinking in integrated HPS and connect it to PSP. This proposed symposium consists of 4 individual presentations, proposing a diverse array of methodological developments in integrated HPS. Rachel Ankeny will provide an overall commentary on these papers]
Mauricio Suárez tackles the familiar methodological quandary of historical case-studies in the philosophy of science, and proposes the re-orientation of such historical work as the study of historical episodes, intended for a hermeneutical illumination of philosophical accounts of scientific practices. These points will be shown through a discussion of the modelling attitude in 19th-century science. Hauke Riesch argues for a more serious injection of sociological perspectives into integrated HPS. Taking up the case of Max Weber's study of prophets and charismatic authority, he will show how these sociological concepts can help us understand some scientific icons such as Albert Einstein and Stephen Hawking. Thomas Bonnin takes up the idea of HPS as complementary science, and makes a novel application of it to the field of architecture. Taking up the example of hospital design, he shows how contemporary designers can learn from studying how early hospital designers conceptualized, operationalized, and measured the health effects of architectural decisions. Jan Baedke proposes to bring methods of digital humanities to help expand the reach of integrated HPS. Pulling together the data embedded in hundreds of herbaria worldwide, he identifies local collectors and their networks that have been erased from conventional historiography. This

creates a novel empirical basis from which to address questions of epistemic injustice and data biases in a more informed way.
This symposium is coordinated by the Joint (history/philosophy) Commission of the International Union of History and Philosophy of Science and Technology (IUHPST), and it will be chaired by Agnes Bolinska, the Chair of the Joint Commission.

Abstracts of Individual Papers

Case Studies, Historical Episodes, and Integrated History and Philosophy of Science
Mauricio Suárez

In this talk I take up the issue of the value of 'case studies' within integrated history and philosophy of science. While the methodology of case studies is a characteristic staple of the philosophy of science in practice, and a much-employed tool in the ethnographic and sociological studies of scientific practice that are typical of Science and Technology Studies units, it is not so clearly appropriate for the integration of specifically historical work within philosophy of science. I argue that in fact at least some parts of integrated history and philosophy of science would benefit from a shift away from the notion of 'case study' towards the different concept of a 'historical episode'. The difference between these two notions in this context is essentially down to the role they play vis a vis philosophical views, theses, or theories. 'Case studies' were taken up by anglophone philosophers of science in the 1970's and 1980's to allow them to decide between competing philosophical accounts of science, scientific theory, or scientific method. They were thus meant to provide philosophers with the equivalent of evidence for or against a scientific theory. In other words, the notion of 'case study' in traditional philosophy of science reflects an underlying commitment to treat historical data as providing evidence for or against philosophical theory, in an analogy to how data relates to theory in science itself.
A main reason to move away from 'case study' talk is thus intimately connected with an abandonment of 'scientism' with respect to philosophy of science itself. Philosophical thesis, theories or views are not themselves scientific but are instead best understood as hermeneutical interpretations of science and its practice. The rather common assumption that 'case studies' may be ad hoc, post hoc, ante hoc, or propter hoc with respect to philosophical theses or views, and that this prevents them from doing their work effectively, is thus of a piece with the thought that philosophy relates to history in the same way that scientific theory relates to data.
'Historical episode' does not carry any of these connotations. A 'historical episode' is a genealogical narrative of events leading up to present science that provides a certain sort of account of how this present came to be. Such 'presentist' history of science need not be whiggish – there is no reason why it should consist in anything like an account of the inevitable triumph of present-day science (Chang, 2022). It is rather perfectly compatible with the thought the science is inevitably a complex contingent intellectual development that can go down many different paths at any given juncture in time. The central role of 'historical episodes' is to illuminate philosophical accounts of scientific practice, by filling in the details and providing striking examples. As an illustration I offer the historical episode of the emergence of what I call the modelling attitude in 19th century science (Suárez, 2024, 2025).

References

Chang, H. (2021), "Presentist History for Philosophy of Science", Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie 52 (1):97-114

Shapin, S. (1996), The Scientific Revolution, University of Chicago Press.
Suárez, M. (2024) "The Emergence of the Modelling Attitude", in T. Knuuttila, N. Carrillo and R. Koskinen, Eds., Routledge Handbook of Philosophy of Scientific Modelling, Routledge: London, pp. 11-25

Complementary Architecture: Contemporary Lessons from the First Modern Hospitals
Thomas Bonnin

Contemporary debates about hospital design are dominated by pressing concerns over obsolescence and adaptability. Hospitals, it is assumed, must keep pace with the rapid and unpredictable evolution of scientific, technological, and political conditions. As a result, current strategies for planning complex healthcare environments emphasize continuous digital monitoring, data-driven decision-making, and built-in flexibility and modularity. This agenda–often framed as an advancement of Evidence-Based Design (EBD)–leaves little room for earlier architectural knowledge, except for occasional recognition of isolated pioneers who appear to prefigure present-day approaches.
This presentation challenges that teleological narrative by proposing an integrated historical and philosophical framework through which past hospital architecture can meaningfully contribute to present design practices. We do so by extending Hasok Chang's concept of complementary science into the architectural domain. Chang argues for the revival of prematurely abandoned scientific practices for pluralistic motives: to expand the repertoire of available practices, stimulate and improve existing methodologies, and question implicit or unexamined conceptual and methodological assumptions. We propose that this framework can similarly help recover architectural approaches that remain epistemically valuable, even if they are no longer technologically current.
EBD seeks to ground design decisions in rigorous evidence, treating the built environment as a therapeutic agent capable of influencing patient and staff outcomes. Its proponents emphasize the objective documentation of spatial effects emphasizing methods–such as controlled experiments–inspired by Evidence-Based Medicine. However, when viewed through the lens of complementary architecture, this project appears far less novel. Late-19th- and early-20th-century hospital design – often associated to the figure of Florence Nightingale – already involved reflexive attempts to measure and theorize the impact of built spaces on health.
We argue that 21st-century hospital designers need not remain caught in a perpetual race for methodological or technological innovations to counter obsolescence. Instead, they can benefit from examining how their predecessors understood and negotiated analogous challenges. Our analysis combines historical scholarship with epistemological analysis to identify how early hospital designers conceptualized, operationalized, and measured the health effects of architectural decisions. At a time when design research is increasingly driven by vast streams of sensor-derived data, the reflexivity embedded in past practices offers valuable guidance for articulating a robust epistemological framework for contemporary hospital design. More broadly, this case demonstrates both the promise of complementary science for applied sciences and the adjustments required to adapt it to architectural inquiry.

Prophecy and Science
Hauke Riesch

Contributions from the sociology of science have shown how scientific knowledge is shaped by the social interactions and rhetorical strategies employed by scientists in their

communications to each other and the wider public. This social dimension of knowledge-making has long been acknowledged in the HPS canon, even though explicitly sociological work has so far been left out of the discipline (and its acronym). This talk will argue through an example of Weber's characterisation of the prophet and Festinger's account of failed prophecy how classical approaches to sociology and social psychology can contribute to HPS by helping us theorise how the presentation of the scientific persona contributes to how people react to the science and explore the similarities to how science and prophecy account for failed predictions.

I will first add to Max Weber's (1978) characterisation of prophecy working mainly through the establishment of charismatic authority by elaborating on his description of the prophetic persona, which includes humbleness, laicity and adversity. The link between scientific rhetoric and persona with the prophetic will be established with comparison to postulated norms of scientific behaviour such as those of Robert Merton (1973): the norm that science should be disinterested for example, maps on to the humbleness of the prophet, as, if the prophecy is to be persuasive, the prophet should not stand to gain from their message (Riesch 2024).

Both science and prophecy predict and inevitably confront the success or failure of their predictions. In Festinger et al.'s (1956) seminal work on what happens to an apocalyptic belief system in case of a disconfirmed prophecy, the believers changed some auxiliary hypotheses in order to save their beliefs. This approach to disconfirmation will be compared to classical theories in the philosophy of science, in particular the scientific research programmes elaborated by Lakatos (1978). Developing my argument in Riesch (2021), I will then theorise a wider account of how rationality and prediction in such widely different fields of science and apocalyptic prophecy can be understood as a wider, intuitive, social epistemology we engage in that is relevant to everyday life and experiences, incorporating insights from Weber's characterisation of the prophetic persona.

These rhetorical and epistemological characteristics that are shared by science and prophecy will be explored through the running example of the Millerite prophecy, one of the case studies used by Festinger, a small selection of influential and charismatic scientists: Albert Einstein's and Stephen Hawking's public image I will argue depended on charismatic prophetic characteristics such as adversity, humbleness and laicity.

References

Festinger, L., Riecken, H. & Schachter, S., (1956). When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world. New York: Harper & Row

Lakatos, I., (1978). Philosophical papers, volume I: The methodology of scientific research programmes. Cambridge: Cambridge University Press.

Merton, R. K., (1973). The sociology of science: Theoretical and empirical investigations. Chicago: University of Chicago Press.

Riesch, H., (2021). Apocalyptic Narratives: Science, Risk and Prophecy. London: Routledge.

Riesch, H. (2024). "Science Communication and Apocalyptic Prophecy", in Kamwendo, T (ed.), Science and Religion: Approaches from Science and Technology Studies. Berlin: Springer.

Weber, M. (1978). Economy and Society. Berkeley and Los Angeles: University of California Press.

Local Plant Knowers and Epistemic Injustice: Linking the History of Botany with Biodiversity Conservation through Virtual Herbaria
Jan Baedke

Herbaria hold nearly 400 million plant specimens worldwide. They are vital repositories of botanical, biodiversity, environmental, and traditional ecological knowledge. After decades of underfunding, recent digitization efforts have begun to transform herbarium research by building "virtual herbaria" that, by drawing on 400 years of plant evolution data, enable new climate change–related biodiversity and conservation research. In this talk, I conceptualize herbarium specimens as 'knowledge hubs' that entangle biological, environmental, historical, and socio-cultural insights. This sets the theoretical foundation for a broader integrated HPS methodology that aims to interlink historical, epistemological, as well as data- and value-related issues in herbarium research.

In this paper, first, I propose a digital pipeline for herbarium-based historical and philosophical research. This includes: (1) systematic extraction, cleaning, and analysis of specimen metadata from biodiversity databanks like GBIF and iDigBio to identify so-far hidden local plant collectors and their medical and agricultural knowledge; (2) visualization of collector networks by using spatiotemporal herbarium data; and (3) identification of geographical, seasonal, and taxonomic patterns and shifts in local plant knowledge production. Second, I show how these 'distant reading' techniques can be combined with qualitative sources, like field notes and correspondences, to develop more detailed biographical narratives about local actors, and to address philosophical questions about epistemic injustice and research biases.

To demonstrate this method, third, I examine key botanical expeditions across the Yucatan peninsula (Mexico, Guatemala, and British Honduras/Belize) since the early 20th century. Particular attention is given to local mega-collectors – such as Percy H. Gentle and Elias Contreras – whose contributions were crucial yet have been largely erased from historical narratives. Drawing on expedition reports, archival materials from the US and Central America, and especially the digital traces left by their collected plants in biodiversity platforms, I reconstruct their role within specific botanical and archaeological expeditions in the Maya region. I then uncover patterns of epistemic injustice that shaped locals' collection practices (what, where, and when they collected) and the dissemination of their botanical knowledge. I argue that these injustices led to serious information gaps, epistemic blind spots, and geographical, taxonomic, and seasonal biases in contemporary biodiversity datasets – affecting their reliability for addressing conservation challenges and predicting plant evolution today.

Fourth, I discuss opportunities and challenges of this methodology. While this approach enables novel reconstructions of overlooked local collector networks and opens new avenues for a computational history and philosophy of botany from the perspective of local knowers, significant obstacles persist: incomplete or biased datasets, inconsistent metadata, and the methodological difficulty of aligning quantitative digital data with qualitative historical sources and epistemological questions. These problems demand a theoretically grounded methodology that bridges digital tools with epistemic critique and historical awareness. Against this background, I reflect on how conceptualizing herbarium specimens as knowledge hubs can contribute to a more inclusive and socially just botany for the 21st century. This approach brings together, archival sources, digital humanities tools, and philosophical insight to reorient big data botany toward historical responsibility and the recognition of local and Indigenous plant knowers.

Exiled Empiricists: Academic Practices and the Reception of Logical Empiricism

Claudia Cristalli
Tilburg University
c.cristalli@tilburguniversity.edu

Nina IJdens
Tilburg University
a.f.ijdens@tilburguniversity.edu

Sander Verhaegh
Tilburg University
a.a.verhaegh@tilburguniversity.edu

Ties van Gemert
Tilburg University
p.t.vangemert@tilburguniversity.edu

Overall abstract for the session:
Logical empiricism emerged from discussions between philosophers and scientists in Vienna, Berlin, and Warsaw. But it transformed philosophy far beyond these cities when its best-known representatives fled fascism and sought refuge outside Central Europe. Allies and students hailed its clarity and rigor, welcoming its "analytic" approach as a fruitful method to understand the logic of science. Critics dismissed it as sterile and reductive and used it as a prime antagonist in formulating alternative approaches to philosophy of science.
How could a small group of academic migrants leave such a deep imprint on mid-twentieth century philosophy? And why were the exiled empiricists particularly successful in the United States? Historians of philosophy of science tend to be "text- and argument-focused" (Richardson 2012, 1) and to reconstruct the development of the discipline as an exchange of philosophical arguments. Yet any such approach seems insufficient to answer the above questions. This symposium argues that we need to turn to academic practices to understand the reception of logical empiricism. In four talks, we will explore several such practices – e.g. pedagogy and education, academic and political culture, and research networks and institutions – using the reception of logical empiricism in the United States and France as case studies.
First, Nina IJdens will use the records of the Emergency Committee of Displaced Foreign Scholars to discuss the importance of pedagogical practices. She will show that American administrators were primarily looking for competent teachers and that the logical empiricists were deemed a better fit than their idealist refugees. Next, Claudia Cristalli introduces Charles Morris's contribution to the American reception of logical empiricism, arguing that both his institutional work and his empirical study of values are instances of philosophy of science in practice. Third, Sander Verhaegh constructs public debates about the role of science in American democratic practice to help us better understand logical empiricism's success in the United States. Finally, Ties van Gemert will discuss the development of a distinct epistemological tradition, articulated through a strong connection between scholarly practices and moral virtues, thereby illustrating how logical empiricism functioned as an antagonist in formation of philosophy of science in France.

All in all, this panel will contribute a fresh perspective on the historiography of intellectual migration by interrogating the "success story" of logical empiricism in the United States and the development of a parallel and distinct epistemological tradition in France.

Individual papers

(1) From Hörsaal to classroom: the reception of émigré philosophers in American education

In an interview with Partisan Review editor William Barrett, German exile Hannah Arendt somewhat frustratedly commented on some of her fellow émigrés' rise to prominence in American academia: "I can't understand why you Americans have taken to these second-rate European positivists that have come here. In Europe people like Carnap and Hempel were jokes … Here you take them seriously" (Barrett 1982, 104). One ought not share her frustration to share her puzzlement. Germanophone philosophy had been a multivoiced conversation between neo-Kantians, phenomenologists, existentialists, and Lebensphilosophen, with logical positivism operating on the fringes of institutional philosophy. While representatives of all movements sought to continue their intellectual work across the Atlantic during the war, logical empiricism quickly came to occupy American philosophy's center of attention.

In this paper, I argue that pedagogical practice played an important role in the divergent reception histories of German émigré philosophers in the US, and the success of logical empiricist philosophers in particular. Historians of philosophy have mostly devoted their attention to a reconstruction of the intellectual context in which the émigrés arrived. But the exiles did not arrive with their philosophical ideas alone, nor did their American counterparts evaluate them exclusively in terms of their philosophies. In 1930, American pragmatist Sidney Hook noted that "American philosophers teach while German philosophers lecture," except for some of the more scientifically inclined philosophers. Logical empiricism intervened in the philosophical landscape by modelling philosophy on the natural sciences, opposing metaphysics and emphasizing rigor, clarity and communicability. I argue that these commitments extended to their pedagogical practice, which was more congenial to American academic culture than the oratory style of their fellow émigré philosophers.

My research is based on the archival collection of the Emergency Committee of Displaced Foreign Scholars, which assisted over three hundred European refugee intellectuals in finding employment at American academic institutions between 1933 and 1945. This collection allows historians to reconstruct the obstacles these refugee philosophers faced in finding positions at American universities.

First, I show that American administrators were primarily looking for competent teachers for undergraduate education. Second, I illustrate how émigré philosophers' pedagogical practices were received at American institutions of higher education, showing that many German philosophers struggled to cater to American students' needs. Third, I show that these different reception histories mapped onto different philosophical orientations. Scientific philosophers such as Carnap and Reichenbach were well-received. Many other émigré philosophers, most notably idealist philosophers, struggled to communicate their thought to American students. Some were considered aloof, detached, contemptuous and "unable to bridge the gap between their intellect and their public." I conclude with a reflection on the fruitfulness of this practice-oriented history.

(2) The Unity of Science in Practice: Charles W. Morris on Value

In the Preface of his major empirical work – Varieties of Human Value, 1956 – Charles W. Morris stated that his aim was "being of service of man in his characteristic activity as a valuer" (Morris 1956: vii). Yet, the book is not a manual on how to orient oneself in a value-laden world. Morris' additional aim was "to bring the socio-humanistic disciplines within the scope of the program of unified science" (1956: vii-viii), and indeed Morris played an active role in the Unity of Science project, besides being one of the editors of the Encyclopedia (Reisch 2005). Alongside his institutional work, Morris also pushed a particular vision of what the unity of science may look like. While positioning himself as a bridge between logical positivists and American pragmatists, he advanced a version of scientific philosophy relying on first-hand scientific practice.

In an earlier essay, he had noted that the study of values would require "not only an operationally defined terminology but the development of experimental programs" (Morris 1949: 219). My proposal is that Varieties can be read as a repository not of new experiments but of exploratory measurements, which future research might take up, drop, or alter significantly. Situated at the intersection of behavioristic psychology, Talcott Parson's school of sociology, and psychometrics, Morris' study is an original attempt at putting his philosophy of science in practice. This paper reads Varieties along two lines: as an expression, in practice, of Morris' original reception of logical positivism; and as a case study for philosophy of science in practice.

My argument is articulated in two steps. First, I reconstruct Morris' philosophical commitment to scientific empiricism through a close reading of his correspondence, particularly with Otto Neurath. Their exchange lasts almost eleven years, between July 1934 and Neurath's death in December 1945. They provide crucial insights on the perceived relevance of a scientific study of value as a gap to be filled in order to make science truly unified (Reisch 2005; Uebel 2013). They also help situate the philosophical discussion in its political and social context, thereby contributing to a historicized understanding of the unity of science project. Next, I turn to Morris' venture into the empirical study of values. While the relevance of his work for the logical positivists diminished, he was not a lonely investigator. In his practice, he blended insights from different disciplines in an attempt to find meaningful correlations between answers to questionnaires on preferred "Ways of Life" and various parameters, from geography (he collected data from many different countries, including India and China) to societal, psychological, and behavioral factors. Shortcomings notwithstanding, Morris' work is an interesting example of a kind of philosophy of science which set to create in practice the kind of science demanded by its philosophical commitments. It thus represents an interesting reversal of the most common approaches to philosophy of science today, where philosophers start from the observation of existing scientific practices to reflect philosophically on some ethical, epistemic, political, etc., aspects of them.

(3) Kulturkampf: The Cultural Reception of Logical Empiricism in North America

The Vienna Circle is in part a product of Europe's cultural conflicts. It emerged in a period when disputes about science were deeply entangled with socio-political divides (Romizi 2012). Once in the United States, the exiled empiricists became embroiled in a different but related Kulturkampf, this time concerning the role of science in American democracy. Progressive forces, including many pragmatists, had long sought to steer public culture in a new direction, promoting science as a resource for strengthening U.S. democratic practice (Hollinger 1995; Jewett 2012). Conservatives, however, defined America in religious terms

and dismissed the progressive agenda as a threat to democracy. If we embrace a pragmatist perspective, they argued, democratic and moral values are not absolute or God-given but contingent and dynamic and, as such, cannot provide a firm foundation for the American way of life.

The sociologist J.D. Hunter, who first introduced the term 'culture wars' to describe America's socio-cultural conflicts, has argued that most of the country's public disputes—e.g. on abortion and civil rights—are at bottom a clash between conflicting "moral epistemologies" (1991, 289). He identifies two coalitions of religious and secular groups and characterizes them in terms of their different perspectives on the locus and nature of moral truth: An orthodox alliance of conservatives who insist that values are absolute; and progressive coalition which regard morals as dynamic and grounded in the evolving needs of society.

Hunter's analysis has been influential in shaping twenty-first-century discussions about culture conflicts in the United States. Yet his framework also offers a useful lens through which to understand the reception of logical empiricism. The exiled empiricists became a prominent voice in U.S. philosophy in the 1930s, right when discussions about the source and nature of American values dramatically intensified in response to the spread of totalitarian ideologies in Europe. An analysis of these discussions, I show, reveals that the logical empiricists were frequently targeted by (1) conservatives who emphasized the need for objective moral standards and by (2) progressives who associated it with a modern, scientific outlook on life.

This paper analyzes logical empiricism's cultural reception and role in America's Kulturkämpfe through a close reading of hundreds of articles in newspapers and magazines, showing how it functioned as a scapegoat for conservatives and as an exemplar for progressives. A study of this cultural reception, I argue, can teach us some valuable lessons about mid-twentieth-century American intellectual life. First, it can help us better understand logical empiricism's popularity and cultural impact, both inside and outside academic philosophy. Second, it can help shed new light on its internal development after the intellectual migration. Existing work on this period tends to focus on the conservative side of the Kulturkampf, showing how anti-communist pressures stimulated the logical empiricists to adopt a more politically neutral guise. My reconstruction, however, suggests that this cultural analysis is too one-sided and that a more complete picture should account for logical empiricism's momentum in progressive circles.

(4) Logical Empiricism in France: Scholarly Practices and Moral Virtues

In response to a 1938 lecture by the philosopher H.J. Pos on the Unity of Science, the philosopher Léon Brunschvicg succinctly set out his view on the moral virtues derived from the scholarly practice of studying the history and philosophy of science. First, Brunschvicg emphasized that the philosophical and historical inquiry into the natural sciences instills in the scholar a concept of necessity that can guide all their moral actions, making one not only a better historian but also a better human being. Second, through the practice of studying the history of ideas, the scholar becomes intimately acquainted with 'the development of the human spirit' and all its 'dramatic phases', thereby overcoming presentist prejudices, and abandoning the notion of history as linear and progressive (Brunschvicg 1938, 51).

Brunschvicg regarded these virtues as endangered by the rising influence of logical empiricism and the Unity of Science movement criticized by Pos. By abandoning the autonomy of philosophy, he argued, they also discarded the scholarly and moral virtues of the practice of studying the history and philosophy of science: "I fear that the unilinear theories

… which your criticism has attacked … should be considered as the expression of a perennial philosophy" (Brunschvicg 1938, 52).

In this presentation, I will trace this conception of the relation between scholarly practices and moral virtues of the historian and philosopher of science in France in the encounter with logical empiricism before the Second World War. My objective is to show that within the French academic milieu of the Third Republic, there existed a distinct humanist ideal associated with the practice of the history and philosophy of science—one that viewed human actions as governed by necessity, and social positions as determined by history. In confronting logical empiricism, this ideal came under increasing pressure, deepening the antagonism toward positivist philosophy of science among French historians and philosophers of science.

First, I will explore this ideal in the writings of Hélène Metzger, both in her letters with Otto Neurath (1882-1945) and her reflections on fascism. Next, I will discuss the relation between philosophy and practice in the works and life of Jean Cavaillès. A resistance fighter shot by the Nazi's, Cavaillès drew on the very same idea of necessity evoked by Brunschvicg when asked about his activities: 'I believe we submit to the necessary everywhere' (Canguilhem 1967, 28). In his book On Logic and the Theory of Science (1947), this concept of necessity played a decisive role in his critique of Carnap. Finally, I will demonstrate how this humanist ideal was appropriated by Louis Althusser during the 1960s, resulting in his idea of the practice of philosophy as "class struggle in theory". I will show how Althusser's concept of neo-positivism as anti-philosophical led him - despite his commitment to scientism - to retain a concept of the autonomy of philosophy, In conclusion, I will critique the idea that there are particular moral virtues accompanying the study of the history and philosophy of science, arguing that the connection between theory and practice is historically mediated, and often arbitrary and contingent.

Maps, Graphs, and Sonofications: Data Viz and Scientific Practice(s)

Corey Dethier
Clemson University
corey.dethier@gmail.com

Brandon Boesch
Morningside University
boeschb@gmail.com

Jamee Elder
Tufts University
jamee.elder@gmail.com

James Nguyen
Stockholm University
james.nguyen@philosophy.su.se

While philosophers of science have recognized that scientific representations extend beyond models and theories since at least the middle of the twentieth century, relatively little attention has been paid to the sheer variety of representational vehicles that can be found in the sciences. In the meantime, "data viz" has exploded as both practice and topic within the sciences themselves. Non-linguistic representations—such as maps, graphs, and sonofications—play an increasingly central role in scientific publications, generate substantial controversy, and open up new pathways for both evidential reasoning and citizen science.

In this panel, we explore the diversity of these representations, the challenges and opportunities that they afford scientists, and the possibility of extending philosophical accounts of models to representations more broadly. Our central theme is that the details matter: though maps, graphs, and sonofications are all used to express and communicate information in non-linguistic ways, the differences in how the vehicles encode that information matter greatly to the practical uses that they are put to and the potential ways in which they can fail or be challenged.

Specifically,

First, Brandon Boesch will examine the use of snowfall maps by meteorologists, arguing that they undertake a variety of practices to legitimate or "license" their maps in a manner that is usefully compared to how scientists license models.

Second, Corey Dethier will turn to the use of graphs by climate denialists, showing how subtle manipulation of axes and plotting choices can be used not just to mislead but to convey outright falsehoods.

Third, Jamee Elder will argue that sonofications—of black holes, in this case—cannot simply be understood in terms of "what it sounds like" or even "what it would sound like"; close attention to how sonoficiations are used complicates traditional vision-first metaphors found throughout the literature on scientific representations.

Finally, James Nguyen will step back to consider how maps, graphs, and sonofications can be fit into a broader picture of scientific representations, arguing that the diversity of such representations is not a fatal challenge to an informative unified account.

Paper abstracts:

1. “Representational Licensing of Snowfall Forecast Maps”
Brandon Boesch

Though they both involve the prediction of the quantity of precipitation in different locations, there are important practical differences between forecasting rainfall and snowfall. In both cases, meteorologists rely upon guidance from computational simulations of the atmosphere to inform their forecasts. But snowfall forecasts have challenges that rainfall forecasts do not. These challenges lead to a variety of potential misunderstandings of the public consumers of snowfall forecast graphics, something meteorologists work hard to control and manage. In this presentation, I will argue that the management of interpretation of snowfall maps is a form of what I have previously described (Boesch 2017; Boesch 2022) as “representational licensing.” In this case, the target community is the broader public which modifies how the licensing works.

Among the many challenges of issuing snowfall forecasts is that not all snow is created equal, in terms of the ratio of snow depth to equivalent liquid water depth. At higher ratios, small errors in predicted liquid can result in large differences in impact--two tenths of an inch of liquid being the difference between a dusting and nearly six inches of snow. Furthermore, snowfall sometimes occurs as part of a system with rain and snow in different areas. When and where the rain changes from rain to snow can make a big difference in snowfall. There’s also the fact that social media’s algorithms are not designed to promote the most recent snowfall forecast maps, but rather those which are most likely to increase your interaction. Old and outdated maps often persist, confusing the public. Finally, the salience of a missed forecast is much higher for snowfall than it is for rain. To measure rainfall, you need a rain gauge which you empty before and check immediately after an event. It is far easier to measure snowfall: all that is needed is a ruler--something my five-year-old can (and does!) do.

Altogether, snowfall forecasting provides challenges for not only the accuracy of a forecast, but also the reputation and perceived legitimacy of the forecaster and their agency (Fine 2008). Missed forecasts threaten that--and meteorologists have a well-documented bias for avoiding these reputational risks (Boesch 2021; Herman et al. 2018). Snowfall maps are no different. Meteorologists use a variety of methods to constrain and invite particular interpretations and readings of snowfall forecast maps. Methods include offering probabilistic maps, highlighting areas of uncertainty, time-stamping maps, identifying whether a boom or bust is more likely, among others.

These methods, I will argue, constitute a public-facing form of representational licensing. In previous work (Boesch 2017; Boesch 2022), I have described licensing as the intended representational uses of a representational vehicle, established via scientific practice--the development, use, and modification of a vehicle over time. The concept of representational licensing was previously understood as internal to scientific practice, as a means of establishing and maintaining “intended uses” of a vehicle. In the context of snowfall forecast maps, the constraints are intended for the external public audience, which extends the concept to science communication.

References

Boesch, B. (2017). There Is a Special Problem of Scientific Representation. Philosophy of Science 84 (5): 970-981. https://doi.org/10.1086/693989

Boesch, B. (2021). Diagrammatic Representations of Uncertainty in Meteorological Forecasting. In: Basu, A., Stapleton, G., Linker, S., Legg, C., Manalo, E., Viana, P. (eds) Diagrammatic Representation and Inference. Diagrams 2021. Lecture Notes in Computer Science, vol 12909. Springer, Cham. https://doi.org/10.1007/978-3-030-86062-2_48
Boesch, B. (2022). A Concrete Example of Representational Licensing: The Mississippi River Basin Model. Studies in History and Philosophy of Science 92: 36-44. https://doi.org/10.1016/j.shpsa.2022.01.002
Fine, G. A. (2008). Authors of the Storm: Meteorologists and the Culture of Prediction. University of Chicago Press.
Roebber, P. J., S. L. Bruening, D. M. Schultz, and J. V. Cortinas , Jr. (2003). Improving Snowfall Forecasting by Diagnosing Snow Density, Weather and Forecasting 18, 2: 264-287, doi: https://doi.org/10.1175/1520-0434(2003)018<0264:ISFBDS>2.0.CO;2
Herman, G.R., Nielsen, E.R., Schumacher, R.S. (2018). Probabilistic verification of storm prediction center convective outlooks. Weather and Forecasting 33, 1: 161-184. Doi: https://doi.org/10.1175/WAF-D-17-0104.1

2. “The Case of the Mislabeled Axis”
Corey Dethier

In 2015 testimony before congress, John Christy (2015) presented a series of graphs purporting to show that climate models are badly wrong about how quickly the mid-troposphere is warming. His graphs have been widely criticized by other climate scientists (see, e.g., Schmidt 2016), who argue that Christy’s graphs are misleading and that the same data can be graphed in a way that makes the apparent divergence disappear. Who's right? What does it even mean for one graph to be right and another wrong, given that they're constructed using the same data?
In this talk, I borrow some plausible principles of graphical semantics from Perrini (2005) and Dethier (2025) to analyze the debate between Christy and his critics. I suggest that a necessary condition on a graph being a “true” or “accurate” representation is that each element of the graph is correctly placed according to the interpretative frame given by the graph itself. True graphs can be misleading when the elements are accurately placed in this sense but where the interpretative frame is poorly (or dishonestly) chosen.
Christy’s graphs are clearly at least potentially misleading in this sense: his choice of interpretative frame leads his graph to suggest that errors are growing when they aren’t. Arguably, however, any graph of the trends Christy is highlighting is potentially misleading: there is no interpretative frame that isn’t open to similar or related misreadings. The question is then the familiar one of which of the potential errors is least worrying. (Alternatively, one might conclude that visual comparison of these trends is simply a misguided endeavor tout court.)
The more interesting question is whether Christy’s graphs are not just misleading but untrue. I argue that the answer here is at least plausibly “yes.” The problem stems from a mismatch between the label of the y-axis and Christy’s plotting choices: the axis is labeled as though the temperatures depicted are actual temperatures when Christy’s plotting choices require that they are interpreted counterfactually.

References

Christy, John. 2015. Testimony of Dec. 8 2015 before the U.S. Senate Committee on Commerce, Science, & Transportation, Subcommittee on Space, Science and Competitiveness

Dethier, Corey. 2025. “How do you assert a graph? Towards an account of depictions in scientific testimony.” Nous 59.3: 678-699.
Perini, Laura. 2005. The truth in pictures. Philosophy of Science, 72, 262–85.
Schmidt, Gavin. 2016. Comparing Models to the Satellite Datasets. https://www.realclimate.org/index.php/archives/2016/05/comparing-models-to-the-satellite-datasets/

3. “Making Black Hole Music”
Juliusz Doboszewski, Jamee Elder (presenting), et al.*

In 2019, the Event Horizon Telescope (EHT) published its first picture of a supermassive black hole (The Event Horizon Telescope Collaboration et al. 2019). But taking this picture to represent what a black hole “looks like”, when black holes neither emit nor reflect light, might strike philosophers as confused. Answers to the question of what the black hole looks like must focus on absence; a black hole is invisible and can be inferred only by what we don’t see—i.e., a shadow. What’s more, the EHT picture is based on radio waves, which are not visible to human eyes.
Things are even more puzzling when we switch from visualizations to sonofications because, prima facie, the question “what does a black hole sound like?” is even more difficult to interpret than “what does a black hole look like?”. For some black hole observations there is a plausible answer: gravitational waves emitted by binary black hole mergers are within human audio range and reflect what a human might hear if close enough to such an event (our eardrums would vibrate!). Sonofication of a the “chirp” of a gravitational wave signal can be judged according to a simple counterfactual criterion: does the sonofication sound like what we would hear if we were located near a binary black hole merger?
However, most sonofications are not easily analysed this way. For a start, most data that are sonofied do not correspond to sound at all, let alone sounds we could hear under some easily imaginable modification of our present circumstances (e.g., moving location or extending human audio range). Indeed, many are parasitic on visualisations, re-purposing visual data (e.g., the brightness of light as you move around the rough ring shape present in black hole images) to be presented with sound.
This paper will provide an analysis of what constitutes a good sonofication of a black hole, sensitive to the different roles that such sonofications can play.
First, I will provide a taxonomy of the main purposes of black hole sonofications, ranging from the purely aesthetic, to science education, to scientific research tools. Throughout, I will draw on examples of black hole sonofications in each category in order to provide a sense of how scientists are currently deploying sonofications in practice.
Second, I will draw on philosophical literature about scientific representation including Goodman (1968), Elgin (2010), Frigg and Nguyen (2016), and Parker (2020) and others, to provide a framework for what counts as a good sonofication, qua epistemic representation.
Finally, I flesh this out by drawing on details of specific sonofications and their uses, providing an analysis that is sensitive to intended audiences and their goals. I draw on histories of radioastronomy, space sounds, and crossovers between astronomy and art, as well as empirical studies of how (non-experts) process visualisations and sonofications.
Overall, I provide an analysis of a largely-neglected kind of scientific representation—sonofication—grounded in the broader context of black hole astrophysics. I argue that sonofications, being less familiar and easily interpretable than visualisations, offer both important opportunities and challenges when used as epistemic representations.
*This talk is based on joint work with members of Algorithms, Inference, and Visualisation Focus Group of the next generation EHT’s History, Philosophy, and Culture Working Group.

Full author list as of writing (alphabetical): Juliusz Doboszewski, Jamee Elder, Jonas Enander, Peter Galison, Silvina Maestro, Helen Meskhidze, Paula Muhr, Vitaly Pronskikh, Emilie Skulberg, and Galina Weinstein.

References

Elgin, Catherine Z. 2010. “Telling Instances.” In Beyond Mimesis and Convention, edited by Roman Frigg and Matthew Hunter, 262:1–17. Boston Studies in the Philosophy of Science. The Netherlands: Springer Netherlands. isbn: 9789048138500.
Frigg, Roman, and James Nguyen. 2016. “The Fiction View of Models Reloaded.” The Monist (Oxford) 99 (3): 225–242.
Goodman, Nelson. 1968. Languages of Art: an Approach to a Theory of Symbols. Indianapolis: Bobbs-Merrill.
Parker, Wendy S. 2020. “Model evaluation: An adequacy-for-purpose view.” Philosophy of Science 87 (3): 457–477.
The Event Horizon Telescope Collaboration et al. 2019. “First M87 event horizon tele-scope results. I. The shadow of the supermassive black hole.” The Astrophysical Journal Letters 875 (1): L1. https://doi.org/10.3847/2041-8213/ab0ec7.

4. “Defending Representationalism”
James Nguyen

I map the conceptual landscape populated by accounts of how models (and epistemic representations more generally) relate to their targets by distinguishing four approaches. Representationalists claim that models license inferences because they represent. Inferentialists claim that models represent because they license inferences. Deflationists argue that no substantial account of epistemic representation is available. Artefactualists argue that we should not think of models as representations in the first place. It’s their artefactual status that matters. These latter three positions all reject representationalism, but they do so in different ways and shouldn’t be conflated. And so I defend representationalism on each front.

Suárez’s (2024) deflationism rests on the observation that substantive similarity- or isomorphism-based accounts cannot capture the diversity of epistemic representational practices. Therefore, assuming that no alternative substantive account can, there is no general story to be told about epistemic representation. Hence, all we have is the trivial claim that models licence inferences about their targets. I respond in three ways. I point out that alternative (at least semi-substantial) accounts exist that impose a common structure on this diversity, one realised by different details across contexts. I answer Suárez (2025) recent response to related criticisms, the claim that `oh, Horatio, there are so many more things in heaven and earth than are dreamt of in your philosophy’ i.e. the variety in epistemic representation cannot be captured by a substantial account. Finally, I argue Suárez’s emphasis on inference over representation supports a substantive inferentialist approach, not deflationism.

Such an inferentialism needs clarification. I sketch a thoroughgoing Brandom (1994, 2000) inspired scientific inferentialism that requires exploring the details of the normative social practices that govern how we draw inferences from models, and then argue that the most developed `inferentialist' account (Khalifa et al. 2022) is not so thoroughgoing. Although inferentialism is viable, representationalism has a better track record in explaining how models work. But both positions demand a substantial framework, concerning the structure of

representation relations and how they license inference, or vice versa, in which to situate the specifics of epistemic representations.
De Oliveira (2018, 2022) argues that representationalism is something we `can't have' because it allegedly commits us to simultaneously claiming that models are representations---which requires a mind-dependent account of representation---and that we learn about their targets because of their representational status---which requires a mind-independent account. He further argues that we `don't need' representationalism; we can treat models as artefacts, i.e. objects of extended cognition, instead. I argue the first claim conflates representation with accurate representation; once we distinguish these, we can understand how we learn about targets by exploiting how models represent them, even if specific epistemic representations realise this relation differently. Moreover de Oliveira's `radical', as opposed to `hybrid', artefactualism (the latter being compatible with representationalism (Knuuttila 2017)) lacks the resources to explain how we use models to explore their targets.
This discussion yields a richer picture of representationalism and the relationship between philosophical theories and scientific practice. If a philosophical account is to help us understand any specific instance of epistemic representation it should offer a framework substantial enough to guide an analysis of the practice of exploiting model–target relations, even though those relations may be realised differently in different contexts.

References:

Brandom, R. B. (1994). Making it Explicit: Reasoning, Representing and Discursive Commitment. Harvard University Press, Cambridge MA.
Brandom, R. B. (2000). Articulating Reasons: An Introduction to Inferentialism. Harvard University Press, Cambridge MA.
de Oliveira, G. S. (2018). Representationalism is a dead end. Synthese, 198(1):209-235.
de Oliveira, G. S. (2022). Radical artifactualism. European Journal for Philosophy of Science, 12(2):1-33.
Khalifa, K., Millson, J., and Risjord, M. (2022). Scientific representation: An inferentialist-expressivist manifesto. Philosophical Topics, 50(1):263–291.
Knuuttila, T. (2017). Imagination extended and embedded: Artifactual versus fictional accounts of models. Synthese, 198(Suppl 21):5077–5097.
Suárez, M. (2024). Inference and Representation: A Study in Modeling Science. University of Chicago Press, London
Suárez, M. (2025). Inference and Representation by Mauricio Suárez: Reply by the Author. Studies in History and Philosophy of Science, 114:102088.

Philosophy of Forensic Science

Enno Fischer
Dresden University of Technology
enno.fischer@tu-dresden.de

Saana Jukola
University of Twente
s.j.jukola@utwente.nl

Alexander Gebharter
Marche Polytechnic University
alexander.gebharter@gmail.com

Michał Sikorski
Marche Polytechnic University
michalpsikorski@gmail.com

Edoardo Peruzzi
Leibniz University Hanover
edoardo.peruzzi@philos.uni-hannover.de

Forensic science plays a crucial role in the judicial system, with forensic results serving as key evidence in countless verdicts worldwide. Established forensic procedures, such as fingerprint analysis or DNA testing, involve comparing samples collected from crime scenes with those obtained, e.g., from suspects. Moreover, scientific theories and models from disciplines such as medicine, psychology, or economics are used to support arguments in court.

Forensic science is an understudied area in philosophy of science. This symposium will show that there is much to gain from a more detailed philosophical look at the evidential practices in the forensic disciplines. Forensic science has been diagnosed to be in crisis (Weyermann & Roux, 2021), calling for increased efforts of methodological reflection. Philosophical work on causal reasoning, bias mitigation, the social epistemology of expertise, and model-based reasoning is an important resource for addressing current challenges in forensic science. At the same time, forensic science is an exciting testing ground for work in the philosophy of science. For example, the adversarial character of legal proceedings helps spotlight the role of biases, values, and background assumptions in causal and model-based reasoning and evidentiary practices more generally.

The symposium will bring together perspectives from philosophy of medicine, philosophy of economics and general philosophy of science. While each of the three proposed talks looks at epistemic challenges from the perspective of a specific forensic discipline, the added value of the symposium will be a better overarching understanding of the epistemic practices in the forensic sciences. Common themes to be discussed include the role of forensic expertise, the management of biases in forensic reasoning, and the kinds of evidence that forensic reasoning is based on.

Talk 1: Social epistemology of expert disagreement in forensic medicine
Saana Jukola, University of Twente, s.j.jukola@utwente.nl
Enno Fischer, Institute of Philosophy, Dresden University of Technology, enno.fischer@tu-dresden.de

Unlike most fields of medicine, forensic medicine is not performed to care or to cure. The primary audience for forensic medicine experts are legal courts and their main aim is to determine the cause of death or injury. It is well acknowledged that expert disagreement in forensic medicine is common (e.g., Timmermans 2006), and that it is related to the underdetermination of evidence that is a general feature of forensic medical inquiry (Fischer & Jukola 2024). Moreover, the adversarial nature of many judicial systems means that there is a market for different expert opinions.
Examples of controversial topics in the forensic medical community include the Shaken Baby Syndrome (e.g., Binney 2023) and the Excited Delirium Syndrome (ExDS) (Fischer & Jukola 2024). Expert disagreement concerning, e.g., whether an individual's death can be assigned to one of these causes, can be related to factors operating at the individual-level, such as cognitive biases (e.g., Dror et al. 2021). In this talk, however, we are interested in social epistemic and institutional causes contributing to the expert community being divided on such matters. By focusing on ExDS as a case study, our aim is to identify social epistemic dysfunctions and their signs that contribute to expert disagreement in forensic medicine.
In the US legal context, the so-called Daubert standard provides guidance for the admissibility of expert testimony in court. According to Daubert, judges make decisions about admitting testimony and evidence on the basis of its scientific reliability and relevance. While there are no fixed guidelines for assessing scientific reliability, judges may consider, e.g., whether the testimony is supported by peer reviewed literature and in line with the generally accepted views of the research community. We use Boyd's (2019) concept of 'groupstrapping' to analyze how epistemic authority is constructed. In particular, we show how in the ExDS case, the peer reviewed literature supporting ExDS being a real medical condition and a possible cause of death has been produced by a small number of individuals who often have ties to each other. This serves to highlight two problematic features of epistemic practices in forensic medicine and forensic science more generally. First, the field has an insular character with little crossdisciplinary links and double-checks. Second, there are potential conflicts of interest between forensic medical research and expert testimony.

Binney, N. (2023). Methods of Inference and Shaken Baby Syndrome. Philosophy of Medicine, 4(1), 1-33.
Boyd, K. (2019). Epistemically pernicious groups and the groupstrapping problem. Social Epistemology, 33(1), 61-73.
Dror, I., Melinek, J., Arden, J. L., Kukucka, J., Hawkins, S., Carter, J., & Atherton, D. S. (2021). Cognitive bias in forensic pathology decisions. Journal of Forensic Sciences, 66(5), 1751-1757.
Fischer, E., & Jukola, S. (2024). Bodies of evidence: The 'Excited Delirium Syndrome' and the epistemology of cause-of-death inquiry. Studies in History and Philosophy of Science, 104, 38-47.
Timmermans, S. (2006). Postmortem: How medical examiners explain suspicious deaths. Chicago: University of Chicago Press.

Talk 2: The relevance principle from a causal perspective
Alexander Gebharter, Department of Biomedical Sciences and Public Health, Marche Polytechnic University, alexander.gebharter@gmail.com

Michał Sikorski, Department of Biomedical Sciences and Public Health, Marche Polytechnic University, michalpsikorski@gmail.com

The relevance principle plays a central role in the methodology of forensic science. It is considered the gold standard for determining which information should be used when evaluating a piece of evidence for a hypothesis of interest. In essence, it states that one should consider only task-relevant information - that is, information that would affect the likelihood of the outcome of one's evaluation if the hypothesis were true (or false).
Case studies and simulations demonstrate that using irrelevant contextual information (e.g., results from other forensic disciplines or a confession of guilt) can lead to biased forensic results, double counting, and ultimately wrongful convictions (e.g., Garrett and Neufeld, 2009, Cuéllar et al., 2022, Sikorski, 2022). Due to these concerns, the relevance principle has gained popularity in forensic science and inspired significant methodological reforms. For instance, sequential unmasking (Krane et al., 2008, Robertson and Kesselheim, 2016) involves concealing nonessential information from the expert at a given stage of the analysis. Recently, it has been argued that the relevance principle and sequential unmasking could also be applied to other scientific disciplines (see, e.g., Sikorski, 2022, or Shimizu and Dror, 2023).
In this presentation, we will put the relevance principle to the test by examining its performance from a causal perspective. Specifically, we will consider the most elementary causal relationships among a hypothesis, a piece of evidence, and additional information, and then explore whether the relevance principle successfully prevents bias and double-counting in these scenarios.
Our analysis reveals that, in some of the basic causal structures, the original version of the relevance principle fails in this regard. We will then delve deeper to define general conditions under which double counting occurs and develop a causal version of the relevance principle that successfully prevents the double-counting fallacy in all elementary causal settings.

Cuéllar, M.-F., Mauro, J. A., & Luby, A. (2022). A probabilistic formalisation of contextual bias: From forensic analysis to systemic bias in the criminal justice system. Journal of the Royal Statistical Society: Series A (Statistics in Society), 185, S620–S643.
Garrett, B., & Neufeld, P. (2009). Invalid forensic science testimony and wrongful convictions. Virginia Law Rev., 95.
Krane, D. E., Ford, S., Gilder, J. R., Inman, K., Jamieson, A., Koppl, R. G., Kornfield, I. L., Risinger, D. M., Rudin, N., Taylor, M. S., & Thompson, W. C. (2008). Sequential unmasking: A means of minimizing observer effects in forensic DNA interpretation. Journal of Forensic Sciences, 53.
Robertson, C. T., & Kesselheim, A. (2016). Blinding as a solution to bias: Strengthening biomedical science, forensic science, and law. Academic Press.
Shimizu, T., & Dror, I. (2023). History information management strategy for minimising biases and noise for improved medical diagnosis. BMJ Open Quality, 12.
Sikorski, M. (2022). Is forensic science in crisis? Synthese, 200(3), 1–34.

Talk 3: On the Role of Scientific Models in the Law
Edoardo Peruzzi, Institute of Philosophy, Leibniz University Hannover, edoardo.peruzzi@philos.uni-hannover.de

Scientific models are applied in legal proceedings as forensic tools in a variety of legal realms. Forensic models (e.g., ballistic or bloodstain pattern analysis), psychological models (e.g., risk assessments or mental state evaluations), and financial models (e.g., estimating

damages in patent or fraud cases), for example, routinely find their way into the courtroom. In this article, I argue that applying models in a legal context raises distinctive philosophical and practical challenges, and I advance two main claims. First, I contend that model-based evidence is often discarded by fact-finders not because it lacks relevance or reliability, but because it possesses what I call a broader attack surface—that is, it is especially vulnerable to criticism and dispute within the adversarial setting of the courtroom. Second, I argue that disregarding model-based evidence can severely undermine the quality of judicial decisions by depriving fact-finders of a robust source of evidence.

To support these claims, I focus on the use of economic models by expert witnesses in U.S. legal proceedings involving alleged violations of antitrust or competition law. In this context, models serve as sources of evidence intended to influence judicial outcomes. Economic experts typically construct empirical models by combining theoretical frameworks with econometric techniques and apply them to generate quantitative results about the specific markets under scrutiny. Judges, in turn, evaluate the strength and reliability of this model-based evidence when reaching their verdicts. Despite their widespread use, economic models frequently fail to influence judicial decision-making, thereby undermining the rationale for introducing them into courtroom proceedings. The consequences of dismissing model-based evidence are significant. As I shall argue, such evidence is often overshadowed by weaker and less reliable forms of evidence. This displacement, I suggest, may compromise the quality of judicial decisions, as fact-finders are forced to rely on a less robust evidentiary basis.

To explain why economic models are particularly fragile or vulnerable when used as sources of evidence in legal proceedings, I introduce the notion of an "attack surface". More specifically, I argue that model-based evidence has a large attack surface, that is, it is especially susceptible to criticism from the opposing side in an adversarial setting. This vulnerability arises because economic models used in courtrooms are empirical models with an exceptionally high degree of freedom. By degrees of freedom, I refer to the set of modeling choices available to the modeler. This high degree of freedom provides ample room for critique and even enables the construction of competing models, the evaluation of which often exceeds the expertise of even the most well-intentioned judges. As a consequence, because they are simpler and have a smaller attack surface, other forms of evidence, while less reliable, may take precedence in judicial decisions. I conclude by reflecting on whether similar issues arise in the context of forensic techniques with stronger consensus on methods of evidence gathering, such as DNA analysis.

Navigating Broad Shallows and Narrow Depths in the Study of a Large Interdisciplinary Life Sciences Research Project

James Griesemer
University of California, Davis
jrgriesemer@ucdavis.edu

Kelli Barr
University of California, Davis
barr@ucdavis.edu

Andrew Bollhagen
University of California, Davis
aabollhagen@ucdavis.edu

Shane Jinson
University of California, Davis
stjinson@ucdavis.edu

Edward Wilson
Edward W. Wilson Consulting
edwilsonconsulting@gmail.com

In this session we describe and critically assess a variety of research methods we are using to study a large interdisciplinary research project in the life sciences. Agency, Directionality and Function (ADF), otherwise known as the Purpose Project, brought together more than 200 scholars -- biologists from various specialties, philosophers, modelers, and others -- on the assumption that diverse perspectives were needed to advance the understanding of teleological phenomena in living systems. The John Templeton Foundation, which supported ADF, funded an independent but concurrent project known as the Sociology of the Purpose Project ("SoPP" JTF #62385) to study ADF as it developed. Comprising a team of social scientists and philosophers, SoPP has had the unusual opportunity to investigate interactions within a large, disparate collection of scholars that potentially could lead to new scientific theories, innovative research practices and novel discoveries, and disciplines. Although ADF comprised 24 individual research projects, it also included a number of cross-project forums designed to foster interactions among members of different teams, such as summit conferences, reading groups, and "cluster" meetings that convened three or four project teams to encourage cross-project communication and collaboration. This project design afforded ample occasions for informal interactions, such as lunchtime discussions or meetups in bars. Our SoPP team has theorized that such "informal organizations" are an important first step that may lead to more established groups working together to produce integrative interdisciplinary research and, perhaps, new scientific discoveries. Understanding where, how, and to what extent the trajectory from casual organizations to new, integrative research efforts has unfolded within ADF has required multiple research methods, each providing a window onto complex social dynamics while leaving key questions unanswered. These methods can be characterized as either broad but shallow or deep but narrow. The papers of this session explore various scales of social organization and interaction among ADF participants, from local interactions among small numbers to mid-sized project groups to networks formed across the whole participant membership and beyond.

James Griesemer
Natural History of Thickety Social Experiments in Interdisciplinarity: The ADF Program, 2022-2024

The Agency, Directionality and Function (ADF) research program explored three core concepts that may be central to an interdisciplinary, life-science-focused "science of purpose." The program, funded by the John Templeton Foundation from 2022-2024, sponsored 24 sub-awards, each with its own principle investigator and research group. The program was administered by an executive committee of staff located at the University of Minnesota and overseen by a scientific advisory board, comprising ~200 participants. The program organized the 24 projects in 7 clusters of 3-4 projects each. Some groups were geographically and organizationally distributed, others located in a single university location. Meetings of clusters and of the whole program's participants included outside commentators. This paper reports on a partial, ongoing "natural history" of the project. It offers a timeline of the ADF's development and a characterization of its participants and component projects. The larger project on which this report is based is designed to describe the participants and their projects, characterize the scientific practices and relationships among participants (those prior to, continuing in, and established during the ADF program), as reported in publications, interviews, and surveys, and through observation of sub-award and cluster meetings. The natural history aims to contribute to a sociological, philosophical, and historical study of emergent forms of social organization involved in the pursuit of inter- and multi-disciplinary research in the life sciences. Here, we report on a pilot study bringing together qualitative and quantitative methods of analysis to identify and characterize practices, participant relationships, and modes of interacting that appear to be connected to the forms of social organization that took shape during the ADF program. The methodology is to describe, characterize and analyze a stratified sample of publication outputs from ADF projects, which number somewhere between roughly 200 and 7,000, depending on inclusion criteria. The study design randomly samples up to 4 papers from each of the 24 sub-award project teams. We read each of the sampled papers and code them for practices (identified as tasks and sequences of tasks organized as work "flows" reported by the authors as leading to the findings of the published papers) and relationships among the authors (as reported in the publications and other sources from our interviews, surveys, and information gathered in meeting observations and the ADF website). The aim of the natural history is to discover interaction patterns among participants, projects, and clusters on display in a timeline of the whole ADF program. An interpretation of ADF's social complexity is offered using tools from Wimsatt's account of "interactional complexity." We extend his characterization of "causal thickets" to social processes of interdisciplinary science in practice.
Wimsatt, William C. 2007. Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality. Cambridge, MA: Harvard University Press.

Kelli Barr and Andrew Bollhagen
Talking It Over: Toward a Social Epistemology of Deliberation in Science

Social epistemological theorizing in the philosophy of science seeks to understand how social processes shape the knowledge scientific communities produce. Recent work has analyzed a range of epistemically relevant social interactions. Scientists disagree, testify, advise, communicate in networks, and form consensus. They also deliberate. The first goal of this paper is to characterize deliberation as a scientific practice and an interaction of social epistemological interest (following Longino 2022). Our second goal is to develop cases from our ongoing philosophical and sociological investigations of a large-scale, interdisciplinary

project in evolutionary biology – the Agency, Directionality, and Function (ADF) project (2022-2024) – to analyze the epistemic and epistemological import of deliberative interactions. Agents deliberate when they engage an open-ended discussion about possibilities viz-a-viz an observation, assessment, or evaluation, with the aim of making a determination about what to do. Deliberation is open-ended in that new possibilities can be put on the table during the interaction. Take the following situation as an illustrative example: A frog embryo has been prepared so that a researcher can observe the gastrulation phase and measure notochord development. The researcher looks through the microscope and sees an embryo with an elongated, oblate shape. This seems unusual. The researcher asks the PI of their lab whether the embryo is exhibiting exogastrulation, a classic defective phenotype familiar to developmental biologists. The PI looks through the microscope and, instead of answering directly, notes the position of the blastopore, the location of which is more characteristic of normal gastrulation than of exogastrulation. The PI and researcher determine that the embryo is gastrulating but is not undergoing exogastrulation. They name this novel phenomenon xenogastrulation (Bollhagen, Mayne, and Merzdorf 2025). In response to the indeterminate character of the situation, the investigators drew on background knowledge and assumptions, checked them against one another, proposed options for proceeding, and reached an agreement about how to classify what they saw. Deliberation exemplifies what we will call a constructive interaction. In situations where no existing belief, attitude, assumption, or fact of the matter is at stake, agents must construct their bases for judgment. Further, they often do this in constitutively social ways. We think this is both epistemically and epistemologically significant. Constructive interactions are one way researchers overcome situations of practical underdetermination. Determinations reached through deliberation likely rest on some combination of epistemic standards, background knowledge and assumptions, values, organizational policies, and perhaps other phenomena. Further, since practical underdetermination is pervasive in science, our analysis suggests that an adequate social epistemology of science needs an account of both the right kinds of critical interactions and the right kinds of constructive ones.

Bollhagen, A., Z. J. Mayne, and C. S. Merzdorf (2025) “The Entanglement of Individuation and Explanation in the Discovery of Xenogastrulation.” History and Philosophy of the Life Sciences, 47 (4): 53. doi: 10.1007/s40656-025-00706-1.
Longino, H. (2022) “What’s Social about Social Epistemology?” Journal of Philosophy, 119 (4): 169–195.

Shane Jinson
Bringing It All Back Home: Interdisciplinary Workers from Diverse Social Worlds Unite Large Collectives with Common Knowledge

In this paper I break down the claims two or more Purpose Project participants jointly made in a given collaborative research paper when they introduced a biological problem and framed it in the context of conventional or contemporary understanding. I find that the wider the gap in knowledge between these collaborators, the broader and far-reaching their framing is regarding applicability to the goal of the overall project. Similarly, participants with narrower gaps between (or especially overlapping) specialized knowledge areas tend to address more specific problems that can obscure project aims. This tendency of collaborators from more remote specialties to be far-reaching and broad is likely the result of gaps in common knowledge that presumably prevent a melding of ideas at more specific and zoomed in areas. However, they deftly represent a project’s overarching theme due to the need for common ground separate from each collaborator’s specialty.

Edward Wilson
Webs and Thickets: Uses and Limitations of Social Network Analysis

One example of a broad, shallow research method is network analysis, a set of tools for visualizing and describing relationships among a set of actors — in this case, the researchers and advisors who took part in ADF. Our applications of network analysis revealed some meaningful patterns. For example, they showed that ADF effectively fostered relationships between members of different project teams, and between researchers with diverse disciplinary backgrounds. They identified individuals who played central and influential roles owing to the large number of relationships they had formed. They revealed that adherents of one distinctive theoretical approach had formed relationships throughout the network of ADF participants, suggesting that ADF was helping to disseminate their ideas beyond what had been a somewhat insular academic community. And they showed that the "clusters," designed to bring together researchers presumed to have overlapping interests, were not especially effective at fostering new relationships; most new relationships were between members of different clusters, suggesting that casual meetups outside of structured discussion forums were more likely to result in productive interpersonal relationships. While our network analyses have produced insights into the configuration of relationship webs, they were necessarily based on very thin information about the nature of those relationships. Networks constructed from survey responses relied on self-reports about people who have influenced the respondents' research, but the resulting data were inconsistent as different respondents interpreted the same question in different ways. We also constructed networks based on membership in new teams that had formed to prepare proposals for follow-on research, but we lacked information on social processes within those teams and the interactions that led to their formation. Similar comments apply to co-authorship networks. Network analyses treat interpersonal linkages as essentially equivalent and portray social worlds as webs of uniform dyadic relationships. Following Wimsatt, however, we believe social worlds are more aptly envisioned as "thickets" comprising multifarious, dynamic relationships among entities of different kinds and different hierarchical scales. Network analyses simplify complex social realities by modeling messy thickets as if they were web-like structures composed of similar nodes linked together in similar ways. A more nuanced understanding social thickets requires complementing such broad, shallow research methods network analysis with narrower, deeper methods like those discussed in other papers in this session — although these, to be sure, are also limited, producing analyses that are particularistic and perhaps unrepresentative of the larger whole. Broad, shallow perspectives and narrow, deep methods are both needed to produce rich descriptions of social processes and patterns such as those that have developed and continue to unfold within the ADF cohort.

Technology-driven scientific change: Conceptual, historical, and ethnographic perspectives on single-cell biology

Fridolin Gross
University of Bordeaux
fridolin.gross@u-bordeaux.fr

Meike Wiegand
University of Bordeaux
meike.wiegand@u-bordeaux.fr

Ignazio Arrabito
University of Bordeaux Montaigne
ignazio257.ia@gmail.com

Patrick Ferree
University of Copenhagen
patrick.ferree@sund.ku.dk

Data-driven sequencing technologies are central to research across the life sciences. Over the past 15 years, single-cell technologies (SCTs) have emerged as some of the most prominent—and contested—of them, with researchers caught between enthusiasm about their potential for discovery and concern over their disruption of standard scientific practices. Traditional sequencing technologies rely on bulk processing, i.e. gathering information about the average molecular content of all cells in a sample, while SCTs allow scientists to profile molecular properties of individual cells. This not only adds to the promise of this technology, but also refreshes debates about the classification of cells into different types and inspires the production of data-rich resources for the community, such as comprehensive "cell atlases" of tissues, organs, and even whole animals. The goal of this symposium is to challenge and contextualize some of the dominant narratives about SCTs by drawing on tools from philosophy, history, and social studies of science.

Our first two contributions focus on theoretical implications. SCTs foster a more data-driven form of knowledge production in a move away from theory, thus generating classifications that are supposedly not biased by background assumptions. Fridolin Gross advocates for replacing the ideal of theory-free classification by a weaker version of "theory-minimizing" given that novel technologies are always situated within a history of epistemic iteration and therefore shaped by existing knowledge and practices. Meike Wiegand locates SCTs at the intersection of two distinct research traditions, biological and computational, with the latter moving closer to the aforementioned theory-free ideal. By comparing sequencing-based and imaging-based practices of cell-type classification, the author shows the significance of data integration when it comes to understanding technology-mediated scientific change.

Our last two contributions focus on social and epistemological dimensions of cell atlases. The Human Cell Atlas (HCA) is one of the largest-scale biological endeavors of the century, employing single-cell technologies to generate an open-access database about all the cells of each tissue and organ of the human body. Ignazio Arrabito conducts a comparison between the HCA and the Human Genome Project, another big-science project aimed at producing a foundation for the field. The author questions the effective value of such resources for scientific practice and highlights both changes and continuity in the social organization and infrastructure building capacity of large-scale sequencing projects. Building on the previous paper, Patrick Ferree turns our attention to the fact that many labs continue to produce their

own cell atlases as references and tools for intervention. Drawing on ethnographic fieldwork and philosophical analysis of the production and use of three particular cell atlases, the author shows how epistemic aims and values interact and shape the usefulness of novel technologies.
Collectively, these contributions develop a multifaceted perspective on SCTs and technology-driven scientific change, foregrounding both the production of knowledge and its uneven translation into scientific practice.

Towards theory-free classifications in single-cell biology?
Fridolin Gross

The concept of cell type has recently become problematic among biologists, as evidenced by a series of programmatic articles proposing new ways to classify cells. Often, these articles call for a purely "data-driven" approach to classification, implying that traditional classifications are inadequate due to their reliance on background knowledge and assumptions. New technologies, such as single-cell sequencing combined with sophisticated computational pipelines, are proposed to provide the right tools for such "theory-free" classifications. From a philosophical point of view, however, it seems obvious that there are fundamental reasons that argue against the possibility of purely theory-free classification. On the one hand, novel technologies are always situated within a history of epistemic iteration and thus adopt background assumptions from earlier methods. On the other hand, the generation and processing of sequencing data involves many decisions, such as the choice of specific methods or parameters, which must be based on background knowledge if they are not to be arbitrary. To argue against the possibility of theory-free classifications is important, but not necessarily very illuminating. Therefore, beyond arguing against the plausibility of the "theory-free ideal" for cell classification, I consider the weaker version of a "theory-minimizing ideal." Does it make sense for biologists to strive to reduce background assumptions? Even if it is not possible to eliminate them completely, biologists could take a path that increasingly reduces dependence on background assumptions. Would this be desirable in view of the goals of the discipline, or would it be wiser to work on consolidating the background knowledge on which existing classifications are based?

From Microscopy to Single-Cell Sequencing: Epistemic Cultures and Multi-Modal Data Integration in Cell-Type Classification
Meike Wiegand

Genomics-based methods often require deference to external protocols (such as imaging and functional assays) for validation. Single-cell sequencing technologies have promised to overcome such constraints by profiling cells within "theory-free" or "theory-reduced" frameworks, offering insight into the organization of cells by reducing cellular identity to molecular information through transcriptome analysis. However, this paper locates the use of single-cell technologies in the life sciences at the intersection of two research traditions, biological and computational, and argues that the different technologies operate within their own epistemic cultures. While single-cell and next-generation sequencing technologies rely heavily on computational analysis pipelines, more traditional methods like microscopy remain relevant—a situation that has resulted in the need for data integration across modalities. Building on existing research in the philosophy of experiment, this paper presents a methodological framework appropriate for addressing the problem of multi-modality, while remaining sensitive to distinct epistemic backgrounds. By comparing protocols and data practices in light microscopy and single-cell RNA sequencing, the paper first traces the

history of their underlying experimental systems. Then, focusing on cell-type classification as a case study, the paper elucidates how data from distinct research traditions are integrated in biologically meaningful ways through an analysis of selected scientific papers. Abstracting from there, the paper aims to provide a broader framework for addressing co-existing scientific methods stemming from different research traditions. Overall, the paper has consequences for how we think about the integration of data across scientific practices.

The Human Cell Atlas project as a case study: Are foundational data resources still relevant for the biomedical community?
Ignazio Arrabito

The Human Cell Atlas (HCA) project is one of the largest-scale biology endeavors of the century. By implementing single-cell technologies, it intends to produce a database providing information on the cells of each tissue and organ of the human organism, with the aim of serving as a reference for the research community. Such an ambitious enterprise is not the first of its kind, as both the Human Genome Project (HGP) and ENCODE aimed to act as a foundation for biology. Here, I conduct a comparative study of HCA in the context of its relationship with the latter two. I highlight their common core idea of "completeness" and argue that the same pitfalls plaguing HGP and ENCODE might impede HCA's overall aim. I focus on their objective to create an open-access database that should serve as a resource for the whole community. At the time of the HGP, independent labs were not able to generate their in-house databases and needed the resources offered by big consortia. This is not the case in today's research landscape, which casts doubts on the promoted utility of the HCA database for the entire biomedical community. Closing on a positive note, I assess other possible contributions of the project in terms of the social infrastructures it generates as a community-driven endeavour.

Why small science lives on: The case of cell atlas makers
Patrick Ferree

A new wave of large-scale biology projects, including the Human Cell Atlas consortium, has started to attract the attention of philosophers of science—and it should. The thought of profiling all of the cells in the human body raises important philosophical questions (e.g., what is the point?). But so too does the continued use and production of cell atlases by smaller labs. Why do smaller labs and collaborations continue building and using their own cell atlases when there are plenty of "comprehensive" reference atlases on the way? The goal of this paper is to explain their motivations and analyze how different epistemic aims and values interact and influence these practices. In doing so, I connect with recent research on the relationships between "users" and "producers" and the various epistemic goals that drive mapping campaigns. Drawing on multi-sited ethnographic fieldwork (15 months) in the single-cell omics community and prolonged engagement with a particular research group studying the regulation of energy homeostasis in the brain, I argue that "local" epistemic needs continue to play a significant role in the shapes and qualities of cell atlases. As my primary case study, I focus on the development of three cell atlases, all from the same group, but each differing in content (source of biological tissue), method (molecular measurements), and purpose (representation and intervention). The first atlas is comparative and covers a small region of brain tissue in mice, rats, and macaques. The second atlas integrates spatial features of brain anatomy with molecular profiles of cells and uses tracer technology to map neural projections. The third atlas combines single-cell transcriptomics with chromatin

accessibility in human brain tissue. By analyzing the production of these three atlases alongside one another, I show how certain epistemic values (e.g., completeness) and non-epistemic values (e.g., utility) interact and regulate the process. Overall, this work contributes to ongoing discussions about the relationship between big and small science.

Scientific Consensus in the Wild: The Methodological, Epistemic, and Political Foundations of the Institute for Ascertaining Scientific Consensus (IASC)

Jesse Hamilton
University of Oxford
jesse.hamilton@bsg.ox.ac.uk

Peter Vickers
Durham University
peter.vickers@durham.ac.uk

Rory Jubber
University College London
rory.jubber@ucl.ac.uk

Owen Towler
University College London
owen.towler.24@ucl.ac.uk

Haixin Dang
Case Western Reserve University
haixin.dang@case.edu

Scientific consensus is routinely invoked in public discourse and policy justification, yet until recently, no institution existed to systematically measure it. This absence is striking: science possesses elaborate infrastructures for collecting data, evaluating evidence, and distributing credit, but nothing for capturing the distribution of expert judgment itself. To address this gap, the Institute for Ascertaining Scientific Consensus (IASC) was established in 2022 and piloted a novel, scalable method for surveying global scientific communities using institution-verified contact lists and single-use tokens, demonstrating that systematic consensus measurement is both feasible and repeatable. With feasibility established, central questions now arise concerning IASC's methodology, the political implications of institutionalizing this type of expert opinion, and how its structure distinguishes it from Collins and Evans' Owls.

This symposium brings together some of IASC's members to critically evaluate the project's aims, methods, and implications. Peter Vickers opens by outlining IASC's survey architecture and pilot results, clarifying what the surveys can and cannot capture. Two subsequent talks offer critical reflections on the epistemic and political consequences of measuring consensus. The symposium concludes with critical commentary that highlights the project's strengths, limitations, and future directions. SPSP provides an ideal forum for this work. By treating the measurement of consensus as an evolving institutional practice, this symposium aligns with SPSP's mission to advance an empirically informed, socially engaged, and critically reflective philosophy of science.

Peter Vickers introduces and evaluates IASC's methodological architecture. Drawing on pilot studies, he shows that consensus measurement is not simply 'counting votes' but requires careful attention to survey design, question framing, and response interpretation. Systematic surveys illuminate areas of agreement alongside zones of uncertainty and disagreement.

Owen Towler and Jesse Hamilton examine IASC's political and normative implications. Using HIV/AIDS denialism in South Africa and 'following the science' rhetoric in the UK's

COVID-19 response, they argue that IASC strengthens democratic accountability by making expert opinion transparent, preserving the distinction between epistemic input and political authority.
Rory Jubber evaluates IASC against Collins and Evans's 'Owls,' arguing that IASC's non-statutory, external, methodologically explicit, and descriptive character distinguishes it from earlier proposals and avoids key pitfalls of elitism and technocracy.'
Haixin Dang (Case Western Reserve University) will conclude the symposium with critical comments on the IASC project, its methodology, its challenges, and its future directions.
Collectively, the symposium makes a unified case: IASC is theoretically defensible, methodologically feasible, and practically significant. It offers a way to move beyond anecdotal claims about "what scientists think" by providing systematic, reproducible measurements that can inform public debate without collapsing into technocracy. As scientific issues grow increasingly complex and politically charged, the capacity to carefully and transparently assess expert scientific opinion becomes an essential component of responsible, democratic governance.

Towards an Institute for Ascertaining Scientific Consensus: Lessons from Three Global Pilot Studies
Peter Vickers, University of Durham

This talk introduces and assesses the emerging idea of an Institute for Ascertaining Scientific Consensus (IASC). The starting point is a simple puzzle: we often place great weight on scientific consensus in public discourse and policy, yet there is no standing body tasked with systematically measuring and reporting the level of agreement within scientific communities. Instead, consensus figures tend to arise ad-hoc, from small teams and highly variable methods. The IASC concept proposes a lightweight, transparent, and repeatable way to capture large-scale snapshots of scientific opinion over time, descriptive rather than prescriptive, and independent of government.
To date, three pilot studies have tested this vision. The first established feasibility, developing a rapid, low-burden survey method using institution-verified email contact lists and single-use tokens to ensure trust, anonymity, and high response rates. This proof-of-concept work demonstrated that it is possible to survey thousands of researchers across many institutions with response rates far above typical academic survey benchmarks.
The second pilot, conducted within the global astrobiology community and published in Nature Astronomy, examined expert judgement regarding the likelihood of extraterrestrial life. It revealed strong majority agreement on the existence of basic life elsewhere in the Universe, alongside far higher uncertainty for more ambitious claims. Crucially, the study illustrated how methodological 'decision points' (especially the presence of a neutral response option) shape consensus figures. Neutrality emerged not merely as an expression of middling confidence, but also as a strategic stance under evidential uncertainty, reinforcing the need for transparency in reporting full response distributions.
The third pilot applied the method in real time to two major 2025 announcements of possible extraterrestrial life: first on exoplanet K2-18b, and later in the Cheyava Falls rock on Mars. The results showed markedly higher confidence in the Martian case, with a significant softening of strong disagreement compared to the exoplanet claim. Yet respondents generally remained cautious about crossing from neutrality into agreement, a pattern consistent with well-known scientific reticence in frontier contexts. This dynamic supports the value of repeated, comparable surveys: they can reveal how expert confidence shifts (or does not shift) as evidence accumulates.

Across all three pilots, a central theme emerges: consensus measurement is not a matter of simply 'counting votes'. Rather, it requires methodological clarity, consistent practice, and careful framing of questions at an appropriate level of abstraction. Extremely narrow, highly technical claims risk excluding all but a handful of specialists; extremely broad statements risk irrelevance. The pilots suggest a productive middle ground, where questions remain meaningfully tied to real-world issues whilst still drawing on a sufficiently large and informed respondent base.

Together, these early studies show that systematic, global, repeatable measurement of scientific opinion is feasible, and that doing so can illuminate both areas of strong agreement and areas of caution or emerging disagreement. The talk will outline this progress, discuss remaining philosophical and practical challenges, and consider what a future IASC could be.

Between Deference and Denial: How Scientific Consensus Mapping Democratizes Public Policy

Owen Towler, University College London
Jesse Hamilton, University of Oxford

In the late 1990s and early 2000s, South African President Thabo Mbeki rejected the overwhelming international scientific consensus that HIV causes AIDS, instead endorsing heterodox views that framed the disease as a consequence of poverty or drug toxicity (Boseley, 2008). The resulting delay in antiretroviral treatment led to hundreds of thousands of preventable deaths. Two decades later, during the COVID-19 pandemic, the UK government repeatedly justified policy decisions by appealing to "the science," often treating advice from small expert advisory bodies as standing in for Science as such, despite significant uncertainty and disagreement within the scientific community (Winsberg et al. 2020). Although these cases differ sharply in their political contexts and failure modes, they share a common pathology: political authority was exercised by misrepresenting scientific agreement, denied in South Africa and overstated in the UK.

Can an institution like IASC solve this problem? Such institutions raise familiar concerns. Within STS, scholars worry about technocracy and elitism (Collins and Evans 2002; 2003; 2007), while political philosophers fear the displacement of political judgment by those wielding epistemic authority (Estlund 2008). These critiques suggest a dilemma: when political actors defer to claims of scientific consensus without examining the breadth of actual agreement, they risk lending unwarranted authority to limited expertise; when they ignore or downplay scientific disagreement, they risk substituting their own judgment without constraint. Yet this dilemma rests on a false choice between treating scientific consensus as decision-making authority and disregarding it entirely. A third institutional approach—one that makes scientific agreement and disagreement transparent to the public without claiming authority over political decisions—can escape this trap.

Democratic legitimacy in science-informed policymaking depends on institutionally preserving a visible distinction between epistemic input and political authority. Institutions that systematically map scientific agreement and disagreement, like IASC, can strengthen democratic accountability by making transparent where consensus exists and where it does not. This prevents policymakers from hiding behind false claims about what the science says, forcing them to explicitly own their political choices.

The two cases serve as counterfactual tests of IASC's democratic value. In the UK COVID-19 response, an IASC-style mapping could have constrained technocratic rhetoric by making uncertainty publicly known and preventing policymakers from deflecting responsibility by claiming to "follow the science." In South Africa, public evidence of an overwhelming global consensus on HIV/AIDS would have raised the justificatory burden on political leaders, like

Mbeki, who rejected it. Both failures, we argue, reveal an institutional gap that IASC is designed to fill.
The paper defends IASC as a democratizing epistemic institution that forces transparent justification of political choices by making visible the actual state of scientific community opinion, offering both a normative account of epistemic legitimacy and a practical model for science-informed policymaking without technocratic overreach.

IASC for the birds?
Rory Jubber, University College London

Amongst the STS and socially informed philosophy of science community, IASC has faced a common objection. Namely, that IASC closely resembles aspects of the work of Harry Collins on expertise, scientifically informed policymaking, and scientific consensus. In particular, it is noted that IASC bears a very close resemblance to the Owls project envisioned by Collins and Evans (2017). As such it has been taken that problems which the Owls face will be inherited by IASC, most notably that the Owls and IASC present an elitist vision of the role of scientific expertise in democratic societies. This talk will set out to show that the respects in which IASC differs from the Owls means that the problems faced by the Owls are not faced by IASC.
To achieve this, I will first set out what Collins and Evans (2017) envision the Owls to be, namely, a body of scientific and sociological experts that advises governments on policy making on issues and crises that require insights from technical scientific domains. As Collins and Evans (2017, 86) argue, the job of the Owls would be to examine the current state of expert knowledge in domains relevant to policymaking and then pass those conclusions on to policymakers and politicians to inform their decision processes. This would, critically in the comparison to IASC, involve reporting on the state of consensus in the relevant scientific fields and domains, a task which, interestingly, Collins and Evans take to be a relatively easy one (2017, 86-90). In pursuit of showing the differences between the Owls and IASC, I will spend some time in the talk articulating what Collins and Evans (2017) take to be the purpose and composition of the Owls.
With a sense of the Owls in place, I will move on to the points of divergence from IASC. Firstly, and most critically, I will show that the Owls, if brought into being, would be a statutory body of the state. IASC would have no statutory place within the state. It would be external to the state as a part of the wider polity. This has important implications in debates around expertise, technocracy and democracy and allows IASC to mitigate against accusations of epistemic elitism. Secondly, we do not agree that the task of ascertaining consensus or scientific opinion on a given matter is as easy as Collins and Evans (2017,86) take it to be. While they do not explicitly state a method for attempting to do that, IASC does, as Vickers et al (2022,2024) have shown. Thirdly, unlike with the Collins and Evan's, IASC does not require something like a categorisation of the personality types of individual scientists. Finally, we are far more cautious about how we envision our findings on consensus to inform policymaking. As such, we take IASC to be a body more interested in describing the state of scientific opinion on given scientific matters and less interested in the prescriptive task of informing policymakers on the correct course of action.

Public, Democratic, Amateur: Different Modes of Doing Science [double symposium]

Elis Jones
Technical University Munich
elis.jones@tum.de

Sophie Veigl
University of Vienna
sophie.juliane.veigl@univie.ac.at

Rena Alcalay
Technical University Munich
rena.alcalay@tum.de

Evan Josselin
Institute Jean Nicod
evan.josselin.sorbonne@gmail.com

Richard Williams
Technical University Munich
richard.williams@tum.de

Anna Alexandrova (commentator)

Various labels have been attached to the word 'science' when it is conducted with non-scientists, each invoking a range of positive and negative aspersions: 'public' or 'citizen' science might invoke an active (but also superficial) role for non-scientists; 'democratic' science the idea of exposing science to political oversight (for better or for worse); 'amateur' science conjures images of (at best) unpaid science or (at worst) science which is undertaken poorly.

These labels point to interesting questions: how can we produce reliable and applicable knowledge given the diversity of human interests and perspectives? How can groups typically considered to be part of the 'general public' be involved in different stages of the scientific process? How are some groups excluded from these processes, or when might they choose to operate outside of them? Who counts as an expert and how are these different forms of expertise (and ignorance) choreographed?

This double session will address these questions, looking both at science broadly and at individual cases of knowledge production specifically, exploring issues arising when different perspectives and forms of expertise come into contact with one another. Speakers will discuss: how scientists make judgements of 'feasibility', and what happens when these come into conflict with value judgements; the blurry boundaries of expertise in the experiences of medical patients; the production of biomedical knowledge by people actively choosing to work outside of scientific institutions; how citizen science can be used to democratise different aspects of scientific research; and how the perspectives of diverse groups of people, and of non-human organisms, can be used to develop reliable measurements of ecosystem health.

The session will take place in two parts. Session one features three short talks, time for individual questions for each speaker, and remarks from an invited commentator. Session two

includes two longer talks, remarks from a second invited commenter, and semi-structured discussion of overarching themes, issues, and questions (from both sessions).

The Feasibility Power of Experts
Richard Williams, Technical University Munich

In the philosophy of science, the values in science debate largely explores how the value judgments of experts may influence politics in unacceptable ways. In contrast, I will foreground the underexplored "feasibility-power" of experts. In politics, experts often make feasibility judgments about what is possible and what is necessary. The feasibility-power of experts can serve the public. Once the public has decided on what basic aims to pursue, the experts can decide how to make such aims possible. However, the feasibility-power of experts can also dominate the public. The public may decide an aim is good, but the experts decide that the aim is impossible. Alternatively, the public may decide an aim is bad, but the experts decide that the aim is necessary.
In practice, methodological choices rather than moral choices often shape the feasibility judgments of experts. One methodology may show that an aim is possible/impossible/necessary, and another methodology may show the opposite. So, the public risks the social harm of fallible feasibility judgments. Firstly, the experts may wrongly decide that a good but impossible aim is possible. So, the public pays the price of pursuing the impossible aim, but the public will never gain the benefits of achieving it. Secondly, the experts may wrongly decide that a good and possible aim is impossible. So, the public foregoes the benefits of achieving the good aim. Thirdly, the experts may wrongly decide that a bad and unnecessary aim is necessary. So, the public unnecessarily undergoes the costs of the bad aim.
I will argue that methodological pluralism is a critical check on the fallible feasibility judgments of experts. However, I will argue that no particular individual or institution must practice methodological pluralism. In practice, individuals inside institutions may need consensus on methodological questions to collaborate effectively. On a much bigger scale, I will argue that the research ecosystem as a whole should cultivate methodological pluralism. In practice, different institutions should seek contestation on methodological questions as a critical check on each other. The epistemic benefit of contestation is more reliable feasibility judgments. Contestation helps to eliminate errors. The social benefit of contestation is more socially responsible feasibility judgments. Contestation helps to protect the public from errors. When contestation does not protect the public, contestation protects the trustworthiness of experts as responsive to their own fallibility and the social harm they may cause as a result.

Creeping Doubt of Chronicity
Rena Alcalay, Technical University Munich

Biomedical data practices increasingly rely on quantification and predictive modeling, often sidelining the lived, fluctuating experiences of chronically ill patients. This talk explores how such practices generate epistemic harm through survivor bias and the exclusion of bodily doubt. Drawing on auto-ethnographic reflection, I examine how chronic illness disrupts bodily certainty and exposes the limitations of biomedical epistemologies that privilege measurable data over subjective testimony.
A key example lies in the foundational challenges of establishing baseline measures. In endocrine diagnostic guidelines, for example, baseline thyroid function is typically assessed

using measures of thyroid-stimulating hormone (TSH), thyroxine (T4), and triiodothyronine (T3). Though these hormone levels are quantitatively measured within established normal reference ranges—such as TSH ranging approximately 0.4 and 4.0 mIU/L and free T4 about 0.8 to 1.8 ng/dL—there remain significant diagnostic uncertainties. Symptoms frequently fail to align with these clinical baselines, leading to profound mismatches between biomedical markers and patient experience. This epistemic uncertainty highlights how, even with quantitatively precise biological markers, normative frameworks and diagnostic protocols influence interpretation and treatment decisions. Baseline diversity estimates can vary dramatically and lack convergence, illustrating complex interactions between data, measurement, and human judgment.

I argue that survivor bias in predictive systems constitutes a form of epistemic disadvantage (Goldstein 2022), systematically excluding symptoms that resist population-level categorization. Endocrinology struggles with persistent uncertainties about hormone secretion, disease triggers, and autoimmune causation. A key issue is the disjunction between biomedical markers and lived illness, which fosters epistemic disadvantage in clinical practice by sidelining patient uncertainty (i.e., bodily doubt) in favor of rigid diagnostic certainty. When biases are introduced at the level of clinical data measurement, they are magnified by institutional structures that prioritize standardization and algorithmic prediction, further distorting both nosological practices and patient agency. Survivor bias in biomedical predictive models thereby perpetuate a feedback loop in which doubt itself becomes pathologized rather than acknowledged as epistemically legitimate.

The broader question emerges: in biomedicine's data practices and institutional structures, when and how do human variables introduce biases that compromise scientific objectivity, and when is their role indispensable for producing responsible, reliable, and actionable metrics? These questions are urgent in endocrinology and metabolic clinical practices, where emerging data systems and automated diagnostic protocols risk deepening epistemic disadvantage and injustice. The exclusion of patient experience in this biomedical field not only undermines patient agency but also perpetuates bodily doubt—an uncertainty sustained by the absence of data and the fragility of clinical interpretation.

Building on Havi Carel's (2013; 2015) concept of bodily doubt and Margaret Price's (2015) bodymind lens, I explore how chronic illness challenges assumptions embedded in biomedical data ethics. I call for more integrative, interdisciplinary approaches to chronic illness, and specifically autoimmune, care. This work reconsiders the institutional foundations of data systems and design of clinical standards advocating for models that genuinely account for uncertainty, patient testimony, and the ethical complexity of lived experience.

References

Carel, Havi. 2013. "Bodily Doubt." Journal of Consciousness Studies 20 (7–8): 178–97.
Carel, Havi. 2025. "A Taste of Your Own Medicine: The Experience of a Transplant." The Lancet 405 (10486): 1220–21. https://doi.org/10.1016/s0140-6736(25)00691-9.
Goldstein, Rena Beatrice. 2022. "Epistemic Disadvantage." Philosophia 50 (January):861–1878. https://doi.org/10.1007/s11406-021-00465-w.
Price, Margaret. 2015. "The Bodymind Problem and the Possibilities of Pain." Hypatia 30 (1): 268–84. https://doi.org/10.1111/hypa.12127.

The democratic promise of citizen science: Clarifying an ambiguous concept
Evan Josselin, Institute Jean Nicod

Since what Jasanoff (2004) has termed the "participatory turn," science has increasingly shifted toward participatory models of governance, moving from a public perceived as merely seeking to understand scientific knowledge to one invited to engage directly in its production. This shift has fueled the idea that science should be "democratized." One policy tool frequently advanced to achieve this democratization is the promotion of citizen science projects, which are often presented as remedies to the perceived decline of trust in science within democratic societies and as a means to restore the legitimacy of scientific research (Mattei, 2023). As Strasser et al. (2019) note, citizen science carries strong promises: greater democratization of science, improved scientific literacy, and even new scientific breakthroughs. Yet despite these claims, the notion that citizen science democratizes science remains opaque. Although the democratizing potential of citizen science is frequently asserted, it has rarely been subjected to systematic philosophical scrutiny, and the concept of "democratization of science" is used in multiple and may be incompatible ways.
In this presentation, I assess this widespread assumption and examine the different senses in which citizen science is said to democratize science, in order to clarify the visions of science and democracy that underpin them. To do so, I draw on John Dewey's (1916) framework and his articulation of the intimate relationship between democracy and education, two dimensions that are deeply intertwined in citizen science practices. I develop this analysis through two case studies: (1) Spipoll, a French project in which participants photograph insects on flowers to document plant–insect interactions; and (2) a collaborative project developed by Les Ateliers des Jours à Venir (a third-sector research association), oceanographers, and residents of a precarious neighborhood in Almada (Portugal), aimed at understanding why winter waves regularly flood the area (knowledge that became crucial when the municipality considered demolishing the neighborhood). Through these examples, I identify three distinct modalities of the democratization of science: 1) Democratizing scientific content — enabling participants to understand scientific methods and the role of scientists, which entails a view of democratization as showing the legitimacy of science and scientists, 2) Democratizing inquiry processes — opening both the production and the use of scientific knowledge to broader publics in the idea that participants in citizen science projects can become full-fledged epistemic agents, and 3) Democratizing scientific institutions — granting citizens a role in shaping research priorities, methods, and socio-political applications. These three modalities, I argue, reveal different and may be conflicting conceptions of what it means for science to be democratic.

References:

Dewey, J. (1916). Democracy and education. Macmillan.
Jasanoff, S. (Ed.). (2004). States of knowledge. Abingdon, UK: Taylor & Francis.
Mattei, P. (2023). Democratizing Science: The Political Roots of the Public Engagement Agenda (p. 141). Bristol University Press.
Strasser, B., Baudry, J., Mahr, D., Sanchez, G., & Tancoigne, E. (2019). "Citizen science"? Rethinking science and public participation. Science & Technology Studies, 32(2), 52-76.

Use-based experts are turning away from science - but where are they going?
Sophie Veigl, University of Vienna

Steroid hormones such as testosterone, estrogen, and progesterone are extensively studied substances in biomedical research and clinical practice. Yet their uses extend well beyond the institutional boundaries of biomedicine. In diverse contexts, ranging from bodybuilding communities optimizing muscle growth and recovery to DIY trans* communities navigating gender-affirming hormone use, individuals administer steroid hormones without direct biomedical oversight. These "extra-biomedical" practices often emerge not from a wholesale rejection of scientific knowledge, but from highly specific frictions with biomedical knowledge systems. Many bodybuilders experience biomedical practitioners as lacking situated understanding of performance-oriented hormone regimes, while many trans* people report discriminatory treatment, gatekeeping, or insufficient access to gender-affirming care. As a result, users may "turn away" from institutionalized science and develop community-specific epistemic practices to assess, share, and mitigate biomedical risks.
This paper investigates these communities as use-based knowledge formations that challenge conventional philosophical assumptions about expertise, epistemic authority, and the boundaries of legitimate knowledge production. Their practices raise several interconnected questions: To what extent should these communities be considered experts, given their extensive experiential knowledge, collective troubleshooting, and context-sensitive risk management? How should philosophers of science conceptualize epistemic authority when biomedical knowledge is experienced as insufficiently responsive to the concrete needs and contexts of users? And what responsibilities do biomedical practitioners bear in addressing the resulting fragmentation of knowledge practices?
Drawing on examples from bodybuilding steroid users and DIY trans* hormone users, I argue that these communities do not straightforwardly reject biomedical science but rather selectively disengage from it in response to epistemic, social, and structural failures. Their practices illuminate the limitations of a monolithic view of biomedical expertise and instead support a pluralistic, use-based account of expertise grounded in the situated demands of hormone use. I further explore the prospects and challenges of integrating extra-biomedical knowledge practices with institutional biomedical frameworks. Such integration, while potentially desirable for reducing harm and enhancing mutual understanding, is complicated by mismatches in epistemic norms, institutional incentives, and the social position of marginalized users.
Ultimately, this talk situates steroid-using communities as an important site for rethinking expertise and epistemic authority in contemporary philosophy of science. By examining how and why communities turn away from institutionalized biomedicine—and how they build alternative epistemic infrastructures—we gain insight into the dynamics of epistemic fragmentation and the possibilities for more responsive, context-sensitive forms of biomedical knowledge practice.

Measuring ecosystem health
Elis Jones, Public Science Lab, Technical University Munich

Scientists regularly characterise the condition of ecosystems, more commonly in terms of 'degradation' but also regularly in terms of 'health'. Much has been written debating whether ecosystems can literally be healthy or not. I first briefly offer some arguments in favour of retaining the concept of ecosystem health (and related attempts to characterise ecosystem condition more generally). I then highlight some issues shared between attempts to defining

health in humans and ecosystems, and use these to help explore the following question: assuming ecosystems can be (un)healthy, how can this be measured?
To answer this question, I draw on literature from a few different areas: combined naturalist and constructivist accounts of health from philosophy of medicine (Conley & Glackin); philosophical investigations into wellbeing measurements (Alexandrova); work from philosophy of ecology which treats ecosystems as 'indexical' (Lean); and the notion of indexicality more broadly from the philosophy of language.
Using these ideas, I present a 'mixed methods' approach to conceptualising and measuring ecosystem health, which capitalises on the perspectival nature of the concept. Here, a feature of ecosystem health which is regularly presented as a problem - namely, that outcome of ecosystem health assessments is heavily dependent on methodological choice and underlying value judgements - becomes a useful and attractive feature for allowing the formation of scientific concepts with input from those outside of science, including consideration of the interests of non-human organisms.
I offer a three-stage process which can produce informative, reliable and relevant characterisations of changes to ecosystems:
1. First, a qualitative stage whereby the perspectives of people who interact with the ecosystem in question are explored, primarily with the aim of distinguishing important species, processes, and other matters of concern when assessing the condition of that ecosystem.
2. Second, further exploration of that ecosystem in light of the results of the first step of the process, in particular focusing on the features of the ecosystem raised as particularly valuable. At this stage, the ways in which the relevant living systems interact with their environment can be factored in, allowing further space for important features of the ecosystem to emerge.
3. Finally, a third stage which looks for fundamental ecological processes underlying the valuable features highlighted in stages one and two, which offers potential for broader indicators of changes widely agreed to be negative or positive in a specific case.
This process can be iterated to incorporate further perspectives or respond to social and ecological change. The appropriate boundaries of the ecosystem itself may also be revised throughout. Whilst not without problems, this approach offers the possibility of assessment and characterisation of ecosystem health whilst avoiding the traps of absolutism and nihilism. Importantly, it allows for both more democratic and less anthropocentric scientific activity, recognising various forms of expertise, and allowing for the intertwined nature of fact and value.

Making data speak: Philosophy of Archaeology in Practice

Mathilde Lequin
PACEA, University of Bordeaux
mathilde.lequin@u-bordeaux.fr

Adeline Le Cabec [has co-author]
PACEA, University of Bordeaux
adeline.le-cabec@u-bordeaux.fr

Andra Meneganzin
Institute of Philosophy, Centre for Logic and Philosophy of Science (CLPS), KU Leuven
andra.meneganzin@kuleuven.be

Anton Killin
Department of Philosophy, Bielefeld University
anton.killin@uni-bielefeld.de

Ross Pain
Department of Philosophy, University of Bristol
ross.pain@bristol.ac.uk

Chiara Pertile
PACEA, University of Bordeaux
chiara.pertile@u-bordeaux.fr

Solange Rigaud
PACEA, University of Bordeaux
solange.rigaud@u-bordeaux.fr

Archaeologists face multiple epistemic challenges. These include both a scarcity of primary data, such as fossils and artifacts (Currie 2021), but also issues concerning the interpretation of that data. "Data don't speak for themselves" is a common mantra in all data-based sciences—and nowhere is this more evident than in light of the epistemic situation and resources that characterize archaeological practice (Chapman and Wylie 2016). Between the data and the evidential claims the data are intended to support, there often lies a vast inferential space. Within this space, uncertainty can be generated, amplified, or reduced through successive inferential steps. A central task of archaeological research is therefore to turn material traces into evidence about past hominin features, abilities or lifeways. This interpretive work requires a combination of theoretical resources—ranging from overarching frameworks to field-specific mid-range theories, and including fundamental background knowledge—and technical skills, involving diverse research communities, institutions, and infrastructures. For data to serve as evidence, archaeologists often need to generate multiple data from limited traces and integrate external sources of data to assess and constrain interpretive possibilities.
Given the diversity of strategies involved in developing evidential arguments about the deep human past, the inferential space between data and evidential claims has become a rich site of investigation for philosophers. Practicing philosophy with and in archaeology (Pradeu et al. 2024) thus involves analytic work aimed at disentangling the threads that support

archaeological inferences. But it also engages in a constructive approach that attends to the specific epistemic challenges archaeologists encounter, in order to offer original accounts or heuristics to advance research in productive, new directions (Currie et al. 2024).
This session brings together philosophers (Killin, Lequin, Meneganzin, Pain, Pertile) and archaeologists (Le Cabec, Rigaud) to explore topical philosophical perspectives on the diverse ways archaeologists generate, interpret, and mobilize data. The contributions range from analyses of the role of digital data in archaeology (Lequin and Le Cabec) and methods for detecting signals of innovativeness in the Neanderthal archaeological record (Meneganzin), to the integration of ethological and archaeological data for understanding the evolution of normative behaviour (Killin and Pain), and to the construction of a comparative framework that organizes ecological, social, cognitive, and material variables to improve interpretations of early hominin tool-use (Pertile, Lequin, and Rigaud).
While addressing a range of methodological, epistemic, and conceptual challenges archaeologists encounter in working with data, these contributions also highlight several transversal perspectives: the relationship between technical enhancements in data acquisition and the strength of the claims derived from those data; the integration of diverse types of data; and the potential and limits of cross-species comparisons for understanding hominin evolution. Together, the speakers show how philosophical engagement can deepen our understanding of archaeological practices and strengthen the evidential arguments developed from these data.

References

Chapman, R., Wylie, A. (2016). Evidential reasoning in archaeology. London: Bloomsbury.
Currie, A. (2021). Stepping forwards by looking back: underdetermination, epistemic scarcity and legacy data. Perspectives on Science, 29(1), 104-132.
Currie, A., Killin, A., Lequin, M., Meneganzin, A., & Pain, R. (2024). Past materials, past minds: The philosophy of cognitive paleoanthropology. Philosophy Compass, 19(6), e13001.
Pradeu, T., Lemoine, M., Khelfaoui, M., & Gingras, Y. (2024). Philosophy in science: Can philosophers of science permeate through science and produce scientific knowledge?. The British Journal for the Philosophy of Science, 75(2), 375-416.

Investigating the Value of Digital Data in Archaeology

Mathilde Lequin, Adeline Le Cabec
Univ. Bordeaux, CNRS, Ministère de la Culture, PACEA, UMR 5199, F-33600 Pessac, France

Archaeologists work not only with the material traces of the deep past—such as fossils and artifacts—but also with the data generated from these traces. Increasingly, these outputs take the form of digital data: computational representations of specimens, materials, and sites that encode information about their properties, structure, and relations. By combining 3D imaging and mathematical modelling, this "digital revolution" (Weber 2011) makes it possible to reconstruct missing parts, correct distortions, examine internal structures, and perform virtual analyses that would otherwise be impossible. Yet unlike the "radiocarbon revolution" (Chapman & Wylie 2016; Wylie 2020), the digital turn has received little philosophical scrutiny (but see Wylie 2021), despite its deep integration into archaeological practice.
This talk examines the philosophical challenges raised by digitization, by investigating the nature and epistemic value of digital data in archaeology. While archaeological data are rare, non-renewable, fragile, and demanding—exemplifying a situation of "epistemic scarcity"

(Currie 2021)—digital data are abundant, indefinitely generable, dematerialized, and easily accessible. However, their epistemic value cannot be understood merely in terms of how they compensate for the limitations of physical objects. Drawing on discussions about relational (Leonelli 2016; Wylie 2020) and pragmatic-representational (Bokulich & Parker 2021) views of data, we identify three characteristics of digital data that are crucial for understanding their distinctive role in shaping archaeological knowledge.

First, the epistemic value of digital data extends beyond what they enable archaeologists to observe and measure. Their significance lies in the evidential roles they can acquire, resulting from methodological choices, anchoring arguments, and ultimately supporting archaeological claims. Understanding these evidential functions requires analysing digital data as situated products of specific practices, raising new perspectives on how and when fossils become data (Watkins 2024).

Second, digital data circulate. Archaeologists routinely store, share, and reuse data produced by others. Despite being less materially demanding than physical specimens, digital data raise pressing issues concerning ownership, curation, and reliability. Using the framework of "data journeys" (Leonelli & Tempini 2020), we show how the ways digital data are documented, preserved, and repurposed—through "packaging strategies" (Leonelli 2016) and efforts to ensure "traceability" (Wylie 2020)—can expand or constrain their epistemic value.

Third, digital data are not just dematerialized objects. A relational approach requires considering how they relate both to the physical specimens from which they derive and to the other data types generated from those specimens. This includes asking whether digitization, often assumed to be non-destructive, truly preserves the integrity of specimens or risks foreclosing future analyses (e.g., the recovery of ancient DNA data), and whether digital representations of human remains should themselves be treated as sensitive materials (Le Cabec & Toussaint 2017). Far from dematerializing archaeology, digital data return us—via new pathways—to the materiality of archaeological research objects.

By analysing the relational status of digital data across their production, circulation, and integration, this talk contributes to a more nuanced understanding of archaeological evidence and to strategies for strengthening what is claimed through digital data.

Bokulich, Alisa, and Wendy Parker. (2021). Data Models, Representation, and Adequacy-for-Purpose. European Journal for Philosophy of Science 11:31.

Chapman R, Wylie A (2016) Evidential reasoning in archaeology. London: Bloomsbury.

Currie, A. (2021). Stepping forwards by looking back: underdetermination, epistemic scarcity and legacy data. Perspectives on Science, 29(1), 104-132.

Le Cabec, A., & Toussaint, M. (2017). Impacts of curatorial and research practices on the preservation of fossil hominid remains. JASs. Journal of anthropological sciences, 95, 7-34.

Leonelli, S. (2016). Data-Centric Biology: A Philosophical Study. Chicago: University of Chicago Press.

Leonelli, S., & Tempini, N. (2020). Data journeys in the sciences. Cham: Springer.

Watkins, A. (2024). When Are Fossils Data?. Philosophy, Theory, and Practice in Biology, 16(2).

Weber, G. W., & Bookstein, F. L. (2011). Virtual anthropology: a guide to a new interdisciplinary field (p. 423). Wien: Springer.

Wylie, A. (2020). Radiocarbon dating in archaeology: triangulation and traceability. In Data Journeys in the Sciences (pp. 285-301). Cham: Springer.

Wylie, C. D. (2021). Preparing Dinosaurs: The Work Behind the Scenes. Cambridge, MA: MIT Press.

Taming Comparative Innovativeness In the Middle Pleistocene

Andra Meneganzin
Institute of Philosophy, Centre for Logic and Philosophy of Science (CLPS), KU Leuven

The behavioral repertoire of Neanderthals, along with the Mousterian lithic technocomplex, has traditionally been characterized by notable "conservatism" (i.e., "continuity" or "persistence") when compared to the quasi-contemporaneous African Middle Stone Age (300-30 Kya) and later Upper Paleolithic hominins (Mellars 1998; Kuhn and Stiner 1998). Yet, it seems unlikely that a species with such a vast geographic range, from the Iberian Peninsula to Siberia, could persist for hundreds of thousands of years with such behavioral inflexibility. Recent evolutionary approaches are offering more nuanced perspectives on this assessment, distinguishing between claims of a relative lack of substantial technological innovations and those concerning the absence of a clear directional trend in technological evolution (de la Torre et al. 2013; Kuhn 2020).

While philosophers have engaged in dichotomous discussions of stability and change (see, e.g., contributions in Charbonneau 2024 ed., Sterelny 2017), and particularly regarding modes of information transmission and learning mechanisms, less attention has been dedicated to the construction and detection of the relevant archaeological signal and how such practices are mobilized in comparative assessments. This paper takes a step back to examine claims of comparative innovativeness, foregrounding features of the material evidence and the philosophical challenges these claims present.

Drawing on recent developments in the philosophy of biology regarding comparative assessments of innovativeness (Ramsey and Meneganzin 2025), the philosophy of paleoanthropology—particularly in relation to 'trait' or character identification (Meneganzin et al. 2024)—and theoretical archaeology, especially concerning cultural taxonomy (Reynold and Riede 2019), I identify two key problems hindering progress in this area: (1) the inertia created by treating named assemblages as relevant units of analysis and proxies for loci of innovation, and (2) the neglect of a multidimensional, multilevel approach to comparative innovativeness for understanding what the material record is ultimately a record of. By outlining productive ways forward, this paper aims to contribute to the redefinition and reassessment of the material explanandum of innovativeness in the Middle Pleistocene archaeological record.

References

Charbonneau, M. (Ed.). (2024). The Evolution of Techniques: Rigidity and Flexibility in Use, Transmission, and Innovation. MIT Press.

de la Torre, I., Martínez-Moreno, J., & Mora, R. (2013). Change and stasis in the Iberian Middle Paleolithic: considerations on the significance of Mousterian technological variability. Current Anthropology, 54(S8), S320-S336.

Kuhn, S. L., & Stiner, M. C. (1998). Middle Paleolithic 'Creativity. In: Creativity in Human Evolution and Prehistory, 104. Edited by Steven Mithen. Routledge.

Mellars, P. (1989). Technological changes across the Middle-Upper Palaeolithic transition: economic, social and cognitive perspectives (pp. 339-365). Princeton university Press.

Meneganzin, A., Ramsey, G., & DiFrisco, J. (2024). What is a trait? Lessons from the human chin. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 342(2), 65-75.

Ramsey, G., & Meneganzin, A. (2025). Multilevel innovativeness and cross-species comparisons. European Journal for Philosophy of Science, 15(3), 1-25.

Reynolds, N., & Riede, F. (2019). House of cards: cultural taxonomy and the study of the European Upper Palaeolithic. Antiquity, 93(371), 1350-1358.
Sterelny, K. (2020). Innovation, life history and social networks in human evolution. Philosophical Transactions of the Royal Society B, 375(1803), 20190497.

Locating Normativity in the Archaeological Record

Anton Killin[1] & Ross Pain[2]
[1]Department of Philosophy, Bielefeld University, Bielefeld, Germany; [2]Department of Philosophy, University of Bristol, Bristol, United Kingdom.

Evolutionary cognitive archaeologists aim to investigate the evolution of human cognition by way of the archaeological record. Using artifacts like stone tools, they employ models from the cognitive sciences or other theoretical frameworks as mid-range theories in order to license inferences about cognition from the material traces of the past. This can include comparative thinking, marshalling data from non-hominin species (e.g., Wynn et al. 2011). Over the last 20 years, archaeology has expanded beyond the hominin lineage and excavated the sites of 3 other species of primate: chimpanzees, macaques, and capuchins (Falótico et al. 2019; Haslam et al. 2016; Haslam et al. 2017; Mercador et al. 2007). This talk will place a spotlight on the comparative dimension of evolutionary cognitive archaeology—and its prospects—by mobilizing normativity as a case study.
A major evolutionary-cognitive transition in the human lineage is often taken to be the onset and enrichment of normative thinking. But understood behaviourally, normative regularities may be a phylogenetically widespread phenomena—as argued recently by Kristin Andrews and colleagues (e.g., Andrews et al. 2024; Westra et al. 2024). Indeed, Andrews et al. offer a 6-dimensional, incrementalist framework for gauging different grades of normative behaviour, and in this talk we apply this kind of thinking to the archaeological signatures of chimpanzees, other primates, and early hominins. While Andrews et al. put their framework to use in characterizing norms across (extant) species, we will argue that it is a promising tool for (cognitive) archaeological and phylogenetic thinking: the general strategy allows for the creation of a baseline reading of normativity among related taxa. In particular, it allows us to compare ethological data relating to normative behaviours involved in toolmaking with excavated archaeological data. This strategy is especially promising in the case of our closest primate relative, the chimpanzee. In comparing the results of this process with the earliest stages of the hominin record, we can get a more accurate picture of the evolution of normative behaviour.

References

Andrews K, Fitzpatrick S, Westra E (2024) Human and nonhuman norms. Phil. Trans. R. Soc. B 379: 20230026.
Falótico T et al. (2019) Three thousand years of wild capuchin stone tool use. Nature Ecology & Evolution 3: 1034-1038.
Haslam M et al. (2016) Archaeological excavation of wild macaque stone tools. Journal of Human Evolution 96: 134-138.
Haslam M et al. (2017) Primate archaeology evolves. Nature Ecology & Evolution 1: 1431-1437.
Mercador J et al. (2007) 4,300-Year-old chimpanzee sites and the origins of percussive stone technology. PNAS 104(9): 3043-3048.
Westra E et al. (2024) In search of animal normativity: a framework for studying social norms in non-human animals. Biological Reviews 99: 1058-1074.

Wynn T et al. (2011) “An Ape’s View of the Oldowan” revisited. Evolutionary Anthropology 20: 181-197.

Bridging Observed and Inferred Behaviour: Toward a Primate Model for Archaeology

Chiara Pertile, Mathilde Lequin, Solange Rigaud
Univ. Bordeaux, CNRS, Ministère de la Culture, PACEA, UMR 5199, F-33600 Pessac, France

Over the past two decades, research on non-human primates has increasingly highlighted the complexity of their cultural and technical repertoires, including stable behavioural traditions, intergroup variation, and diverse mechanisms of social learning (Whiten 2019; 2021). Within this broader landscape, primate archaeology has emerged as a way to apply archaeological methods to the study of primate tool-use, systematically documenting raw material choices and modification produced for and by these activities. Empirical studies (i.e. on chimpanzees, macaques and capuchins) have shown that specific technical practices can generate identifiable material signatures. These include recurrent patterns in raw-material selection and spatial distributions, while also displaying marked inter- and intra-group variation shaped by ecological context (Haslam et al. 2009; Pascual-Garrido et al. 2024). A central assumption is that the relationships between behaviour, local environment, and material culture in extant primates may provide insights relevant for interpreting early hominin technological evolution (Haslam et al. 2017). On this basis, comparative primate data have been proposed as a resource for archaeological inference, particularly for (i) contextualizing patterns of emergence and maintenance in early technical behaviour; and (ii) constraining hypotheses on the cognitive and cultural capacities attributed to hominin populations.

Despite its heuristic potential, the integration of primate behavioural data within archaeological researches faces several epistemic and methodological obstacles. A first difficulty concerns the lack of a shared conceptual framework capable of aligning the evidential standards of archaeology and primatology. The two fields operate with distinct inferential regimes: whereas primatology proceeds from observed behaviour to its material outcomes, archaeology must infer technical and cognitive capacities from fragmentary record of stone or bone tools. These reconstructions depend on the mediating models developed in cognitive archaeology, which link artefacts to the unobserved behavioural and cognitive processes that produced them (Coolidge, Wynn 2016; Currie, Killin 2019; Killin, Pain 2025). This asymmetry raises foundational questions about how, and to what extent, the relationships between behaviour, ecological context and material modification documented in living primates can inform hypotheses about early hominin behaviour. Furthermore, the comparative use of primate data is constrained by well-documented limitations of cross-species comparisons. As emphasized by Wood et al. (2024), differences in ecological affordances, motivational structures, attentional biases, task constraints and opportunities for social learning can generate behavioural divergences unrelated to underlying cognitive differences. Even within primate archaeology, caution has been raised regarding equifinality in material signatures, the influence of local ecological conditions, and the difficulty of distinguishing convergent from homologous behavioural patterns (Pascual-Garrido et al. 2024).

To address these challenges, this talk examines both the potential and the limits of using the primate model within archaeological reasoning. We argue that progress requires the construction of a comparative framework, understood here as an epistemic strategy that organizes ecological, social, cognitive and material variables in a way that makes them explicitly comparable across disciplines. Such a framework clarifies the assumptions and limits of comparability that shape cross-species analogies, specifying when primate evidence

can, or cannot, legitimately inform archaeological hypotheses, and providing a more robust basis for productive dialogue between primatology and archaeology.

References

Coolidge, F. L., & Wynn, T. (2016). An introduction to cognitive archaeology. Current Directions in Psychological Science, 25(6), 386–392.
Currie, A., & Killin, A. (2019). From things to thinking: Cognitive archaeology. Mind & Language, 34, 263–279.
Haslam, M., et al. (2009). Primate archaeology. Nature, 460(7253), 339–344.
Haslam, M., et al. (2017). Primate archaeology evolves. Nature Ecology & Evolution, 1(10), 1431–1437.
Killin, A., & Pain, R. (2025). Four heuristics for theory selection in evolutionary cognitive archaeology. Journal of Cultural Cognitive Science, 9(2), 273–290.
Pascual-Garrido, A., et al. (2024). Primate archaeology 3.0. American Journal of Biological Anthropology, 183(3), e24835.
Whiten, A. (2019). Cultural evolution in animals. Annual Review of Ecology, Evolution, and Systematics, 50, 27–48.
Whiten, A. (2021). The burgeoning reach of animal culture. Science, 372(6537), eabe6514.
Wood, L. A., et al. (2023). Cross-species comparisons of human and non-human culture: Approaches, discoveries, limitations, and future directions. In: The Oxford Handbook of Cultural Evolution. Oxford University Press.

The Institutional Turn in Science and Values

Anna Leuschner
University of Wuppertal
leuschner@uni-wuppertal.de

Faik Kurtulmus
Sabanci University
faik.kurtulmus@sabanciuniv.edu

Kevin Elliott
Michigan State University
kce@msu.edu

Ahmad Elabbar
University of Cambridge
ae423@cam.ac.uk

In recent years, philosophers of science concerned with science and values, including those contributing to this symposium, have increasingly turned to the institutional foundations of science: they are examining how norms, rules, and standards—embodied in organizations such as universities, funding agencies, and regulatory bodies—structure what kinds of knowledge are produced, whose voices are heard, and which values guide scientific practice (e.g., Elabbar forthcoming; Elliott 2025; Gutiérrez Valderrama forthcoming; Irzik & Kurtumulus 2024; Leuschner & Fernández Pinto forthcoming). This symposium seeks to explore this institutional turn in the science and values debate. Our central premise is that focusing on institutions and organizational structures can provide important insights about how science can be more effectively deployed to address major contemporary global challenges such as persistent social inequalities, global injustices, climate change, and environmental degradation. This is especially important as it has become increasingly apparent how fragile many science-related institutions are, and how they are therefore being undermined or exploited—in some cases deliberately and strategically.

Thus, we argue that philosophers of science should further explore the institutional environment in which science takes place if they are to help address these problems. This means examining how science is shaped and guided by institutional arrangements, thereby promoting or hindering epistemic and social justice, structuring power relations, and protecting or damaging public trust in science. We will explore which conceptual tools could help to critically analyze, understand, and respond to the institutional challenges underlying the contemporary problems facing science and society, and what institutional and organizational conditions would need to change or be stabilized, and in what ways, for these problems to be effectively addressed.

The symposium brings together scholars working on agnotology and moral, feminist, and political philosophy of science. By examining how institutions shape value management in science, how evaluation practices anchor values under conditions of inequality, how methodological standards can both protect and jeopardize epistemic integrity, and how deliberate deviations from epistemic conventions reveal the limits of conventional representations of bias, the symposium will shed light on which institutions shape the relationship between science and society in a positive way and how they can be improved and

stabilized. It will thus contribute to a better understanding of how science can help address major global challenges of injustice and the erosion of democracy.

References

Elabbar, Ahmad (forthcoming). Varying Evidential Standards as a Matter of Justice. British Journal for the Philosophy of Science, just accepted. Doi: https://www.journals.uchicago.edu/doi/10.1086/727429
Elliott, Kevin C. (2025). Alleviating Epistemic Injustice in Environmental Health Research: Strategies from Science and Values. Topoi. https://doi.org/10.1007/s11245-025-10268-3
Gutiérrez Valderrama, Juliana (forthcoming). Values in Global Science and the Relevance of Geographic Diversity. In: Ted Richards and Kevin Elliott (eds.), Routledge Handbook of Science and Values.
Irzik, Gürol and Faik Kurtulmus (2024). Distributive Epistemic Justice in Science. The British Journal for the Philosophy of Science 75 (2), 325–345.
Leuschner, A. and M. Fernández Pinto (forthcoming). Epistemic Intimidation and Illegitimate Value-Influence in Science. In: Ted Richards and Kevin Elliott (eds.), Routledge Handbook of Science and Values.

1. Managing Values in Science: Why Background Institutions Matter (Falk Kurtulmus)

Approaches to the management of values that gave a central role to individual scientists are being supplanted by more institutional approaches. This shift is guided by three insights. First, many value-laden decisions in scientific research arise not from deliberate evaluative choices but because scientists' decisions themselves carry implicit value judgments. Second, many such decisions are governed by disciplinary conventions, such as how variables are operationalized or where significance levels are set. Third, since these decisions are normatively complex and sensitive to a wide range of considerations, scientists cannot be expected, on their own, to address them satisfactorily. As a result, the focus has broadened from individual scientists' choices to institutional mechanisms—such as methods for integrating public values into research and the design and norms of scientific advisory bodies at the science-policy interface. This paper argues that we should widen our institutional focus further by examining the role of institutions beyond those directly involved in scientific research or its application.

The talk offers two reasons for this broadened focus. The first concerns the prospects of realizing the objectives we have in managing values in science. These objectives may include ensuring that policies informed by science are democratically legitimate or consistent with the relevant requirements of justice. We may also aim to create conditions conducive to public trust in science and to promote the public's voluntary compliance with science-based policies. Whether these aims are realized depends not only on the organization of science but also on how well other institutions—such as the political system, the media, and institutions of education—function. Furthermore, which value-laden decisions within science best promote these objectives depends on such background institutional conditions. There is no single set of practices or value-laden choices that will reliably advance these aims under all background conditions. Thus, managing values in science to attain its objectives requires attending to these background institutions and, where necessary, considering reforms to them alongside reforms to scientific institutions.

The second reason for a broadened focus arises from the various ways institutions outside science influence the value-laden choices within science. The reporting of science in journalism, for instance, not only shapes how far our objectives are realized but also affects

research priorities and how specific value-laden decisions are made. Social movements can challenge research priorities by drawing attention to neglected questions and can also identify and question value-laden choices that researchers may be unaware of. Their ability to perform this role depends crucially on the wider political system. The education system, to cite one last example, influences who becomes a scientist and thereby, indirectly, some of the value-laden decisions in science. Given these influences, proposals for managing values in science that do not attend to these external factors will be incomplete.
The talk concludes by discussing the limits to managing values in science—how much control is possible and how successfully such approaches can serve their broader aims—once their interactions with the wider institutional framework are taken into account.

2. Institutionalizing Values and Science: The Value-Ladenness of Methodological Standardization in Troubled Times (Anna Leuschner and Kevin Elliott)

Philosophers working on the topic of values and science increasingly recognize that they should investigate not only the value-laden choices made by individual scientists but also the value-laden practices associated with scientific organizations and institutions. This talk contributes to this effort to "institutionalize" the literature on values and science by exploring how the creation of standardized methodological rules, norms, and conventions (e.g., for study design, data collection, and analysis) supports some values while detracting from others. Building on a case study from regulatory toxicology, the talk argues that in our present social context, in which we grapple with widespread polarization, misinformation, and corruption, methodological standardization provides valuable resources for identifying and limiting bias.
Nevertheless, methodological standardization in science does not always succeed in preventing social problems, as we will illustrate with three examples. First, the tobacco industry's promotion of "distraction research" illustrates how scientifically sound studies on the many possible causes of lung cancer were strategically funded to deflect attention from the causal role of smoking (Proctor 2012). Second, as Elabbar (2024) has shown in the context of climate change, there is "morally subversive expertise," when scientifically correct and methodologically rigorous research (e.g., from IPCC reports) is paired with morally subversive narratives to undermine people's moral agency. Third, the chemical industry has used the call for "sound science" as a justification for demanding exceptionally high amounts of evidence before concluding that industrial chemicals are toxic, thereby paralyzing efforts at chemical regulation (Michaels 2008). In all these cases, the very standards that secure science's credibility are strategically mobilized to mislead. "Good science," thus, can become weaponized against the promotion of a "good society."
We will argue, however, that this does not render methodological standardization toothless. Rather, institutional standards remain crucial: although they cannot prevent disinformation or moral subversion, they promote high epistemic quality and transparency, which makes such strategies easier to detect and contest. Even in "troubled times," shared methodological standards constrain how far manipulative rhetoric can depart from accepted knowledge without becoming publicly challengeable. Additionally, promoting institutional structures that facilitate justified deviation from methodological standards while taking account of the broader social context in which the standards are embedded can help to promote the benefits of standardization while minimizing its weaknesses and lessening its potential to be weaponized.
References:
Elabbar, A., 2024. Expertise, moral subversion, and climate deregulation. Synthese, 203(5), 147.

Michaels, D. 2008. Doubt Is Their Product. Oxford University Press.
Proctor, R.N., 2012. Golden Holocaust: Origins of the Cigarette Catastrophe and the Case for Abolition. University of California Press.

3. Collective Inquiry and Conscientious Disobedience: Epistemic Conventions, Bias, and the 'Greenhouse Detection Problem' circa 1990 (Ahmad Elabbar)

Once upon a time, the epistemic analysis of preference bias seemed straightforward: Research projects are biased, and scientists' conduct displays bias, to the extent that core methodological choices are shaped by social, ethical, or political values, presumed external to science. In recent philosophy of science, this traditional account of bias, wedded to the value-free ideal, has been widely rejected. Wilholt's (2009) conventional account of bias is among the most promising alternatives. On the conventional account, research is biased not when it is shaped by non-epistemic values per se, but when such values lead to methodological choices that violate the epistemic conventions of the research community. Such deviations break the fabric of epistemic trust in science, as complex communities, whether of science or any other kind, rely on shared conventions to coordinate their activities. The account is attractive for recovering an epistemic analysis of bias while accepting two major insights of contemporary philosophy of science: the value-ladenness and social character of scientific inquiry.

Yet, the conventional account faces a challenge, which I call the 'James Hansen problem', in honour of the climate scientist who, I argue from reconstructing the historical episode, knowingly violated his community's conventions during his 1988 U.S. congressional testimony, where he claimed, with 'high confidence', that the 'observed warming' was caused by the 'greenhouse effect'. Prima facie, the conventional account does not distinguish among reasons for deviating from conventions, nor forms of deviation, condemning all as instances of bias. Thus, the conventional account lacks the resources, as it stands, to distinguish the actions of scientists in Hansen's position from the actions of pharmaceutical and other industry-sponsored researchers whose convention-breaking actions motivated the account, lumping all such deviators as biased – and hence untrustworthy – epistemic agents. This is problematic: the form of convention violation exemplified in Hansen's actions seems genuinely less worrisome from an epistemic point of view, yet the conventional account is incapable of illuminating that intuition. The fact that Hansen largely retained his standing as a trusted member of his community, despite violating its conventions, suggests we should take this intuition seriously.

My aim is to develop the conventional account of bias in light of this challenge. I do so by introducing a distinction between two forms of deviation from epistemic conventions: deviations of opportunity and deviations of conscience. Deviations of opportunity occur whenever an opportunity arises to tilt a study towards the preferred outcome of a sponsor, as exemplified by industry-sponsored researchers discussed in the literature. Such deviations occur frequently, covertly, and are directionally inconsistent. Because of these features, deviations of opportunity are incompatible with the coordinative role of conventions; therefore, their widespread presence in an epistemic community would break the fabric of trust. By contrast, deviations of conscience occur rarely, overtly, and, owing to the consistency of reasons triggering them, are directionally consistent; therefore, they do not break the fabric of trust in an epistemic community. The paper thus works towards reconciling a tension between the agential and institutional strands in the values in science literature.

Investigating Biological Variation

James Lowe
University of Exeter
j.w.e.lowe2@exeter.ac.uk

Cristina Villegas
Konrad Lorenz Institute for Evolution and Cognition Research
cristina.villegas@kli.ac.at

Alan Love
University of Minnesota
aclove@umn.edu

As Dan Nicholson (2025) has observed, "Variation is a primary, irreducible, and inescapable feature of the living world." Indeed, the identification, classification, analysis, interpretation, intervention, control and relation of different forms of variation is central to the life sciences. This session aims to further develop a concerted endeavour to characterise and understand biological variation and its place in scientific practices.

The term 'varipraxis' has been coined to capture the entangled methodological, epistemic and ontological features of the elucidation of variation. Varipraxis is defined "as the set of practices involved in the apprehension, measurement, investigation and analysis of different forms of variation. It both captures forms of variation that are identified and used, and the articulation of them in particular domains" (Lowe and Ingram 2023).

Formulating the concept of varipraxis opens the way to establishing systematic means for elucidating and interpreting the particular aspects of biological variation salient in a given research programme or endeavour. How are forms of variation conceptualised, explained, and theorised in relation to each other?

To further develop the articulation of varipraxis as a way of comprehending the treatment and conceptualisation of biological variation in scientific practice, the contributions in this session will further explore these entangled features of the elucidation and deployment of variation with diverse examples across distinct areas of fundamental and translational life science research.

In 'The Constitution of Variables in Causal AI', James Lowe will argue that understanding how and why particular variables are formed and enter scientific and translational practices requires a holistic approach: constituting variables. He will use causal AI – a suite of approaches that applies causal principles to AI modelling – to identify key features of the constitution of variables and consequent lines of inquiry.

In 'Modeling Evolutionary Variation with Causal Probability', Cristina Villegas will analyse some of the abstractions and idealisations embedded in developmental evolution models of phenotypic variation. She will argue that some of these models can be understood as causal probability models that measure the causal tendency of phenotypic traits to produce variation under certain conditions.

In 'What is a Trait That It Might Vary? Genetics and Variation in Developmental Biological Practice', Alan Love scrutinises the differences in how traits and variation are conceptualised in genetics and developmental biology in light of new theoretical perspectives on character identity mechanisms. He will argue that these divergent frameworks and standard practices can be integrated by distinguishing between different kinds of traits and keeping trait identity

separate from trait state, thereby yielding a practical basis for the meaningfulness of trait attributions and the measurement of variation.
All three talks articulate ways that requirements and goals of explanations, modelling and theory interact with scientific practices concerning biological variation. This will help us to identify salient avenues for developing our investigation and understanding of varipraxis.

References

Lowe, J.W.E., Ingram, D.S. DNA barcoding and the changing ontological commitments of taxonomy. Biol Philos 38, 32 (2023).
Nicholson, D.J. Population Thinking and the Uniqueness of Biological Entities. Acta Biotheor 73, 8 (2025).

Constituting Variables in Causal AI
James Lowe
Egenis, Centre for the study of Life Sciences; University of Exeter & Causality in Healthcare AI (CHAI) Hub

To be made tractable for scientific investigation, the variation that exists in nature – or could be generated through manipulations – has to be turned into variables. However, as Benjamin Jantzen (2021) has noted, "Despite their central role in science […] variables have received scant direct attention and remain remarkably obscure."
The literature on variables concerns itself with the nature of scientific variables, variable definition, choice, operationalisation and entrenchment in particular data practices and investigative systems. I argue that understanding how and why particular variables are formed or otherwise enter into scientific and translational practices requires a more holistic approach that encompasses the full technical, methodological, conceptual, practical and social/organisational array of means by which variables are apprehended, construed, selected, incorporated, related, analysed and evaluated over time in a given endeavour. This is captured under the notion of the constitution of variables, or constituting variables.
This talk seeks to establish the basis for this holistic approach. This will allow us to better understand and characterise the use of variables in scientific practices, and the consequences of this. In so doing, it will provide us with a framework and a set of tools to engage with scientific and translational practitioners around the decisions and constraints involved. It is also intended as a contribution towards the wider project of articulating varipraxis, "the apprehension, measurement, investigation and analysis of different forms of variation […that] both captures forms of variation that are identified and used, and the articulation of them in particular domains" (Lowe and Ingram 2023).
I will use causal AI to identify key features and salient lines of inquiry that are raised by examination of the constitution of variables. Causal AI encompasses a suite of approaches that seek to apply causal modelling and the apparatus of causal inference to AI models and modelling.
Causal AI is a useful for explicating the constitution of variables, as it involves: explicit formal modelling of variables and their relationships; modelling the processes of data generation, processing, analysis and interpretation; heterogeneity in the kinds of variables, how they enter into a given varipraxis and how they are used within it; and novel encounters with established research endeavours that engenders a variety of ways in which existing variables are engaged with. This may result in adaptation of variables for new purposes or even the constitution of new variables.

I will examine how variables are constituted in the framework of the creation and deployment of causal graphs, which allows me to highlight the impact of variable constitution as well as its highly contextualised nature, contingent as it is on the epistemic goals and conditions of the causal modelling.

References

Jantzen, B.C. Scientific Variables. Philosophies 6, 4: 103 (2021).
Lowe, J.W.E., Ingram, D.S. DNA barcoding and the changing ontological commitments of taxonomy. Biol Philos 38, 32 (2023).

Modeling Evolutionary Variation with Causal Probability

Cristina Villegas
Konrad Lorenz Institute for Evolution and Cognition Research

In recent years, philosophers have paid attention to the probabilistic aspects of classical evolutionary models. In particular, the nature of natural selection and genetic drift have been approached both from a propensity view (Pence and Ramsey 2013) and a causal probability approach (Abrams 2023). However, the role of probabilities has not been philosophically considered with respect to current central concerns in evolutionary biology, such as the origin and impact of phenotypic variation in evolution, a key contribution of the developmental perspective to evolutionary thought (Müller 2021). Some preliminary work has associated the new developmental models of phenotypic variation to the idea of propensities (Nu√±o de la Rosa and Villegas 2022), but much philosophical work remains to be done in this regard.
In this paper, I analyze the models of developmental evolution (devo-evo) from a philosophy of probability and chance viewpoint in order to fill this gap. Devo-evo studies the impact of organismal developmental properties in the course of evolution, building formal models of how variation is produced in specific lineages (Pavli*f*çev 2021). These models explain and predict evolutionarily relevant phenotypic change in biological traits on the basis of their developmental properties, building a mapping function from genotypic variants to phenotypic ones: the genotype-phenotype map (Alberch 1991). While some of these models are qualitative, I approach them through the framework of imprecise probability, focusing on those that still permit a quantification of so-called variational tendencies: variability, evolvability, robustness, modularity, and plasticity (Wagner 2014). Following recent tendencies in philosophy of biology (Abrams 2023), I apply a causal probability view to these variational tendencies, focusing on modelling aspects such as the abstractions and idealizations embedded in their derivation. In doing so, I argue that phenotypic traits work as idealized chance set-ups bearing variational propensities in these models, and I argue that the levels of determination or modalities of traits can be associated with levels of abstraction of initial conditions for such chance set-ups. I conclude with some philosophical implications about the roles of chance as randomness and as contingency in evolution.

References

Abrams, M. (2023). Evolution and the machinery of chance: philosophy, probability, and scientific practice in biology. University of Chicago Press.
Alberch, P. (1991). From genes to phenotype: dynamical systems and evolvability. Genetica, Vol. 84, No. 1, 5-11.
Müller, G. B. (2021). Evo-devo's contributions to the extended evolutionary synthesis. Evolutionary developmental biology: a reference guide, 1127-1138.
Nu√±o de la Rosa, L., & Villegas, C. (2022). Chances and propensities in evo-devo. The British Journal for the Philosophy of Science.

Pavli*f*çev, M. (2021). Developmental evolutionary biology (devo-evo). In Evolutionary developmental biology: a reference guide (pp. 1033-1046). Cham: Springer International Publishing.
Pence, C. H., & Ramsey, G. (2013). A new foundation for the propensity interpretation of fitness. The British Journal for the Philosophy of Science.
Wagner, G. P. (2014). Homology, genes, and evolutionary innovation. Princeton University Press.

What is a Trait That It Might Vary? Genetics and Variation in Developmental Biological Practice

Alan C. Love
Department of Philosophy and Minnesota Center for Philosophy of Science, University of Minnesota

The split between genetics and embryology in the early 20th century led to distinct fields that matured independently and incommensurately. Different central questions organized the sciences with consequences for their respective scientific aims. Genetics is concerned with reliably predicting the relationship between genetic and phenotypic variation, whereas embryology (developmental biology) is concerned with elucidating ontogenetic mechanisms. This disconnect makes it difficult for genetics and developmental biology to jointly pursue shared problems. Variation is the central focus of genetics, but it is largely absent or ignored by developmental biology, despite the importance of "genetic mechanisms" in its practices. Traits are considered continuously variable in genetics and polygenic models are deployed to account for quantitative variation in traits. In developmental biology, the genetic architecture of characters is often conceptualized as qualitatively distinct features (e.g., a body part) underlain by specific genetic regulatory networks. Other differences manifested in scientific practices include divergence in features tracked (quantitative variation of traits in populations in genetics versus mechanism details in model organisms in developmental biology) and styles of experimentation (controlled breeding and mutational analysis in genetics versus intervention on mechanism constituents in developmental biology). How, if at all, can these divergent frameworks be reconciled?

I explore how to integrate these two sciences by drawing attention to how trait categories can be theorized in standard practices. Using new theorizing about character identity mechanisms (DiFrisco et al. 2020), which draws a distinction between different kinds of traits (characters versus properties) and keeps trait identity separate from trait state, I develop resources for deciding whether something is a trait of a particular kind through a specific tripartite strategy: conceptually distinguish different kinds of traits, differentiate between the genetic architecture for character identity or state and quantitative variation of properties, and strategically use model systems where one can study these types of architecture simultaneously. This provides a practical basis for identifying differences in the meaningfulness of trait attributions and the measurement of variation that can guide novel forms of developmental biological inquiry.

Genetics can then be understood as concerned with the quantitative variation of character states and properties for which developmental genetic mechanisms provide a qualitative basis or substrate (characters). This framing offers a bridge between the standard developmental biology question where variation is orthogonal—how do physical and biochemical resources contribute to the mechanisms of development that yield an integrated, functional trait?—and the standard genetics questions where variation is foregrounded—how do differences in the genome and other resources relate to phenotypic differences that shape the genotype-phenotype map? Articulating this conceptual framework to simultaneously address

differences in these approaches through the work of integration has consequences for how future inquiry into the nature of traits is pursued.

Reference

DiFrisco, J., A.C. Love and G.P. Wagner. 2020. Character identity mechanisms: a conceptual model for comparative-mechanistic biology. Biology & Philosophy 35(4):44.

Mathematics in Practice: A DMRCP Symposium

Benedikt Löwe
Universität Hamburg
loewe@math.uni-hamburg.de

Deborah Kant
Vrije Universiteit Brussel
deborah.kant@uni-hamburg.de

Mikkel Willum Johansen
University of Copenhagen
mwj@ind.ku.dk

Matthew Inglis
University of Loughborough
M.J.Inglis@lboro.ac.uk

Coordinator: Benedikt Löwe (Hamburg, Germany & Cambridge, U.K.)

The study of scientific practices has become a central theme in contemporary philosophy of science, offering rich accounts of how scientific knowledge is generated. In contrast, mathematical practices have only very relatively recently started to attract philosophical attention. A field that could be called "Empirical Philosophy of Mathematics" has emerged as a vibrant and methodologically diverse research area in its own right, understood as a sub-discipline of the philosophy of mathematical practice more generally. It integrates insights from philosophy, history, sociology, cognitive science, and educational research in order to better understand mathematics as an evolving human activity rather than merely a corpus of abstract results, and focuses especially on incorporating empirical data on mathematical practices.

Within this landscape, the research network "Diversity of Mathematical Research Cultures and Practices" (DMRCP) based at the Universität Hamburg plays a central role in promoting empirical approaches to the philosophy of mathematics. The network fosters collaborations that combine cognitive investigations with corpus-based methods, interviews, ethnographic observation, and other qualitative techniques. These interdisciplinary approaches illuminate the variety of ways in which mathematicians reason, communicate, organize their work, and evaluate the status of mathematical knowledge.

Our symposium will showcases several of the main strands currently shaping the empirical philosophy of mathematics. It will consist of three talks (by Deborah Kant, Mikkel Willum Johansen, and Matthew Inglis) which, in turn, discuss an overview of the uses of qualitative studies of mathematical practice (Kant), a study on the use of diagrams in mathematics applying a multitude of different empirical methods from the social sciences (Johansen), and a comparison of well-established studies from empirical linguistics about the use of conditionals with the situation within mathematical contexts (Inglis). The symposium will conclude with a panel discussion session during which researchers in empirical philosophy of mathematics aim to relate to and contrast with the analogous research in the wider field of philosophy of science in practice (PSP), aiming to learn from the experiences that researchers in PSP have made.

Talk 1.
Deborah Kant (Brussels, Belgium).
Qualitative Approaches to Mathematical Practices

This talk provides an overview of qualitative approaches to the study of mathematical practices, focusing on two central aspects: the wide variety of topics that have been investigated, and the methodological interplay between qualitative data collection and philosophical analysis. Qualitative methods are relatively time-consuming and therefore not yet widely applied within the philosophy of mathematics. Nonetheless, their ability to uncover previously unknown dimensions of mathematical activity—often with a high degree of contextual richness and conceptual nuance—has motivated a growing number of scholars to engage directly with mathematicians through interviews, observations, and related methods.

A recurring theme in qualitative investigations has been peer review in mathematics. Christian Greiffenhagen's (2023a,b) interviews with editors of mathematics journals reveal the delicate balance editors strike between collecting quick judgements about the significance of a submission and considering more time-intensive evaluations of the rigour and correctness of its arguments. Complementing this editorial perspective, Line Edslev Andersen (2017) has studied the practices and criteria of referees, shedding light on how they interpret their role, negotiate uncertainty, and justify their recommendations. Further topics explored in interview-based studies include academic integrity (Goddiksen et al. 2021), mathematicians' work with definitions (Coumans and Consoli 2023), and their views on philosophically challenging issues such as set-theoretic independence (Kant 2025). These investigations collectively demonstrate the capacity of qualitative methods to illuminate philosophically relevant aspects of mathematical reasoning, evaluation, and communication.

In engaging with this body of work, the talk will also examine the methodological structure of qualitative approaches in the empirical philosophy of mathematics (Kant and Löwe 2025). It is important not to conflate these approaches with qualitative research in sociology tout court. While sociological methods—such as semi-structured interviews and thematic coding—provide essential tools and standards of rigour, empirical philosophy of mathematics must also satisfy the conceptual and argumentative standards internal to philosophy. In practice, this means that qualitative studies begin from recognisably philosophical questions about mathematics or mathematical practice, and culminate in conclusions that contribute to philosophical debates. Maintaining this bidirectional connection between empirical findings and philosophical significance is both a methodological challenge and an opportunity for innovation.

The concluding part of the talk will address this challenge directly. How can descriptive accounts of what mathematicians do, believe, or value yield philosophically relevant insights? What kinds of inferences, conceptual distinctions, or theoretical frameworks are needed to connect empirical observations to normative or explanatory claims about mathematics? By reflecting on these questions, the talk aims to clarify the distinctive contribution of qualitative research to the philosophy of mathematical practice, and to articulate strategies for integrating empirical richness with philosophical analysis in a systematic and productive way.

Mikkel Willum Johansen (Copenhagen, Denmark)
Investigating mathematical diagrams using empirical methods

In this talk I will report the methods and results of a series of empirical investigations of mathematicians' use of mathematical diagrams. The methods include qualitative interviews, qualitative case analysis, large-scale quantitative survey and digital humanities methods. The different methods were chosen in an iterative way through the series of investigations to explore different aspects of mathematicians' diagram-related practices.
The investigation began with a qualitative interview study exploring mathematicians use of diagrams (Johansen & Misfeldt, 2018). The main results showed 1) that diagrams play a central role in mathematicians' work practices, and 2) that mathematicians nevertheless are reluctant to use diagrams in their published work. To explore the last result further, we conducted a large-scale quantitative investigation mapping the prevalence of diagrams in mathematics journals from 1885 to 1915 (Johansen & Pallavicini, 2022). The investigation showed that diagrams were relatively common in mathematical research publication around year 1900, then they disappeared for half a century and reappeared during the 1960s. In 2015 about 65% of the surveyed papers included at least one diagram. Focused qualitative analysis of selected parts of the corpus revealed that there are significant differences in the diagrams used in the period around 1900 and in the period after 1960. In short, the diagrams used in the latter period conform more closely to ideals of rigor connected to formalist ideologies (Johansen & Pallavicini 2021). In-depth analysis of more recent diagrams used in preprints published on arXiv showed a return to a more relaxed use of diagrams. Both the types of diagrams used and the epistemic function they were allowed to play were less restricted by formalist norms (Sørensen & Johansen, 2022). In the most recent (still unpublished) investigation, the use of diagrams in arXiv preprints is compared with the use in published versions of the same manuscripts. In this investigation quantitative methods are combined with qualitative close-readings of selected papers to form a multi-dimensional comparison of the use of diagrams in preprints vs published papers.
In the talk I will give an overview of the main results of the investigations outlined above, and I will discuss the challenges and opportunities connected to a multi-methodological approach to empirical investigations of mathematical practices.

Matthew Inglis (Loughborough, U.K.)
How does the believability of true mathematical statements influence the inferences that mathematicians draw from them?

Conditionals of the form 'if A(x), then B(x)' are ubiquitous in both natural language and mathematics. In natural language, decades of psychological experiments have established that the inferences reasoners draw from a conditional are affected by the conditional's believability. Is this also true in mathematics?
Answering this question first requires the development of an empirical method to systematically measure the 'believability' (or 'credence') of mathematical statements, and to establish whether true mathematical statements – despite being equally true – systematically differ in believability. I report two studies which used a comparative judgement method to assess mathematicians' and students' perceptions of which mathematical statements are more or less believable. Participants were shown successive pairs of conditionals and simply asked which was the more believable. These pairwise judgements were fitted to the Bradley-Terry statistical model and parameter 'believability' scores found. Four main results were established. First, participants are willing to make such judgements, even when the two conditionals are both (obviously) true, suggesting that believability is a genuine notion that

differs from perceptions of truth. Second, participants showed high levels of agreement with each other (high internal reliability), suggesting that there is a shared conception of believability. Third, research-active mathematicians and undergraduate mathematics students had, at least for these conditionals, similar views on believability (the Bradley-Terry parameter values from the two groups were strongly correlated). Fourth, some false mathematical conditionals are apparently more believable than some true mathematical conditionals, again indicating that believability and truth are not identical concepts.

Having established that mathematical conditionals differ in believability, we can ask whether the believability of a given conditional influences the inferences mathematicians are willing to draw from it. I report two experiments which compared the inferences undergraduate mathematics students made make from more- and less-believable mathematical conditionals (every conditional was true, but had a false converse). The first experiment demonstrated that participants were willing to make more valid (modus tollens) inferences from more-believable conditionals than less-believable conditionals, and that this effect was larger than the equivalent effect with day-to-day conditionals. The second experiment extended this finding to explore the effect of varying the believability of conditionals' converses. It found more endorsement of invalid (affirmation of the consequent and denial of the antecedent) inferences from (true) conditionals that had more-believable converses than those which had less-believable converses.

A final experiment explored the possibility that these results are due to the possibility that more-believable conditionals are easier to think about than less-believable conditionals. We found that differences in inference endorsement are best understood as being due to believability, not easiness.

In the talk I will discuss what it might mean for different (true) mathematical statements to differ in believability, suggest that these kinds of experiments shed light on different approaches to mathematical inference, and reflect on some open questions.

Infrastructure as method: scaffolding technologies in translational research and practice

Robert Meunier
ICI Berlin
robert.meunier@ici-berlin.org

Sophia Efstathiou
Norwegian University of Science and Technology (NTNU)
sophia.efstathiou@ntnu.no

Catherine Kendig
Michigan State University
kendig@msu.edu

Ariane Hanemaayer
Brandon University
HanemaayerA@BrandonU.CA

Infrastructure has been an important research focus in Science and Technology Studies, Human Geography, and Anthropology. Philosophers of science have been less engaged with infrastructures. Even though important work has addressed research infrastructures, esp. data infrastructures (e.g., Leonelli 2013), many other ways in which science and infrastructures interact are still understudied. This interdisciplinary session explores the potential of investigating these interactions for philosophy of science in practice. We focus on the ways in which scientific activity on the one hand relies on and is shaped by infrastructures and, on the other hand, contributes to the making, interpretation, and transformation of infrastructures.
Scholars of infrastructure typically reject the reduction of infrastructure to its material aspects, the pipes, tracks, and wires that help provide utilities, transport, or communication. Instead, infrastructures are understood as heterogeneous constellations that also involve social and institutional structures. According to one definition by anthropologists Collier & Ong (2003), the term 'infrastructure' "designates specific institutional, material, or social conditions through which the functioning of a certain technology, ethical regime, form of regulation, or mode of communication is either enabled or impeded." (423)
'Scaffolding' in our title points to a dynamic interaction between a given structure and the structures it enables/impedes. In science, as elsewhere, infrastructures are structured constellations, but they can also impose a structure on the relations they engender. In turn, activities coordinated by an infrastructure can feed back into that infrastructure. This dynamic aspect is often expressed by the term 'infrastructuring'. Focusing on infrastructure in science, further, highlights how systems of social and material technologies that scaffold certain relationships are typically themselves scaffolded by other technologies.
These general insights motivate specific case studies. In the four individual contributions from philosophy and sociology, we investigate the ways infrastructures interact with and engage epistemic and normative aspect of science, esp. concerning translational practices. The papers discuss 1) how infrastructures shape performances and ethical relations in science, e.g., concerning ethical encounters with experimental animals (Efstathiou); 2) how experimental systems can work as technological infrastructures that facilitate new explanatory tools, e.g. the import of concepts from soil microbiology to the field of faecal microbiota transplantation (Kendig); 3) how infrastructures serve as interfaces, e.g. when mediating between biomedical research and clinical practice (Meunier); and 4) how

infrastructures of governance shape research practice, e.g. when health care infrastructures are persistent and hence impose historically grown relations on new initiatives (Hanemaayer). Together the cases show how infrastructure constitutes an often-unacknowledged part of scientific method. Hence, analysing infrastructure expands the methodological toolbox of philosophy, while contributing to the broader study of infrastructure from new perspectives. The session will feature a 10 min introduction, four talks, each 15-min plus 5 min for clarification, and a 30 min collective Q&A session.

Collier, S. J., & Ong, A. (2003). Oikos/Anthropos: Rationality, Technology, Infrastructure. Current Anthropology, 44(3), 421–426.
Leonelli, S. (2013). Global data for local science: Assessing the scale of data infrastructures in biological and biomedical research. BioSocieties, 8(4), 449–465.

Infrastructures of effacement – structuring human-animal encounters in the animal lab
Sophia Efstathiou

This talk explores the role of scientific infrastructure in instituting, routinising and managing ethics in science. By 'science infrastructure' I understand purpose-designed, or found, material, social and symbolic structures that operate to support, maintain and promote the work of science and scientists in its institutional, but also extra-institutional, societal or business functions. Science infra-structure can be understood as what is underlying (-infra meaning 'below' or 'under'), supporting or grounding, the structure(s) of science: enabling the stabilisation, reproduction, and continuation of the work of science and scientists, while providing a foundation for scientific expansion, translation and transformation.
This paper proposes that certain types of infrastructure involve removing moments of ethical encounter as part of the institution of 'good' science. I focus specifically on a set of technologies I identify as 'technologies of effacement' (Efstathiou 2019). Technologies of effacement have manifest aims like ensuring safety, hygiene, care, or expediency, but they in effect operate to block the moral potentialities of unique face-to-face encounters between morally significant beings (cf. Levinas 1969). By scripting and routinising encounters into expected, already known events, technologies of effacement can block ethics from happening: they remove opportunities to pause one's spontaneity and to meet the face of a unique, secret, Other (Levinas 1969, Efstathiou 2019).
Considering the context of an animal lab, technologies of effacement include: 1. Built environments, enclosures and architectures, e.g. cages, animal houses, 2. Rules for entering and exiting protected spaces, e.g. protocols for changing attire, or washing, 3. Personal Protective Equipment, e.g. lab clothes, gloves, masks, hats, 4. Protocols for handling animals, e.g. picking up a rat from their tail versus holding them in one's palm, 5. Techniques for naming and identification, e.g. drawing lines on an animal's tail or punching holes in their ear, versus calling them with a personal name (Efstathiou 2018, 2019, and see Kirk 2016, Nelson 2016). These five types of technologies of effacement are cited as technologies that become integrated into the broader infrastructure of animal experimentation, legitimating and reinforcing each other. They manifestly support the proper work of the lab by providing equipment for the performance of scientific labour in a safe, controlled, and epistemically reproducible setup – grounding good science. At the same time, these technologies operate to structure and routinise human-animal encounters in the animal research lab away from ethics: away from facing unique Others, and towards encounters between known types, whether lab animals or scientists.

Efstathiou, Sophia. 2019. "Facing Animal Research: Levinas and Technologies of Effacement," in Atterton Peter and Tamra Wright. Eds. Face-to-Face with Animals: Levinas and the Animal Question, 139-163. New York: SUNY Press.

Kirk, Robert G. W. 2016. Care in the Cage. In Kristian Bjørkdahl & Tone Druglitrø. Eds. Animal Housing and Human-Animal Relationships: Politics, Practices and Infrastructures, 167-184. London and New York: Routledge.
Levinas, Emmanuel. 1969 [1961]. Totality and infinity. Trans. A. Lingis, Pittsburgh, PA: Duquesne University Press.
Nelson, Nicole. 2016. Model homes for model organisms: Intersections of animal welfare and behavioral neuroscience around the environment of the laboratory mouse. BioSocieties, 11(1), 46-66.

Experimental systems as technological infrastructures: soil ecology-inspired gut microbiome research

Catherine Kendig

Once framed primarily through physiological and host-centric mechanisms, the gut is increasingly conceptualized as an ecological system structured by community-level interactions. Ecological concepts originally developed in soil microbiology now play central explanatory roles in gut microbiome research. The clinical success of faecal microbiota transplantation (FMT), particularly in treating recurrent Clostridioides difficile infections, exemplifies this shift (Flores et al. 2025). Despite clinical success, the mechanisms underlying FMT remained largely underspecified: which taxa were critical, how they interacted, and how the community stabilized were largely unknown. Traditional host-centric and molecular approaches were insufficient for capturing the causal complexity of microbial community dynamics. I examine how FMT researchers successfully captured this causal complexity by revising the technological and clinical infrastructure of the experimental systems as well as the body of knowledge that they relied upon to underwrite their causal claims.

The surprising robustness of FMT, the availability of high-throughput sequencing, and the absence of pre-existing mechanistic theories created a context in which FMT researchers were compelled to navigate empirical complexity, without stable theoretical expectations. In response, researchers imported ecological frameworks from soil science. These concepts allowed microbiologists to reconceptualize the gut not as a static organ but as a structured ecological system.

My paper investigates how concepts, such as habitat filtering, ecological succession, priority effects, keystone species, metabolic cross-feeding, and community coalescence from soil science restructured new ways to investigate microbe/gut systems experimentally. Relying on the case study of FMT clinical experiments, I show how the development of new experimental methods was not just scaffolded on soil science research, but their usefulness required significant infrastructural work to coordinate community-level technologies to investigate microbe/gut systems experimentally including sequencing skills, tools, and data repositories. In doing so, my focus is on the coordination of different epistemic, technical, and managerial tasks undertaken by investigators to make experimentation possible, what Pitt (1995) calls “technological infrastructures” of science and Creager and Landecker (2009) describe in their discussion of tissue culture, as the techniques that become infrastructural insofar as they scaffold thinking and offer a new mode of perception.

Using the notion of technological infrastructures, I show how soil microbial ecology—long concerned with diversity, spatial heterogeneity, and disturbance—provided precisely the kinds of resources microbiome researchers needed when confronting FMT engraftment patterns and post-antibiotic successional trajectories. However, it is the infrastructural work of coordinating the concepts, technologies and clinical practices that facilitated the asking of new ecological questions, use of new tools, and heuristics that supported decision-making in

a complex clinical and scientific environment—ranging from donor selection to therapeutic design in FMT. As such, or so I'll argue, experimental systems are themselves an important instance and technological site of infrastructuring and re-infrastructuring activities.

Creager, A., Landecker, H. 2009. Technical matters: method, knowledge and infrastructure in twentieth-century life science. Nature Methods 6(10): 701-705.
Flores, C., Millard, S., Seekatz, A. 2025. Bridging ecology and microbiomes: Applying ecological theories in host-associated microbial ecosystems. Current Clinical Microbiology Reports 12:19. https://doi.org/10.1007/s40588-025-00246-z
Pitt, J. 1998. Explaining Change in Science. Philosophy and Technology 3(3): 135-140.

Epistemological consequences of infrastructures as interfaces of translation in biomedicine
Robert Meunier

The relation between biomedical research and the clinical contexts in which biomedical knowledge is used is bi-directional and, or so I argue, in important ways mediated by infrastructures. I will address two kinds of infrastructures: shared infrastructures for the exchange of materials and reproduced infrastructures that enable the translation of research outcomes. In both cases, infrastructures not only provide a structure that scaffolds relations between research and clinic, but also structures the way disease-related phenomena can be known.
Designing biomedical research projects typically requires an understanding of the clinical context that enables researchers to identify relevant research problems and estimate the usability and utility of their research outcomes for clinicians. Additionally, research often requires access to sites, material, or data from the contexts of applications. Hence, biomedical research needs an interface with clinical contexts, which is often realized as shared infrastructures of collaboration, communication, and exchange. The first example then focuses on the role of biobanks in making material and data from the clinic available for research.
The second example is concerned with molecular high-throughput infrastructures. These are constitutive for experimental systems in biomedicine. As researchers often re-formulate clinical problems in terms of their own established methods, they assume that clinical practices can be assimilated to research practices: While sometimes technology-intensive research can be translated into low-tech clinical test, the utilization of insights from omics research often requires that the same technological infrastructures are present in the clinic. That is, parts of the experimental infrastructures of research need to be reproduced as platforms in clinical laboratories to perform molecular tests for clinical decision-making. In this case, infrastructures are not necessarily shared or involve direct interaction, but they nonetheless mediate the relation between research and clinical contexts.
STS scholars have investigated aspects of governance and standardization of biobanks (Metzler et al. 2023). Furthermore, studies of biomedical platforms have shown how they reconfigure clinical practice and its relation to research (Cambrosio et al. 2018). This work sets the stage for addressing epistemological effects of infrastructures in biomedicine. While scaffolding translational relations between research and clinic, they structure the way biomedical phenomena can be accessed by researchers and physicians. Regarding research, the standards and global decisions in setting up biobanks pre-configure the design of experiments, for instance, in terms of the selection of cases and controls. Regarding clinical practice, with high-throughput platforms physicians are given new ways to inhabit the molecular landscape of disease. This, however, re-configures their approach to these disease phenomena. Such a focus on infrastructures as interfaces between the contexts of research

and application might provide important insights into the epistemology of knowledge translation even beyond biomedicine.

Cambrosio, A., Vignola-Gagné, E., Nelson, N., Keating, P., & Bourret, P. (2018). Genomic platforms and clinical research. In Routledge Handbook of Genomics, Health and Society (2nd ed.). Routledge, 142-150.
Metzler, I., Ferent, L.-M., & Felt, U. (2023). On samples, data, and their mobility in biobanking: How imagined travels help to relate samples and data. Big Data & Society, 10(1), 1-13.

From Research to Standard of Care: Infrastructures and institutions in the making of a newborn screening program in the NHS
Ariane Hanemaayer

This paper traces the material and discursive conditions that enable and constrain the creation of an emerging newborn screening infrastructure in the United Kingdom. In October 2024, the National Health Service (NHS) announced that it had successfully begun a new study, in collaboration with Genomics England. Called "the Generation Study," the NHS partnership aims to sequence the genomes of 100,000 newborn babies by the end of 2025. The study recruits pregnant individuals from 20 weeks to participate. Participants' healthy newborns will have blood drawn and the whole genome will be sequenced, looking for more than 200 known conditions that could lead to illness later on (Rahman et al. 2024). The study promises to be a major research leap forward in England's reputation of being a leader in genomic medicine.
Politically, the study has been heralded as the Government's commitment to preventative medicine, ostensibly "leapfrogging disease" (Donnelly 2025). For example, some subspecialties have praised the study's selection of known biomarkers in dermatology (e.g., Bick et al. 2025), and new doctors believe that the study will enable preventative care (Seed et al. 2024). Critics, however, have been wary. Some have raised issues about the public's ability to understand both genomics and the ways that risk is interpreted with respect to biomarkers (Stroppa & Wilson 2025), while others have questioned the selection process for the genes used by the screening program on ethical and scientific grounds (Baple et al. 2024, Kaplanis et al. 2025), and key bioethical issues have been framed around questions of communication (Stroppa & Wilson 2025). Despite these concerns, humanists have yet to examine the institutional and material conditions under which the study is being rolled out. This paper will describe the landscape of NHS research infrastructure as the antecedent conditions upon which the Generation Study aims to institutionalize a newborn screening. Practically, the study is a form of research, with aspirations of translating newborn screening into standard care as one method of preventive medicine in the UK. The paper draws on published and grey materials about the study design, roll out, and evaluation in real time. I will trace the discursive and historical conditions, or "legacy effects," of governmental organization and infrastructures that impacts and shapes the current program. By examining the field of possibility that gave rise to the integration of genomic sequencing technology and the Generation Study, the paper informs the normative debate about genetic screening through a focus on the creation of new institutional infrastructure and the mechanisms that enable its success or failure.
Baple, E.L., Scott, R.H., Banka, S. et al. "Exploring the benefits, harms and costs of genomic newborn screening for rare diseases." Nat Med 30, 1823–1825 (2024).

Bick, D., Rajan, N., & Fassihi, H. “Newborn genome screening (the Generation Study): the advent of presymptomatic diagnosis for genodermatoses.” British Journal of Dermatology, 193(4), 767–769. (2025).
Donnelly, L. “DNA Screening for Every Baby on the NHS.” The Telegraph. (2025)
Kaplanis, J., et al. “Assessment of the variant prioritization strategy for genomic newborn screening in the Generation Study.” Genetics in Medicine. (2025).
Rahman, S., Bick, D., Scott, R.H. “Whole genome sequencing to screen 100 000 newborns for treatable genetic disorders.” J Inherit Metab Dis.47:7–8. (2024).
Seed L, Scott A, Pichini A, et al. “Perceptions of genomic newborn screening: a cross-sectional survey conducted with UK medical students.” BMJ Open 14:e089108 (2024).
Stroppa, L., Wilson, E. “The NHS plan to genetically test all newborns sounds smart – until it creates patients who aren’t sick.” The Conversation. (2025).

Ethical and Methodological Challenges in Human Population Genetics

Celso Neto
University of Exeter
C.Alves-Neto@Exeter.ac.uk

Yasmin Haddad
University of Quebec Montreal
yasmin.haddad@mail.mcgill.ca

Matilde Carrera
Boston University
mcarr@bu.edu

The field of human population genetics involves a variety of scientific practices of increasing social relevance. For instance, advanced modelling techniques for ancestry estimation have become key tools for genetic risk prediction in diverse (and often underrepresented) populations (Martin et al. 2017). These modeling practices face numerous challenges, such as biases in data collection and interpretation of results (Lawson et al. 2018; Lewis et al. 2022). Philosophers of science have discussed some of these challenges in the past, but these discussions rarely engage with recent methodologies in genomics (Gannett 2014; Winther 2022). Furthermore, while philosophers of science have increasingly turned their attention to social dimensions of science (e.g., "values in science"), this shift has not produced effects in the philosophical analysis of genomics (Haddad & Neto, forthcoming). The talks of this symposium aim at addressing this gap by adopting a practice-oriented approach and by bringing together a practicing geneticist and philosophers of science in practice in direct conversation.

The goal of this symposium is to examine ethical and methodological challenges arising in the modeling practices of current human population genetics. To do so, the symposium comprises 4 talks exploring complementary aspects of modelling practices around ancestry estimation and evolution in population genetics. For instance, Celso Neto discusses how ancestry estimation methodology has been accused of tautology, "self-fulfilling prophecy", or vicious circularity by social scientists (Lipphardt et al. 2021). Neto dispels these criticisms and shows how they are entangled with legitimate social/ethical concerns. Yasmin Haddad takes another step into clarifying the uses of models for ancestry inference. Haddad argues that adopting the "adequacy-for-purpose" view of models in population genetics helps to improve the conceptual and methodological basis of that inferential practice (Bokulich & Parker 2021). In turn, Matilde Carrera investigates modelling practices through the concept of underdetermination (Turnbull 2018). Carrera identifies types of underdetermination that have been overlooked by philosophers, connecting them to the torn challenge of parameter unidentifiability in inferences about human history and evolution. Aylwin Scally explores different modeling practices, but also gives special emphasis to the context of human evolution. Scally shows how both researchers and the public misinterpret models in problematic ways. In doing so, Scally also offers a "practitioners" perspective on some of the topics explored by the previous philosophy talks: how it is easy to misunderstand human population genetics; how it is easy to overlook purposes and the difference between models and reality; consequences of misinterpretation. Taken together, the talks explore a range of scientific case-studies, engaging with multidisciplinary scholarship and relying on interdisciplinary collaborations. Furthermore, the symposium identifies pressing challenges

coming out of the current scientific practice in human population genetics, reflecting on how geneticists, social scientists, philosophers, and the broader public (should) react to them.

Participants/Talks:

1. Celso Neto (Associate Professor, University of Exeter)
Signal and Noise in Human Population Genetics

Contemporary human population genetics is a data-intensive, model-based science. The field quickly evolves as new models analyse increasing amounts of data from samples across diverse populations. In recent decades, one important task of population genetics models is ancestry estimation: identifying the relative percentage of genetic ancestry in populations and individuals (Royal et al. 2010). This practice offers valuable insights into evolutionary history and medicine, contributing to better genetic disease risk prediction. However, ancestry estimation practices are also subject to controversy. Geneticists and social scientists worry that ancestry categories resemble 19th-century racial divisions, evoke fraught notions of biological race, and oversimplify human genetic relatedness (Lewis et al. 2022).
In this talk, I examine worries about the methodology of ancestry estimation. Social scientists have noted that ancestry estimation often relies on arbitrary or even circular methodological practices (Lipphardt et al. 2021; Schramm 2021). For instance, Lipphardt and colleagues argue that genetic studies on the European Roma population begin with the assumption that Roma evolved as an isolated population, which leads researchers to sample only isolated groups within contemporary Roma people. This sample reinforces the initial assumption, resulting in a biased reconstruction of Roma history. According to the authors, those genetic studies ignore important data ("signal") from mixed individuals, generating a "self-fulfilling prophecy."
The criticism from social scientists speaks to important underlying questions: what makes methodological choices arbitrary or circular in problematic ways? What makes data a signal rather than noise in human population genetics modelling? I address these questions by revisiting the Roma case and recent work on ancestry estimation (All of Us 2024). Contrary to what critics have argued, I show that often this research does not involve vicious circularity or pernicious wishful thinking. Assumptions about "isolated populations" enable scientists to initially distinguish signal and noise in the data without overdetermining results. Finally, I consider the entanglement between ethical and methodological challenges in human population genetics. In some cases, legitimate ethical worries get mistakenly framed in methodological terms, as the Roma case-study illustrates. In other cases, such as the representation of ancestry in the All of Us database, ethical and methodological concerns might be harder to disentangle.

Works cited

All of Us Research Demonstration Project Teams (2024). Genomic data in the All of Us research program. Nature, 627(8003), 340-346.
Lewis, A. C., Molina, S. J., Appelbaum, P. S., Dauda, B., Di Rienzo, A., Fuentes, A., ... & Allen, D. S. (2022). Getting genetic ancestry right for science and society. Science, 376(6590), 250-252.
Lipphardt, V., Rappold, G. A., & Surdu, M. (2021). Representing vulnerable populations in genetic studies: The case of the Roma. Science in Context, 34(1), 69-100.

Royal, C. D., Novembre, J., Fullerton, S. M., Goldstein, D. B., Long, J. C., Bamshad, M. J., & Clark, A. G. (2010). Inferring genetic ancestry: opportunities, challenges, and implications. The American Journal of Human Genetics, 86(5), 661-673.
Schramm, K. (2021). Race, genealogy, and the genomic archive in post-apartheid South Africa. Social Analysis, 65(4), 49-69.

2. Yasmin Haddad (Postdoctoral Fellow, Université du Québec à Montréal)
Adequacy-for-purpose and ancestry inference

Human population genetics studies similarities and differences among human groups, what is often called human genetic diversity. A central tool for studying this diversity is ancestry inference, the estimation of an individual's genetic background using information about their genetic ancestral origins (Padhukasahasram 2014). Studies sometimes may show discordance between self-identified ethnicity or race and genetic ancestry in some contexts. As such, self-reported ancestry does not always align with genetic patterns and may be incomplete for some analyses. Therefore, a common practice is to establish an individual's genetic ancestry based on a set of genetic markers. Using such ancestry information, researchers make predictions about predominant ancestry components and, in some contexts, about traits or risk factors correlated with those components.
Ancestry inference depends heavily on models – for example, models of population movements across time, models of current genetic distribution of genetic diversity, and statistical models used in clustering or admixture estimation (Bolnick 2008; Durand et al. 2021; Kidd et al. 2012). The parameters and standards guiding these models vary across research aims: genome-wide association studies rely on ancestry estimations for confounding control, whereas gene-discovery or demographic-history studies use ancestry to characterize patterns of variation. These approaches all presuppose assumptions about how human genetic diversity is structured. A common underlying ideal is that models should represent "true" or "natural" clusters of genetic similarity as fully and accurately as possible. However, this ideal tends to obscure the limitations of the models and can misrepresent the complexity of human genetic diversity.
An alternative approach, developed in philosophy of science, is the "adequacy-for-purpose_ view of models (Parker 2020; Bokulich and Parker 2021; Lusk and Elliott 2022). This paper argues that applying an adequacy-for-purpose framework to ancestry inference clarifies what model quality should amount to in practice. On this view, a model should be evaluated relative to the aims it is meant to serve, rather than relative to an idealized representation of the distribution of human genetic diversity. The paper thus contributes to the philosophy of science in practice by showing how adequacy-for-purpose can improve the conceptual and methodological foundations of ancestry inference.

Works cited

Bokulich, Alisa, and Wendy Parker. 2021. "Data Models, Representation and Adequacy-for-Purpose." European Journal for Philosophy of Science 11 (1): 31. https://doi.org/10.1007/s13194-020-00345-2.
Bolnick, Deborah A. 2008. "Individual Ancestry Inference and the Reification of Race as a Biological Phenomenon." In Revisiting Race in a Genomic Age, edited by Barbara A. Koenig, Sandra Soo-Jin Lee, and Sarah S. Richardson. Rutgers University Press.
Durand, Eric Y., Chuong B. Do, Peter R. Wilton, et al. 2021. A Scalable Pipeline for Local Ancestry Inference Using Tens of Thousands of Reference Haplotypes. Preprint. Genomics. https://doi.org/10.1101/2021.01.19.427308.

Kidd, Jeffrey M., Simon Gravel, Jake Byrnes, et al. 2012. “Population Genetic Inference from Personal Genome Data: Impact of Ancestry and Admixture on Human Genomic Variation.” The American Journal of Human Genetics 91 (4): 660–71. https://doi.org/10.1016/j.ajhg.2012.08.025.
Lusk, Greg, and Kevin C. Elliott. 2022. “Non-Epistemic Values and Scientific Assessment: An Adequacy-for-Purpose View.” European Journal for Philosophy of Science 12 (2): 1–22. https://doi.org/10.1007/s13194-022-00458-w.
Padhukasahasram, Badri. 2014. “Inferring Ancestry from Population Genomic Data and Its Applications.” Frontiers in Genetics 5 (July). https://doi.org/10.3389/fgene.2014.00204.
Parker, Wendy S. 2020. “Model Evaluation: An Adequacy-for-Purpose View.” Philosophy of Science 87 (3): 457–77. https://doi.org/10.1086/708691.

3. Matilde Carrera (PhD Candidate, Boston University)
Underdetermination in demographic inferences: the case of parameter unidentifiability

Philosophers of science interested in methodological issues in population genetics have mainly focused on two themes: the modeling approaches used to reconstruct evolutionary processes and the grouping of human populations based on genetic differences. Although this work has been influential, it primarily engages with scientific literature from the 1990s and early 2000s. This time gap raises two important questions. First, how effectively is philosophy of science staying relevant to current scientific practice in fast-moving fields like genomics? Second, in what ways can the philosophical analysis of scientific methods be improved by considering the latest advances in sequencing technologies, statistical inference, and computational modeling?
In this talk, I explore these questions by examining a family of inferential practices that have expanded significantly over the past twenty years: demographic inference from modern genetic variation to population histories (Marchi et al. 2021). These methods have rapidly grown in popularity due to next-generation sequencing and major advances in computational power, allowing us to infer past population sizes and demographic events, such as migrations, admixture, and bottlenecks. This information is crucial for reconstructing population histories over time. However, there is increasing awareness within the scientific community about methodological issues in the approaches used to make demographic inferences. Several mathematical and empirical findings indicate that the parameters of these demographic models are sometimes not identifiable, leading to what is known in statistics as an “unidentifiability problem” (Bhaskar et al. 2014, Mazet et al. 2015).
I argue that parameter unidentifiability can be subsumed under the philosophical concept of underdetermination. This has two main implications. First, it expands philosophical analyses of underdetermination by situating these mismatches within the relationship between data and statistical models, rather than between evidence and theory as traditionally framed (Stanford 2023; Turnbull 2018). Second, it clarifies how common the problem is in demographic inference and how it connects to other practices, such as model building and trade-offs between precision with generality. Overall, I contribute to the initial questions by offering a practice-based analysis of how underdetermination is individuated, how its epistemic impact on the inference process is evaluated, and how it is mitigated in the context of population genetics.

Bhaskar, Anand and Yun S. Song (2014). “Descartes’ rule of signs and the identifiability of population demographic models from genomic variation data”. Annals of Statistics 42.6, pp. 2469–2493

Marchi, Nina, Flavia Schlichta, and Laurent Excoffier (Mar. 2021). “Demographic inference”. Current Biology 31.6.
Mazet, Olivier, Willy Rodríguez, and Lounès Chikhi (Sept. 2015). “Demographic inference using genetic data from a single individual: Separating population size variation from population structure”. Theoretical Population Biology 104, pp. 46–58.
Stanford, Kyle (2023). “Underdetermination of Scientific Theory”. The Stanford Encyclopedia of Philosophy. Ed. by Edward N. Zalta and Uri Nodelman. Summer 2023. Metaphysics Research Lab, Stanford University.

4. Aywlwyn Scally (Associate Professor, University of Cambridge)
Models and their misinterpretation in population genetics

Models are ubiquitous in science, but perhaps less so in biology than in other fields. As a result, models in biology are sometimes less well understood, both by practitioners and the wider public. There is often confusion about their purpose and about the distinctions between model, data and reality.
The mathematical framework of population genetics is perhaps the most prominent example of a theoretical model in biology, and is central to how we study and interpret the differences between genomes and how they have evolved. I will discuss the nature of models in population genetics, particularly in the context of human evolution, and the consequences of their misinterpretation, both for researchers and for the public understanding of human genetics and biology.

Is Philosophy of Science in Practice the Best Way to Realize the Science-Philosophy Continuum?

Thomas Pradeu
ImmunoConcept, CNRS & Univ. Bordeaux
thomas.pradeu@u-bordeaux.fr

James Ladyman
University of Bristol
James.Ladyman@bristol.ac.uk

Héloïse Athea
Gustave Roussy & IHPST
heloise.athea@gustaveroussy.fr

Marie Kaiser
Bielefeld University
kaiser.m@uni-bielefeld.de

Session 'Philosophy of science in practice' (PSP) advocates direct engagement with the actual practices, methods, and contexts of scientific work rather than focusing exclusively on formal logical structures or abstract theories of scientific knowledge. Building on earlier key figures such as Thomas Kuhn and Ian Hacking, PSP is often explicitly contrasted with both logical empiricism and purely theory-focused approaches, which it critiques for creating idealized and inaccurate accounts of how science actually operates. Central themes include attention to experimental practice, the material culture of science (instruments, specimens, laboratories), the social organization of research, and case studies drawn from actual scientific episodes rather than hypothetical examples.
In this session, we propose to reconceptualize PSP as one realization of the science-philosophy continuum. At least three key features of PSP support this approach: its commitment to empirical engagement, its pursuit of relevance to practicing scientists, and its emphasis on direct collaboration with scientists. We ask to what extent PSP contributes to rethinking this continuum, how it addresses this question, and how it compares with other HPS approaches concerned with exploring and enacting this continuum. Through four talks by scholars who hold quite distinct views regarding the science-philosophy continuum, we will compare PSP with 'natural philosophy' (Ladyman), 'philosophy in science' (Pradeu), 'philosophy in the lab' (Laplane and Athea), and 'collaborative interdisciplinarity' (Kaiser). We argue that situating PSP within these various trends offers one of the most productive ways to understand the distinctive features and contributions of PSP.

Natural Philosophy: Science in Philosophical Practice and Philosophy in Scientific Practice
James Ladyman

The relationship between science and philosophy is contested. Some philosophers regard science as irrelevant to the philosophical questions they address, and some seek to reform science in line with their philosophical agenda. Even among philosophers of science there are very different attitudes to science and its history, with some attempting to rationally reconstruct the history of science, and some attempting to deconstruct narratives of progress and the accumulation of knowledge. Retrospectively, the logical empiricists are viewed as scientific philosophers attempting to reform philosophy by making it scientific by conceptual

and logical analysis and the repudiation of metaphysics. In current philosophy there are philosophers who take their work to be continuous with science, and there is a society for the metaphysics of science, a society for philosophy of science in scientific practice, and the proposal of philosophy in experimental science in the senses of philosophers in the laboratory. The requirements of the division of labour and the need for disciplinary identities and intradisciplinary branding in the academic political economy aside, what is the difference between doing philosophy and doing science? This paper argues that while philosophy and science are often separate there natural philosophy can be done by both scientists and philosophers. The need for natural philosophy is illustrated with the case of life.

"Philosophy in science vs. philosophy of science in practice: antagonism or complementarity?"
Thomas Pradeu

'Philosophy in science' (or 'PinS' for short) is a recently described trend within history and philosophy of science that consists in using philosophical tools to address a scientific question (Pradeu et al. 2024). In other words, PinS tries to contribute to the resolution of scientific questions, not primarily philosophical ones. For this, PinS uses the five classical tools of conceptualization, critique, contextualization, consistency, and connection. Several examples of PinS work done in the last five decades have been documented. Some key lessons of this analysis are that: i) The vast majority of papers in philosophy of science do not belong to PinS (addressing a scientific question is simply not their goal); ii) All instances of PinS are instances of scientifically informed philosophy of science, but the reverse is not true; iii) Some (rare but interesting) examples of PinS work (e.g., Sober in philosophy of biology, Malament in philosophy of physics) have had a strong impact on science and scientists. Given these features, PinS and philosophy of science in practice (PSP) seem, at least intuitively, to have quite a lot in common. Crucial similarities include acquaintance with cutting-edge science, a perspective informed by the history of the scientific discipline under consideration, close collaborations with scientists, embedment in scientific labs, and a care for how science is done in actual practice (Chang 2011; Ankeny et al. 2011). One major potential difference is that PinS focuses less on describing scientific activities than on producing science (Pradeu 2025). In other words, PinS tends to be interventionist and to aim at producing science via the five aforementioned tools. This dimension brings PinS closer to another trend in HPS, namely "complementary science" (Chang 1999), although the latter is more focused on past science and on questions that current scientists tend to neglect. In this talk, I propose to use qualitative and bibliometric arguments to offer the first systematic comparison between PSP, complementary science, and PinS. I will draw conclusions from this comparison about how we should connect philosophy and science both in research and, crucially, in education.

References:

Ankeny R, Chang H, Boumans M, Boon M (2011) Introduction: philosophy of science in practice. Euro Jnl Phil Sci 1:303–307. https://doi.org/10.1007/s13194-011-0036-4
Chang H (2011) The Philosophical Grammar of Scientific Practice. International Studies in the Philosophy of Science 25:205–221. https://doi.org/10.1080/02698595.2011.605244
Chang H (1999) History and Philosophy of Science as a Continuation of Science by Other Means. Science & Education 8:413–425. https://doi.org/10.1023/A:1008650325798

Pradeu T (2025) ‘Philosophy in science’, a distinctive approach combining philosophy and science. In: Veigl S, Currie A (eds) Methods in the Philosophy of Science: A User’s Guide. MIT Press, pp 271–288
Pradeu T, Lemoine M, Khelfaoui M, Gingras Y (2024) Philosophy in Science: Can Philosophers of Science Permeate through Science and Produce Scientific Knowledge? The British Journal for the Philosophy of Science 75:375–416. https://doi.org/10.1086/715518

PhilLabo: Philosophy in the Laboratory
Héloïse Athéa & Lucie Laplane

In philosophy of science, it is now commonplace to claim that science and philosophy lie on a continuum. Yet, in practice, this continuum is largely one-sided. Scientists routinely engage in conceptual and theoretical work, whereas philosophers of science almost never engage in hands-on experimental practice. They typically reconstruct science from texts, talks, and datasets, as if theoretical and experimental dimensions could be cleanly separated - something that immediately appears illusory once one spends time at the bench. In this presentation, we develop and defend PhiLabo: an approach in which philosophers of science themselves perform experiments in laboratories and treat experimentation as a philosophical tool, not merely an object of reflection.
PhiLabo builds on two recent trends. First, “Philosophy in Science” (PinS) argues that philosophy can and should contribute directly to ongoing scientific research. Second, Hasok Chang’s notion of “complementary science” shows how historically and philosophically informed experimentation can recover forgotten results, extend existing experimental repertoires, and thereby produce new scientific knowledge. PhiLabo combines these trends: it treats experimental biology as one of the legitimate methods of philosophy, while aiming to contribute to mainstream science rather than only to a marginal, non-orthodox domain.
We articulate four epistemic benefits of experimental practice for philosophers:
(1) Experiments can directly generate new empirical knowledge, for instance by testing philosophical hypotheses.
(2) Experiments provide an irreducible source of evidence about laboratory practice, including tacit knowledge and under-described procedures that cannot be reconstructed from publications alone.
(3) Trying (and sometimes failing) to replicate historical or contemporary experiments raises new questions that were not visible from the armchair, thereby reshaping philosophical agendas.
(4) Doing experiments offers a more realistic understanding of how scientific reasoning actually proceeds, including the intertwinement of conceptual, technical, and institutional constraints.
We then argue for two further, more programmatic claims. First, engagement at the bench refines the scientific relevance of philosophical work by forcing prioritization: when one is embedded in an experimental project, some questions clearly emerge as pivotal for practice while others become less salient. Second, PhiLabo enhances the accessibility of philosophy for scientists: abstract conceptual proposals are anchored in manipulable biological systems, experimental readouts, and concrete design choices, making them easier to integrate into ongoing research programs.
We conclude by drawing the institutional and disciplinary consequences of taking PhiLabo seriously. If experimentation is accepted as a legitimate philosophical method, then philosophical research programs, funding schemes, hiring criteria, and training pathways must be reorganized to include laboratory work as part of what it is to “do philosophy of science.” Rather than remaining seated in the proverbial armchair, philosophers can - and, in

some cases, should - enter the “hell” of the laboratory, where the conceptual and the experimental are inseparable and where an enhanced philosophy of science can emerge.

How Interdisciplinarity Works in Practice: Lessons from Collaborations Between Biology and Philosophy
Marie I. Kaiser

Calls for philosophy in science emphasize the value of philosophers contributing to scientific inquiry (Pradeu et al. 2024), yet we still know relatively little about how such collaborations actually function on the ground. Much existing work on interdisciplinarity offers typologies or high-level analyses (Boden 1999; Repko 2008), but it rarely examines the concrete dynamics through which scientists and philosophers jointly produce knowledge. This talk advances a practice-based understanding of philosophy-science collaboration by presenting a detailed, reflexive case study of interdisciplinary research between empirically working biologists and philosophers of biology investigating individualization in animals.
Drawing on close analysis of co-authored publications, collaborative discussions, and joint conceptual-empirical problem-solving, I argue that interdisciplinarity in practice takes two distinct but often intertwined forms. The first involves integrating extant disciplinary epistemic elements – biological concepts, empirical findings, methods, and philosophical analyses – into coherent interdisciplinary products (Huutoniemi et al. 2010). The second, less often recognized, involves a deeper mode of collaboration in which researchers jointly pursue shared research questions and develop novel epistemic products through sustained discursive engagement. In such cases, the interdisciplinary team effectively becomes a “plural subject” of inquiry (sensu Gilbert 2006), generating ideas that neither discipline could have produced alone.
To show how these forms of epistemic integration unfold, I introduce a four-dimensional analytic framework that examines:
(1) Objects of integration: which knowledge elements (e.g., concepts, research questions, methodological assumptions) are being integrated;
(2) Activities of integration: the epistemic actions through which integration is accomplished (e.g., joint conceptual work, iterative figure-building, clarification of implicit assumptions);
(3) Communication: how tailoring research to philosophical or biological audiences shapes not only presentation but the integrative process itself (Pohl et al. 2015); and
(4) Conditions of collaboration: the institutional, organizational, and individual factors that enable or constrain interdisciplinary work (Lyall et al. 2013).
Analysing these four dimensions reveals that communication and audience expectations are not merely external constraints but constitutive of the integrative process: choices about writing style, argumentative norms, the formality of definitions, or the use of visual representations influence which interdisciplinary insights can emerge and how they stabilize. Equally crucial are the enabling conditions of collaboration, such as shared funding structures, research-group organization, and individual epistemic virtues such as humility, openness, and flexibility (Valles 2018, Chapter 8). Moreover, my analysis shows that interdisciplinary collaborations succeed when researchers negotiate common ground, reciprocally adjust disciplinary expectations, and build shared conceptual and methodological repertoires. Reflecting analytically on such collaborations “from the inside” reveals hidden epistemic labor, tacit assumptions, and the mechanisms through which genuinely interdisciplinary knowledge is produced.
I conclude by suggesting that detailed, practice-based case studies of this sort can enrich philosophy of science in practice by offering both conceptual insights into the nature of

interdisciplinarity and concrete guidance for designing, sustaining, and evaluating future philosophy-science collaborations.

References:

Boden M (1999) What is interdisciplinarity? In: Cunningham R (ed) Interdisciplinarity and the organization of knowledge in Europe. Office for Official Publications of the European Communities, Luxembourg, pp 13–24.
Gilbert M (2006) A theory of political obligation: Membership, commitment, and the bonds of society. Clarendon Press, Oxford.
Huutoniemi K, Klein JT, Bruun H, et al (2010) Analyzing interdisciplinarity: Typology and indicators. Res Policy 39(1):79–88. https://doi.org/10.1016/j.respol.2009.09.011
Lyall C, Bruce A, Marsden W, et al (2013) The role of funding agencies in creating interdisciplinary knowledge. Sci Public Policy 40(1):62–71. https://doi.org/10.1093/scipol/scs121
Pohl C, Wülser G, Bebi P, Bugmann H, et al (2015) How to successfully publish interdisciplinary research: Learning from an ecology and society special feature. Ecol Soc 20(2):23. https://doi.org/10.5751/ES-07448-200223

Applying Knowledge in Practice: Epistemology, Methodology, and Legitimacy

Karoliina Pulkkinen
University of Helsinki
karoliina.pulkkinen@helsinki.fi

Mieke Boon
University of Twente
m.boon@utwente.nl

Henk-Jan van den Brink
University of Twente
h.j.vandenbrink@utwente.nl

"I often wish this phrase, 'applied science,' had never been invented. For it suggests that there is a sort of scientific knowledge of direct practical use, which can be studied apart from another sort of scientific knowledge, which is of no practical utility, and which is termed 'pure science.' But there is no more complete fallacy than this. What people call applied science is nothing but the application of pure science to particular classes of problems." – Thomas H. Huxley, "Science and Culture" (1880)

Philosophy of Science in Practice must confront fundamental questions about how scientific knowledge relates to application, technology, and engineering. While the pure/applied distinction has received considerable attention historically (Mayr 1976, Kline 1995), its persistence in scientific culture and institutions demands renewed philosophical scrutiny. This symposium examines application not as a straightforward derivative of "pure" knowledge, but as a complex epistemological, methodological, and political phenomenon requiring careful analysis of actual practices. Our three papers approach application from complementary angles, together revealing its multifaceted nature.

Boon's contribution challenges the traditional view that application simply follows from theory by demonstrating how conceptual modeling—a methodology oriented toward producing applicable knowledge—operates as a full-fledged scientific methodology on par with the hypothetico-deductive method. Through analyzing research articles across engineering fields, she shows that researchers construct conceptual models as epistemic tools incorporating multiple ingredients beyond theory: target phenomena, contextual information, measurement methods, and values. This methodology, grounded in an engineering paradigm of science, reveals that producing applicable knowledge requires distinct epistemic practices from the outset, not merely at a final "application stage."

Van den Brink's paper provides empirical evidence for the epistemic consequences of orienting science toward applications through ethnographic research in biomass chemistry. His analysis reveals that application-orientation is neither epistemically costless nor simply detrimental. While external goals like sustainability actively guide research design—from model selection to experimental regimes—they also constrain the pursuit of internal goals like detailed mechanistic understanding. This creates genuine trade-offs: applicable knowledge is produced, but at the cost of slower epistemic progress on traditional internal goals. His work demonstrates that normative assessment of application-oriented science depends crucially on one's account of scientific progress and cannot be settled by general philosophical arguments alone.

Pulkkinen's contribution shifts attention to the political and rhetorical dimensions of application by examining how the pure/applied distinction functioned in Soviet science. Focusing on physical chemist Nikolai Semenov's efforts to secure a foothold for "chemical physics," she traces how more theoretical approaches evolved from being "maids" to industry into "mothers" of new technologies. This transformation reveals that application serves not only epistemic purposes but also legitimation functions for political projects, both domestically and internationally. Her analysis introduces previously unexplored perspectives on the internal/external distinction by showing how Soviet scientists and philosophers conceptualized it through the science/industry relation.
Together, these papers demonstrate that understanding application in practice requires attention to methodology (how applicable knowledge is produced), epistemology (what trade-offs application-orientation entails), and politics (how application rhetoric functions in legitimation). Rather than treating application as a secondary concern or accepting the pure/applied distinction at face value, Philosophy of Science in Practice must recognize application as central to scientific knowledge production itself—with profound implications for understanding research practices, evaluating scientific progress, and assessing science's role in society.

Epistemic and methodological consequences of orienting science towards applications: The case of biomass chemistry in the context of the chemical industry
H.J. van den Brink
University of Twente, Enschede, The Netherlands
h.j.vandenbrink@utwente.nl

In contemporary society, science is expected to be societally relevant, that is, it should not only contribute to internal goals such as explaining and predicting phenomena, but also to external aims such as sustainability, efficiency and frugality (Houkes, 2016). Despite its prevalence in scientific practice, the epistemic consequences of this re-orientation towards applications have received little attention in the philosophy of science. This prevents normative evaluation of the application-orientation turn in science. To investigate this, I carried out ethnographic research on biomass chemistry in the context of the chemical industry. By chemically investigating complex biomass matrices, this scientific research yields knowledge and methods that are usable for engineering design and in the industrial context.
Specifically, I describe how external goals affect the selection of physical models to be experimentally studied, the selection of experimental regimes, the development of analytical methods, and the construction and validation of mechanistic models. In these cases, the researchers can actively direct their research efforts towards applications by letting the external goals constrain and guide the internal aims of the research. I show that the pursuit of external goals does not exclude the pursuit of internal goals. Instead, the internal goals, such as mechanistic explanations and understanding, are a prerequisite to achieving the external goals. Therefore, the most pessimistic epistemological scenario of the re-orientation towards applications – that the pursuit of external aims always diminishes the pursuit of internal aims – is not warranted.
However, our ethnographic research reveals the methodological challenges of application-oriented science. As a consequence, the biomass constituents, reaction pathways and products are characterised and described in less detail. At least, compared to when the research would have focused on pure lignin or hemi-cellulose fractions instead. It follows that the application-orientation incurs a cost for the internal goals of research. Specifically, there is less epistemic progress with regard to the internal goals. Therefore, the most optimistic

epistemological scenario – that the pursuit of external goals does not hinder the pursuit of internal goals – is not warranted either.
As such, this study helps practising scientists and science policy makers understand the opportunities and limitations of application-oriented science, which goes beyond the optimistic and exclusively pessimistic scenarios. In chemical research, such as in our biomass case study, re-orienting science towards applications is possible, but at the cost of epistemic progress on internal goals. How one assesses the appropriate trade-off depends on one's philosophical account of scientific progress. On the epistemic and noetic accounts of scientific progress (Bird, 2007; Dellsén, 2016), application-oriented science is not desirable because it slows the accumulation of scientific knowledge and understanding. On the other hand, on a functional account of scientific progress (Douglas, 2014; Houkes, 2024), application-oriented science is much more progressive because it enables us to predict and intervene in the world beyond experimental contexts. However, the ability to combine internal and external goals likely depends on domain-specific characteristics such as methods, types of explanations, and the particular external goals. Therefore, general normative assessments of the application re-orientation in science depend on the investigation of individual scientific domains and their characteristics.

References

Bird, A. (2007). What Is Scientific Progress? Noûs, 41(1), 64–89.
Dellsén, F. (2016). Scientific progress: Knowledge versus understanding. Studies in History and Philosophy of Science Part A, 56, 72–83. https://doi.org/10.1016/j.shpsa.2016.01.003
Douglas, H. (2014). Pure science and the problem of progress. Studies in History and Philosophy of Science Part A, 46, 55–63. https://doi.org/10.1016/j.shpsa.2014.02.001
Houkes, W. (2016). Perovskite Philosophy: A Branch-Formation Model of Application-Oriented Science. In M. Franssen, P. Vermaas, P. Kroes, & A. Meijers (Eds.), Philosophy of Engineering and Technology (Vol. 23, pp. 195–218). Springer Nature. https://doi.org/10.1007/978-3-319-33717-3_12
Houkes, W. (2024). Progress in Application-Oriented Research. The Harvard Review of Philosophy, 31, 59–74. https://doi.org/10.5840/harvardreview202522560

From a Maid of Industry to its Mother: a Closer Examination of Output Legitimacy of Science and Technology

Karoliina Pulkkinen
University of Helsinki
karoliina.pulkkinen@helsinki.fi

Policy-relevant research is in a position to impact the lives of the broader public. Correspondingly, concerns of its political legitimacy have become a pressing question. To remedy potential deficits of political legitimacy, some authors have considered increasing democratic input of stakeholders of research (e.g., Intemann 2015, Schroeder 2017, Lusk 2021). Turning attention to nondemocratic contexts, however, reveals that outputs produced by political actors have been used to legitimise political power. This talk explores how science and technology have been used as sources for output legitimacy for political actors and political projects. Existing historical research into Soviet science and technology suggests that especially technology could serve this legitimation function (Bailes 1976, 1978). This gives rise to the question of whether less-applied fields served similar purpose. Through a close engagement with histories of Soviet physical sciences and technology, my talk explicates how more theoretically oriented fields and approaches found a foothold in a

science system that strongly supported applied research. By examining the arguments given by physical chemists to support their approaches, I map out how more theoretical approaches to chemistry went from being mere helpers of technological and industrial projects to become the initiators of new technologies and industries. To show this, I examine the efforts of Nikolai Nikolayevich Semenov, the first Soviet citizen to win a Nobel prize, who was known for his work on chain reactions, chemical kinetics and reactivity. By focussing on Semenov's attempts to secure a position for "chemical physics" within the Soviet science system, this talk brings attention to his strategies to gain recognition and patronage for more theoretical approaches in a research landscape that was strongly application-oriented. By doing so, I track the development and growing authority of fields that were conceived as less applied and argue that they gained the function to legitimise the Soviet political project on the international arena. This suggests that political legitimacy of science and technology was not solely a domestic affair but had an international dimension, too.
Aside from the above, a closer look at both theoretical and applied fields in the Soviet Union allows to introduce previously unexplored perspectives on the applied/theoretical discussion. In particular, I will show that Soviet scientists and philosophers often conceptualised internal/external distinction in terms of science/industry relation. Although initially, more theoretical approaches were conceived as helpers or "maids" to industry, by the 1960s, they had regained a status as their "mother", which problematised the arguments of vulgar materialists who saw science subservient to industry. This indicates that Soviet discussions on internal/external (or science/industry relation) went hand-in-hand with the growing legitimization function of more theoretical approaches to science, allowing for a new dimension for existing reflections regarding the internal/external distinction (Shapin & Schaffer 2011/1985; Shapere 1986).

References

Bailes, K. E. (1976). Technology and Legitimacy: Soviet Aviation and Stalinism in the 1930s. Technology and Culture, 17(1), 55–81. JSTOR. https://doi.org/10.2307/3103253
Bailes, K. E. (1978). Technology and society under Lenin and Stalin: Origins of the Soviet technical intelligentsia, 1917-1941. Princeton University Press.
Intemann, K. (2015). Distinguishing between legitimate and illegitimate values in climate modeling. European Journal for Philosophy of Science, 5(2), 217–232. https://doi.org/10.1007/s13194-014-0105-6
Lusk, G. (2021). Does democracy require value-neutral science? Analyzing the legitimacy of scientific information in the political sphere. Studies in History and Philosophy of Science Part A, 90, 102–110. https://doi.org/10.1016/j.shpsa.2021.08.009
Schroeder, S. A. (2017). Using Democratic Values in Science: An Objection and (Partial) Response. Philosophy of Science, 84(5), 1044–1054. Cambridge Core. https://doi.org/10.1086/694006
Shapere, D. (1986). External and Internal Factors in the Development of Science. Science & Technology Studies, 4(1), 1–9.
Shapin, S., & Schaffer, S. (2011/1985). Leviathan and the air-pump: Hobbes, Boyle, and the experimental life (Revised edition., Vol. 32). Princeton University Press.

How is it Possible to Apply Knowledge? Conceptual Modelling as Scientific Methodology

Mieke Boon
University of Twente, Enschede, The Netherlands
Philosophy
m.boon@utwente.nl

Kant famously reframed the debate between rationalism and empiricism by asking: How is knowledge possible? This inquiry shifted philosophical attention from methodological concerns to the fundamental conditions enabling human knowledge. Today, while much philosophy of science centers on scientific evidence and hypothesis justification, a parallel question in Kant's spirit remains underexplored: How is it possible that scientific knowledge can be applied? How do humans produce knowledge that is practically usable—for designing technologies (to produce specific bio-chemical-physical phenomena or functionalities), creating reliable interventions, or solving real-world problems?

Nancy Cartwright's claim in How the Laws of Physics Lie (1983) revealed that strict deductive application of theories like Newtonian mechanics rarely yields accurate predictions for concrete phenomena. She emphasized models as indispensable bridges between theory and accurate descriptions of reality. This insight gained traction in Models as Mediators (Morgan & Morrison, 1999), positioning models at the center of philosophical inquiry and sparked ongoing debates about model-target representation.

Building on this tradition, my paper addresses a new dimension: How is it possible to apply models for epistemic purposes? Boon and Knuuttila (2009) proposed viewing models as epistemic tools—artifacts constructed to perform specific epistemic tasks. Beyond theoretical principles, models incorporate contextual and practical information that enables their use by epistemic agents. Models are thus not mere instantiations of theory but complex artifacts designed for conducting epistemic tasks.

Through extensive analysis of original scientific articles—ranging from Carnot's essay on the ideal heat engine (Knuuttila and Boon 2011) to contemporary research in chemical, biomedical, mechanical, and materials engineering—I have developed an account of scientific methodology centered on model construction and reconstruction (e.g., Boon 2020). This methodology moves away from the traditional hypothetico-deductive framework toward a modeling framework that better reflects actual research practices. Scientific models typically incorporate multiple ingredients: references to broader societal relevance, epistemic purposes, specific target phenomena, relevant bio-chemical-physical-technological aspects affecting the phenomenon, and methods for measuring these aspects and their effects. My empirical findings demonstrate that scientific articles across these fields can be comprehensively analyzed using this standard set of ingredients.

Crucially, my analyses reveal that conceptual modeling serves as the starting point for scientific research. Before applying abstract formulas to construct mathematical models, researchers must first develop rich conceptual models of their target systems. For instance, modeling a Newtonian system requires conceptualizing it through Newtonian physics concepts before mathematical calculations become possible.

The epistemological foundations of this methodology align with an engineering paradigm of science, where the primary aim is to produce knowledge for real-world applications (Boon and Van Baalen 2019). My elaboration emphasizes that conceptual modeling not only fosters conceptual understanding (cf. Rouse, 2019), but also systematically incorporates what is often referred to as "context"—including disciplinary frameworks, instruments, societal conditions—together with epistemic, societal, and ethical values (cf. Longino, 1990) for assessing the quality of the model. Crucially, this methodology is not algorithmic; it demands that human researchers actively create and employ meaningful concepts and make deliberate

choices about which ingredients to include and how to combine them. These decisions, coupled with professional responsibility, shapes a robust notion of epistemological responsibility.

Through this epistemological underpinning, conceptual modeling emerges as full-fledged and workable scientific methodology. It explains how knowledge can be constructed for application, thereby addressing the question: How is it possible to apply knowledge? Recognizing conceptual modeling as a core methodology represents a significant shift, with profound implications for understanding research practices, advancing science education, and opening new avenues in the philosophy of science.

References

Boon, M., & Knuuttila, T. (2009). Models as epistemic tools in engineering sciences. In Philosophy of technology and engineering sciences (pp. 693-726). North-Holland.

Boon, M., & Van Baalen, S. (2019). Epistemology for interdisciplinary research–shifting philosophical paradigms of science. European journal for philosophy of science, 9(1), 16.

Boon, M. (2020). The role of disciplinary perspectives in an epistemology of scientific models. European journal for philosophy of science, 10(3), 31.

Cartwright, N. (1983). How the laws of physics lie. OUP Oxford.

Knuuttila, T., & Boon, M. (2011). How do models give us knowledge? The case of Carnot's ideal heat engine. European journal for philosophy of science, 1(3), 309.

Longino, Helen E. 1990. Science as Social Knowledge: Values and Objectivity in Scientific Inquiry. Princeton: Princeton University Press.

Morgan, M. S., & Morrison, M. (Eds.). (1999). Models as mediators: Perspectives on natural and social science (No. 52). Cambridge University Press.

Rouse, J. (2019). Articulating the world: Conceptual understanding and the scientific image. University of Chicago Press.

Paleosciences Beyond Earth: Evidence, Narrative, and (the Limits of) Analogy

Franziska Reinhard
Ulm University
franziska.reinhard@uni-ulm.de

Siyu Yao
University of Cincinnati
syyaogy@icloud.com

Matthew Brewer
Boston University
brewer@bu.edu

The paleosciences are no stranger to difficult epistemic situations. Whether in geology, paleontology, or archaeology, historical scientists regularly infer long-past events from fragmentary or ambiguous traces, drawing on a host of indirect methods, models, or analogies. Most philosophical work on these practices has focused squarely on Earth-based historical sciences. Much less attention has been paid to paleoscientific investigations beyond Earth, even though much of extraterrestrial research in fact targets processes, events, or entities in the deep past.

"Extraterrestrial paleosciences"–from cosmology and astronomy, to planetary geology and astrobiology–deserve philosophical consideration, first, because they face greater epistemic challenges than their terrestrial counterparts. Many standard approaches of historical science research–revisiting sites, gathering additional contextual evidence, iteratively refining interpretation–are far more constrained or unavailable when investigating other planets or ancient cosmic events. Extraterrestrial evidence is harder to come by, strongly mediated by instrumentation, and often only available at coarser spatial and temporal resolution. Uniformitarian assumptions that underpin inferences on Earth, become less reliable when dealing with celestial bodies shaped by unfamiliar chemistries, climates, and geological processes. These constraints do not just make extraterrestrial paleosciences "more difficult", they expose features of historical science research taken for granted in the terrestrial context, thereby offering a testing ground for evaluating and extending existing philosophical accounts.

Second, despite their challenges, extraterrestrial paleosciences complement terrestrial research by expanding its epistemic resources. Extraterrestrial systems provide additional instances of planetary formation and evolution, climate histories, and geological processes that can serve as analogues, surrogates, or models for reconstructing the Earth's or our Solar System's deep past. Hence, extraterrestrial paleosciences are an important piece of the puzzle for reconstructing our Earth's history. Focusing on how scientists use analogies, narratives, and models in these settings opens up the opportunity to refine existing philosophical accounts by showing where familiar epistemic tools succeed, and where they require modification when knowledge is transferred from extraterrestrial to terrestrial contexts.

This session brings together three talks focusing on a diverse set of research practices in different extraterrestrial paleosciences. Franziska Reinhard looks into the difficulties facing (micro-)paleontological investigations on other planets, focusing on the current Perseverance mission on Mars. She argues that one of its main challenges compared to terrestrial investigations is the lack of opportunities for iteration and analyses the practices developed and employed to compensate for it. Siyu Yao examines how exoplanetary studies generate

generalizable knowledge about the evolution of planetary systems despite severe observational constraints. Analyzing the case of Hot Jupiter, she argues that narrative construction—stitching together models and mechanisms to form how-possible explanations—generates an epistemic toolkit that enables transferable knowledge about planetary formation and migration, reshaping our understanding of the Solar System's history. Matthew Brewer considers the role of taxonomies and analogical reasoning in planetary science. He highlights a recent debate about whether to classify the dynamics of the Jovian moon Europa as similar to Earth's plate tectonics. He argues that when studying the histories of lesser-known celestial bodies, non-unique classifications will be more adequate, for the sake of the analogical comparisons that they warrant.

Martian-Leopard Spots, Iteration, and One-Shot Science
Franziska Reinhard (franziska.reinhard@uni-ulm.de)
Humboldt Centre for Philosophy, Ulm University

In July 2024, NASA's Perseverance rover discovered something exciting on Mars. While exploring an ancient Martian riverbed, it spotted bright patches of rock surrounded by dark ring-like patterns–quickly dubbed "Leopard Spots". On Earth, such patterns can result from redox reactions associated with microbial activity, and Perseverance even detected organic molecules in some of the patches (Hurowitz et al. 2025). Still, whether all this signals the presence of past Martian life remains very much an open question. While the search for Martian early life resembles terrestrial (micro-)paleontology in some ways, it faces additional constraints. And though some philosophical work has gone into scrutinizing what counts as a 'biosignature' (Malaterre et al. 2023; Gillen et al., 2023), the practical and technological challenges and solutions that underlie investigations like the Perseverance mission have been less discussed by philosophers of science. My argument in this talk is that one such central challenge is their relative lack of iteration, and that much of the scientific practices surrounding Perservance and other missions aim to reintroduce iterativity, either by 'front-loading' it before the mission or by 'simulating' it during operations.

Current "Martian paleoscience" is essentially a one-shot investigation, at least when compared to its terrestrial counterparts. Once a mission has taken off, the technological level of the investigation is fixed. Once the rover collects samples from a location, returning, either before or after sample return, is effectively impossible. And while Perseverance is designed to collect multiple samples, its storing capacities have definitive limitations. Historical sciences on Earth routinely deal with non-repeatable events, but their investigations rely on various types of iterative refinement: revisiting sites and resampling with new technologies or in light of updated background knowledge, thereby building up lines of evidence and interpretations over time (Wylie & Chapman 2016; Currie 2018). This kind of iterativity is much more difficult to achieve for extraterrestrial locations.

Strategies to deal with this challenge aim to restore something like iterativity in the course of the Perseverance mission and fall into two categories. One set of practices basically front-loads iterativity by calibrating expectations and detection strategies before the mission begins. This is done through Earth-based analog fieldwork and laboratory simulations that test and refine our understanding of potential biosignature preservation, instrument responses, and the limits of Earth-Mars analogical reasoning (Cockell et al. 2019; McMahon & Cosmidis 2022). The second set of practices simulate iterativity directly on Mars so-to-say. This involves (i) pre-sampling context capture: the rover's instruments take high-resolution panoramas, multispectral scans, provenance data before a sample is collected, maximizing the informational value of each sample. And (ii) the design and implementation of sequential and adaptive sampling instruments that allow sampling decisions to be updated in real time,

mimicking iterative evidential processes in a constraint window of opportunity (Lawson et al. 2025). These strategies work by shifting a lot of the epistemic load to planning, instrument design, and pre-sampling documentation–processes that are also important to terrestrial paleosciences but whose contribution becomes particularly salient when looking at extraterrestrial investigations.

From Hot Jupiter to Jupiter: Narratives, Knowledge-How, and Exoplanetary Science as Epistemic Resource

Siyu Yao (syyaogy@icloud.com)
Department of Philosophy, the University of Cincinnati

Exoplanetary science can be considered a historical science with special challenges and opportunities. A central part lies in inferring how exoplanets were formed and evolved to their current stage. Despite limited observational access to exoplanetary systems, they provide various types of planets under different environments and developmental stages, complementing the uniqueness of the Solar System and serving as surrogates to enrich accounts of planetary evolution. However, unlike simulations and experiments, exoplanets as surrogates still face common challenges of the historical sciences. A question arises as to how to obtain generalizable knowledge to benefit the study of the Solar System. I answer this question by examining the case of the Hot Jupiter, an exoplanet that has significantly shifted the research direction of planetary evolution.

The 1995 discovery of 51 Pegasi b, aka. Hot Jupiter, posed a challenge to planetary science: contemporary planet formation models failed to explain Hot Jupiter's odd closeness to the central star. To provide a plausible account, astronomers had to revise one of their basic assumptions, that planets were formed in situ, which was intuitive from the arrangement of terrestrial and gaseous planets separated by the frost line in the Solar System. This revision also required developing alternative scenarios of planetary systems' evolution and a toolkit to investigate them.

I argue that narrative construction has been central in both investigating Hot Jupiter and obtaining generalizable knowledge. Narratives have been a key element of the historical sciences' toolkit. They are descriptions of ordered events with a beginning, middle and end structure (Roth 2017). Narratives have been identified to serve roles like organizing evidence, facilitating inferences and constructing how-possible explanations (Currie and Sterelny 2017). Beyond structuring evidence and inference in local contexts, here I also highlight how narratives can produce general, practical knowledge that can be transferred across contexts.

To explain the orbit of Hot Jupiter, scientists performed a line of investigation that is best understood as narrating. In constructing a storyline describing Hot Jupiter's migration history, scientists drew on multiple existing mathematical models that were used for other purposes and systems, adapted them to the special conditions of Hot Jupiter, stitched together their starting and ending conditions, and constructed new mechanisms to fill in gaps. I argue that this produces a set of knowledge-how about using abstract models for the specific problems of planet migration (Cartwright et al. 2022). Narrating a coherent story structures scientists' modeling practices, so that they obtain a set of models that can be used to study the history of planetary systems, and obtain knowledge about how to adjust their parameters and piece them together to form a complex storyline, what their outcome space is, and what potential supplementary mechanisms are. This produced a toolkit that later enabled the reconsideration of the Solar System's history and finally uncovered the turbulent migration history of Jupiter and Saturn. I conclude that while being historical and distant itself, exoplanetary science provides key evidential and methodological sources to mitigate the challenges of knowing the past of our closer environment.

Classification, Analogical Reasoning, and Uniformitarian Assumptions in Planetary Science
Matthew Brewer (brewer@bu.edu)
Department of Philosophy, Boston University

Earth is the only celestial body known to exhibit plate tectonics. Just as biologists face an "N = 1" problem (Earth is the only planet known to host life), planetary scientists face an "N = 1" problem about plate tectonics. In both contexts, scientists have only one example from which to draw generalizations about the phenomenon under investigation. Geoscientists thus frequently describe plate tectonics as "unique" to Earth. Indeed, for much of its history, planetary science was idiographic: our Solar System's planets were considered as individuals, with unique features and causal histories (Gould 1991).

However, recent advancements in planetary science and the discoveries of thousands of exoplanets have complicated the notion that planetary dynamics are "unique." For example, recent observations of the surface of Jupiter's moon Europa have raised the question: does Europa exhibit plate tectonics? Surface images show that Europa seems to "recycle" its icy shell in a way that resembles plate tectonics (Kattenhorn and Prockter 2014). However, some geoscientists argue that we should not classify Europa as exhibiting plate tectonics, because Europa's shell is icy, unlike Earth's silicate rock shell (Stern et al. 2018).

In response to these advancements, some planetary scientists have proposed new taxonomies, which consider plate tectonics as a non-unique phenomenon. In this paper, I argue that the new taxonomies proposed by planetary scientists best serve their current purposes because they provide a warrant to pursue analogical reasoning, and such reasoning is key to the study of lesser-known celestial bodies, such as exoplanets. Classifying plate tectonics as a non-unique phenomenon gives scientists a warrant to pursue fruitful analogies between Earth and other lesser-known celestial bodies. For instance, the non-unique classifications of Earth and Europa have warranted the pursuit of subsequent studies which have analogically compared their dynamics, yielding new insights. Howell and Pappalardo (2019) produced a model identifying potential driving forces for plate tectonics on Europa, while Okazaki et al. (2024) investigated how icy plates might behave under temperature and salinity conditions analogous to those on Europa. I argue that for the purposes of studying lesser-known celestial bodies, non-unique classifications will be more fruitful, for the sake of the analogical reasoning that they warrant.

In the final part of the paper, I consider the implications of my argument for planetary science considered as an historical science. Some planetary scientists have proposed that the "evolution" of planets follows a standard trajectory across typical regimes. Still, some maintain that plate tectonics is a non-standard deviation from the typical trajectory for rocky planets. Whether plate tectonics is assumed to be a contingent or convergent feature implicates the warrant to engage in analogical reasoning in planetary science: e.g., could the dynamics of Mars have resembled Earth's plate tectonics at some point in its history? I argue that the history of the Earth's dynamics should not be considered as atypical. By considering how plate tectonics could be a regular pattern in the cosmos, scientists are more primed to discover other celestial bodies that are (or once were) analogous to Earth.

Institutional Neutrality

Ezgi Sertler
SOCRATES, Leibniz Universität Hannover
ezgi.sertler@gmail.com

Francesco Nappo
Università Cattolica del Sacro Cuore
francesco.nappo@unicatt.it

Ann-Christin Fischer
Ruhr-Universität Bochum
ann-christin.fischer@ruhr-uni-bochum.de

Kevin Elliott
Michigan State University
kce@msu.edu

Institutional neutrality is a central yet intensely contested norm for several organizations operating at the nexus of knowledge production, policy recommendation, and society. This symposium systematically examines the viability, (dis)advantages, and ongoing criticisms of this ideal. The tension that we aim to address is common across different contexts. On the one hand, defenders of institutional neutrality highlight its crucial role for protecting scientific organizations as well as public institutions (e.g. higher education) from political pressure. This, in the case of scientific organizations, ensures that their findings serve as a legitimate basis for political action, and, in the case of universities, is supposed to enable a healthy and balanced institutional environment for research and education. On the other hand, critics stress the inevitable use of neutrality to hide value judgments, enabling institutions to deflect moral or political accountability by presenting value-laden decisions as purely objective or technical.

This session investigates this tension by drawing on three diverse case studies. The first paper analyzes the IPCC's commitment to being "policy-relevant, but policy-neutral," assessing how this institutional boundary shapes its legitimacy and its ability to catalyze policy information. The second paper explores the challenge of neutrality in the standardization of medical practices, investigating how the development and enforcement of biomedical standards for endometriosis, while appearing objective, rely on implicit value judgments that serve specific social, economic, or epistemic goals. The third paper shifts focus to the university setting, unpacking the term 'neutrality' in that context. Drawing from both philosophy of science and political philosophy, the final paper investigates the problems that accompany the term neutrality, its epistemic and political commitments, and the concerns that the term articulates for academic governance.

This symposium advances the current state of discussions in two ways. On a methodological level, it will help draw connections between neutrality in the case of scientific organizations and in the case of public institutions, thereby identifying ways in which the respective literatures in philosophy of science and political philosophy can inform one another. On a more substantive level, the presentations will collectively discuss institutional neutrality as a political and epistemic strategy, rather than an achievable state. Future research should provide a rigorous framework for evaluating not whether institutions are neutral, but rather for whom and for what purposes they claim neutrality, and what consequences follow from that claim in terms of legitimacy, democratic accountability, and impact.

Paper 1: Whose pathways? Evaluating ensembles of emission scenarios (Author 1)

Working Group III (WGIII) of the Intergovernmental Panel on Climate Change (IPCC) regularly assesses collections of emission scenarios submitted by the scientific community, linking long-term projections about socio-economic conditions and technology rollout with the changing climate. Current protocols for assembling and evaluating submitted scenarios have been shaped by the IPCC's institutional position at the science-policy interface, particularly its commitment to delivering a 'policy-relevant but policy-neutral' scientific assessment. WGIII's latest assessments adopt an unstructured ensemble approach, where emission scenarios from different models are collected and assembled with no pre-determined design (Guivarch et al. 2021). Unstructured protocols effectively dispense WGIII authors from making several value-laden choices in ensemble construction, such as whose emission pathways to assess and which research questions are worth pursuing through scenarios. At the same time, the large number of submitted scenarios allows WGIII experts to address an important array of questions about climate change mitigation options.
This talk will point out that, while reducing WGIII experts' involvement in ensemble construction, unstructured protocols also limit opportunities for drawing policy-relevant information from submitted scenarios in distinctive ways. Specifically, the two persisting issues that this talk will address are missing variety, viz., insufficient representation of the relevant space of uncertainty (Giarola et al. 2024), and missing linkages, viz., insufficient elaboration of the regional and country-level impacts of global mitigation pathways. Because current unstructured protocols play a crucial role in securing WGIII's conception of a policy-neutral ensemble assessment, ultimately what emerges is a tension between the aims of policy-neutrality and policy-relevance. The examples provided will help illustrate that current ensemble construction protocols serve some policy purposes at the expense of others, implying that not all intended users of WGIII's assessments are equally well-served and that, to achieve a greater balance, a reform of the unstructured approach may become necessary.
In the final part, this talk will briefly review proposals for reform of WGIII's ensemble construction workflow. The options range from internal reforms, which give an entirely different meaning to WGIII's notion of a 'policy-neutral' assessment (Peters et al. 2023), to more conservative proposals, which aim to address persisting issues while preserving the design choices that inspires current ensemble construction protocols (Pirani et al. 2024). While leaning on the side of reform, a point that this paper will insist upon is that the choice between the different options is not merely one of philosophical principles, but of facts. There are risks involved in abandoning WGIII's unstructured approach, which any proposed reform for WGIII's ensemble assessment ought to consider. These risks derive specifically from the politically contested nature of climate economic modeling, on the one hand, and by IPCC's status as a distinctive actor at the science-policy interface, on the other. A comprehensive perspective is required, which takes into account both the epistemic and the institutional aspects of the problem of ensemble design.

Paper 2: Ignorance by Design: Endometriosis and the Injustice of Scientific Standards (Author 2 & 3)

This paper examines how scientific standards, rather than the values of individual scientists, can systematically privilege certain values while undermining others. While values are woven into scientific practice itself, standards operate as formal criteria that must be met to signal epistemic quality. Using endometriosis research as a case study, we argue that standards for diagnostics or funding in science often operate not as neutral measures of

epistemic quality but as mechanisms that can sustain epistemic injustice and institutionalize ignorance.

Debates around values in science have largely focused on the moral, social, or political values that individual scientists bring into their reasoning and decision-making. Yet this focus risks obscuring how institutionalized standards—those collectively stabilized rules, norms, or conventions that define what counts as legitimate knowledge, good evidence or proper method—direct scientific work in value-laden ways. When standards are treated as epistemically neutral, they can covertly serve existing social hierarchies and epistemic exclusions.

Drawing on Nancy Tuana's (2004; 2006; Tuana and Sullivan 2007) work on ignorance and the literature on epistemic injustice, we argue that ignorance about endometriosis is a predictable outcome of specific biomedical standards rather than mere oversight. First, diagnostic standards systematically privilege visible pathology over subjective reports of pain. Because endometriosis is characterized by variable, cyclical pain, these norms lead to diagnostic delays, dismissal of patients' experiences, and the reinforcement of gendered biases in clinical practice. Second, the standards used to allocate research funding reflect economic and political priorities that favor diseases perceived as life-threatening, measurable or of more relevance to society. Endometriosis, despite affecting millions of women worldwide, remains underfunded and invisible, perpetuating gaps in knowledge and treatment. These standards demonstrate how biomedical ideals of neutrality and objectivity can reproduce gendered epistemic injustice, rendering female experiences of pain systematically illegitimate.

At the same time, we highlight how endometriosis patients' rights groups have played a central epistemic role in disrupting these standards. Through advocacy and networking, activists have challenged assumptions about what diagnostic criteria and funding priorities are most appropriate. Their interventions show that the epistemic renewal of science can emerge from outside its institutional boundaries, precisely because activism operates disruptively in relation to scientific norms. These disruptions can expose the value-laden structure of scientific standards.

The paper contributes to two ongoing debates. First, it extends the values-in-science literature by shifting attention from individual to structural sources of value-ladenness. Second, it introduces justice as a neglected but crucial value in the evaluation of scientific practices. By integrating feminist analyses of ignorance and epistemic injustice with philosophical accounts of scientific standards, we propose a framework for understanding how ignorance has been designed into science and how some forms of activism can be crucial to redesign it. In sum, the case of endometriosis reveals that the pursuit of standards can become a vehicle for injustice. Recognizing this dynamic not only clarifies the epistemic politics of biomedical research but also points toward a more reflexive and justice-oriented conception of scientific responsibility with a strong focus on institutions rather than individuals.

Paper 3: Institutional Neutrality: Problems, Commitments, Concerns (Author 4)

Institutional neutrality is currently a tricky buzzword. It echoes in the corridors of our universities, is evoked in climate change discussions, and plays a big role in legal and journalistic discourses and disputes. The difficulty of dealing with 'institutional neutrality' not only stems from defining what is meant by it – as well as what is kept purposefully ambiguous about it – but also from the impossible task of sketching, in a context-sensitive way, what 'being neutral' might look like and what it implies in practice. Situated in the higher education context, this paper aims to unpack the term 'neutrality' in terms of the

problems it brings with it, the commitments it is bound by, and the concerns it articulates for institutional environments.

The paper consists of four sections. First, it outlines a brief definition of neutrality by drawing on Hugh Lacey's (2013) and Iris Marion Young's (1990) discussions of neutrality and impartiality in philosophy of science and political philosophy, respectively. It compares these definitions to the one outlined in the Kalven Report, a 1967 University of Chicago faculty committee report, which has been quite influential in current discussions of institutional neutrality in the US. It seems that definitions of neutrality presume an existing array of different permissible worldviews or value outlooks. This raises an important question: What should be counted as an admissible worldview or value outlook within an institutional environment? Hugh Lacey, for instance, suggests that alignment with democratic values should be a requirement for an admissible value outlook. How we could implement this rule of admissibility institutionally in the current political context, however, is an open question.

The second section lists different problems concerning neutrality raised in different literatures. Neutrality, for instance, is a fickle term in that, as Sandra Harding argues, it could be used for and against democratic tendencies (2015). In organizational development terms, it is sometimes considered quite abstract compared to its peers such as institutional restraint and dialogue. In addition, 'neutrality' itself is not neutral: it encompasses specific histories and it often has the assumption of a relatively homogeneous society living in harmony built into it (Ludwig 2025, Mills 2007, Young 1990).

The third section examines what the term neutrality is committed to epistemically and politically. I hope to flesh out the distinct relationships neutrality as a term ends up establishing to a. expertise, b. credibility, c. diversity, d. shared reality, and e. inquiry – all important aspects not only of our institutional environments but also of our social landscapes. These distinct relationships, I argue, help us to think about what institutional neutrality as a political and epistemic strategy could amount to.

The paper concludes with a discussion of concerns the term neutrality articulates. In other words, I ask, what institutional neutrality is a safeguard against and whether we need this term for those specific purposes. I also briefly question what is at stake if claims to neutrality and accusations of failing to be neutral are not managed well.

References

Giarola, S., Chiani, L., Drouet, L., Marangoni, G., Nappo, F., Muttarak. R., and Tavoni, M. 2024. Underestimating demographic uncertainties in the IPCC synthesis process. npj Climate Action. 3: 71.

Guivarch, C., Le Gallic, T., Bauer, N., Fragkos, P., Huppmann, D. et al. 2021. Using large ensembles of climate change mitigation scenarios for robust insights. Nature Climate Change. 12: 428-435.

Harding, S. 2015. Objectivity and Diversity: Another Logic of Scientific Research. University of Chicago Press.

Lacey, H. 2013. Rehabilitating neutrality. Philos Stud 163, 77–83.

Ludwig, D and C. H. El-Hani. 2025. Transformative Transdisciplinarity: An Introduction to Community-based Philosophy. Oxford University Press.

Mills, C. W. 2007. White Ignorance. In Race and Epistemologies of Ignorance, ed. by Shannon Sullivan and Nancy Tuana. Albany: SUNY Press, 11-38.

Peters, G., Al Khourdaje, A., Sognnaes, I., Sanderson, B. 2023. AR6 scenarios database: an assessment of current practices and future recommendations. npj Climate Action. 2: 31.

Pirani, A., Fuglestvedt, J., Byers, E., O'Neill, B., Riahi, K., Lee, J., Marotzke, J., Rose, S., Schaeffer, R., Tebaldi, C. 2024. Scenarios in IPCC assessments: lessons from AR6 and opportunities for AR7. npj Climate Action. 3:1.
Tuana, N., 2004. Coming to Understand: Orgasm and the Epistemology of Ignorance. Hypatia 19, 194–232.
Tuana, N., 2006. The Speculum of Ignorance: The Women's Health Movement and Epistemologies of Ignorance. Hypatia 21, 1–19.
Tuana, N., Sullivan, S. (Eds.), 2007. Race and epistemologies of ignorance, SUNY series, philosophy and race. State University of New York Press, Albany.
Young, I. M. 1990. Justice and Politics of Difference. Princeton University Press.

Early Modern Scientific Practice from Margaret Cavendish's Perspective

Kirsten Walsh
University of Exeter
k.walsh2@exeter.ac.uk

Maria Del Mar Yunis
Pontificia Universidad Javeriana
mariadelmaryunis@gmail.com

Peter West
Northeastern University London
peter.west@nulondon.ac.uk

Margaret Cavendish has not traditionally been viewed as one of history's 'winners'. For one thing, there were many idiosyncrasies in her philosophical positions, as well as her methods and output styles. Her criticisms of the emerging experiment-based sciences, for example, seem to put her on the wrong side of history. For another thing, as a woman, she navigated a landscape of intellectual exclusion, engaging with philosophical ideas from the margins rather than within the institutional centres of authority. Consequently, it was all too easy to leave her out of the standard narrative. There has, however, been a recent resurgence of interest in her work. While she may have initially been reconsidered for external reasons of feminism and diversity, she continues to be studied now for internal reasons, such as her innovative metaphysics and astute commentaries on marriage.

In this symposium, we shall explore early modern scientific practice from Cavendish's perspective. Walsh begins by offering a new perspective on Cavendish's objections to experimental philosophy, arguing that her objections offer novel insights into 'agential distance' understood as the relationship between a scientific agent's capacity and their target of inquiry. Yunis' paper focuses on this gap between capacity and target, showing how Cavendish used imagination to fill it, and highlighting the tension between Cavendish's suspicions of experimental methods and her belief in the insufficiency of rationality to guide our understanding. The final paper of the symposium looks at Cavendish from a different perspective. Recognising that early modern science wasn't conducted in a vacuum by isolated actors, West uses digital techniques to study the impacts of intellectual connections and knowledge transfer on scientific practice.

Navigating Agential Distance: Cavendish on the Pitfalls of Experimentation (Walsh)

Margaret Cavendish was critical of the experimental practices of the early Royal Society, likening the early modern experimental philosophers, such as Robert Hooke and Robert Boyle, to "boys that play with watery bubbles or fling dust into each other's eyes, or make a hobbyhorse of snow" (Cavendish 1666/2001, 52), calling them "worthy of reproof rather than praise, for wasting their time with useless sports" (ibid). But her objection was more than simply that experimental philosophy was a childish pursuit. She had serious concerns about the efficacy of experimentation to achieve the very goals she shared with them, namely, to understand the workings of nature so that we might use that knowledge to improve human lives.

In this paper, I use Cavendish's objections to experimental philosophy to explore a notion of 'agential distance' (developed from Murphy, Currie & Walsh, forthcoming), both developing

that concept and showing that it generates a deeper understanding of Cavendish's work. 'Agential distance' refers to the relationship between a scientific agent's capacity and their target of inquiry. In Newton's optical experiments, for example, when he uses a prism to study the refraction of light, we take that distance to be very small, due the immediacy between Newton's interventions and the phenomenon he examined. In contrast, when a team of scientists use data produced by the LHC to study dark matter, distance is much greater: there is significant delay between scientific intervention and phenomenon identification, moreover, the phenomenon operates at a wildly different scale from the agents themselves. Agential distance presents an epistemic challenge; one that might be mitigated with the help of various instruments, interventions and computational techniques. Such techniques might be said to expand 'agential reach' and so mitigate agential distance.

Early modern experimental philosophers recognised that human capacity is limited by their senses. This was potentially a big problem, given that they also believed natural philosophy should be based on matters of fact (information drawn from the senses) rather than speculation. Viewing (a) the senses as assets that could be aided and enhanced and (b) the natural world as something that could be coaxed and manipulated into revealing its underlying processes, experimental philosophers employed all sorts of instruments and interventions to mitigate agential distance. Hooke, for example, used microscopes to drill down into minutiae and study the inner workings of things, and Boyle used the air pump to study the 'spring' of the air. Cavendish agreed that human sensitive capacity is limited. However, for her, using instruments and other techniques would hinder, rather than help. This was because (a') instrumentation would distort rather than enhance and (b') experimental interventions would generate artificial rather than natural phenomena. She agreed that humans could extend their reach. But this could be done by exercising their rational capacities, rather than attempting to enhance their sensitive capacities: "our exterior senses can go no further than the exterior figures of creatures, and their exterior actions: but our reason may pierce deeper, and consider their inherent natures, and interior actions" (Cavendish 1666/2001, 100).

We can use the notion of 'agential distance', and the differing attitudes of Cavendish and Hooke towards distance and its mitigation, to make sense of their differing attitudes towards experimental practices.

Maker's knowledge in the imaginary world: The poetic science of Margaret Cavendish and Giambattista Vico (Yunis)

Both Giambattista Vico and Margaret Cavendish are thinkers who are on the sidelines of the early modern canon, however, they offered a valuable critique of the new natural philosophy by stating a clear opposition to the expulsion of the epistemic value of imagination in science. The purpose of our paper is to defend the crucial role of imagination in early modern science as an epistemically legitimate tool to fill the gaps between the observable and the unobservable in natural philosophy. Based on both Vico's critique of reductive cartesian rationality and Margaret Cavendish's critique of experimental philosophy, we will explore how fantasy becomes a proper (and necessary) tool to understand reality. We argue that Cavendish's defence of the use of imagination to study the natural world can be justified based on Vico's principle of verum factum (maker's knowledge). According to Vico, we can only understand what we can create. For both of these thinkers, imagination is a power of creation whilst rationality is merely an organizing faculty. This means that the imaginary worlds we create, such as Cavendish's Blazing world, are things we make. In this sense, we are closer to understanding an imaginary world made by us than the external natural world of which we are only a part of. Nevertheless, considering that for Cavendish imaginary worlds

are material extensions of the same nature we inhabit daily, fantasy can illuminate reality and become a reasonable artifact to fill the gaps of what exceeds the observable. With this thesis in mind, we will explore three aspects to build our argument: 1) Vico's critique of cartesian rationality by delving into his proposal of an isomorphism between memory, fantasy and imagination, which he used to argue that rationality without imagination becomes withered instead of exalted. Along with the principle of verum factum, which plays a crucial role in the construction of his New Science; 2) Cavendish's claims about the continuity between reason and imagination in order to legitimate the objectivity of fantasy, as well as her arguments on the materiality of imagination, her critique to the new optical instruments, and how her scientific studies were conducted in a single movement with poetic creation; 3) the case of Galileo and other early modern scientists who used imagination and sensible discourses in their science even if not stated explicitly. With these examples we intend to prove Vico's critique regarding an apparent oblivion of the poetic in science, without it actually ceasing to play a role in it, even in the present day. In sum, pure rationality is insufficient to understand the totality of the natural world. The significance of our thesis can be highlighted in different aspects, for instance, by showing how the inclusion of imagination and sensible discourses in science could help avoid overlooking evidence that becomes invisible under a strictly rational scientific method, as imagination enables new pathways for scientific discovery that led natural philosophers like Cavendish to reach conclusions that were ahead of her time.

What if… Cavendish had been a member of the Royal Society? Counterfactual theorising about historical and intellectual networks via computer network simulations (West)

In 1667, Cavendish became the first woman to attend a meeting of the Royal Society. But no women would be admitted to the Society until Kathleen Lonsdale and Marjory Stephenson in 1945. How might things have been different if Cavendish were not just invited to attend a meeting of the Royal Society, but were actually made a Fellow?
Historians tend to take a very cautious approach to counterfactual questions (it is novelists, not historians, who write counterfactual histories). First, because, of course, we can never really know the answers. Second, because a whole host of further facts of the matter would have had to be different (e.g., societal prejudices against women would have to be different or non-existent) for it to be the case that the Royal Society admitted women as fellows in the seventeenth-century. In this paper, however, I will outline an approach to counterfactual theorising about historical questions that, while remaining cautious, can enable us to speculate in an informed way, based on what we know about certain figures (in this case, Cavendish') and their intellectual networks.
I will present the findings of a digital humanities project (led by researchers in London and Boston) that uses computer network simulations methods to examine how the transfer and dissemination of knowledge is impacted by the structure and size of intellectual networks. Put simply, the method being described here involves identifying a network (a group of individuals) and then running simulations on that network where each individual in the network are asked for their credence level (i.e., level of belief) in a given hypothesis – and then shares that belief with the other members of their network. So, for example, a group of individuals might be asked whether they believe a coin is weighted towards heads. Having been asked, they share their view the other members of the network – and are then asked again. Each step in the simulation plays out like this.
Using a dataset obtained from the project Six Degrees of Francis Bacon, which contains information on the connections between 13,309 contemporaries of Francis Bacon (and Bacon himself), we reconstructed three networks – one that really existed, two which did not but might have if Cavendish were made a Fellow of the Royal Society. Network (1) contains the

eight founding members of the Royal Society. Networks (2) and (3) contain the founding members plus Cavendish. The difference between (2) and (3) is that, in (2), Cavendish only retains her actual connections to those other eight individuals whereas, in (3), Cavendish is connected to everyone else in the network, making her the most well-connected member. We called this counterfactual iteration of Cavendish 'super-Cavendish'. We then ran the kinds of simulations described above on these three networks to test how Cavendish's (and super-Cavendish's) exclusion and inclusion from the Royal Society impacted the transfer of knowledge. This findings of these simulations indicate that scientific practice is not a solo endeavour but is deeply embedded – and influenced by – intellectual connections and networks.

After the Practice Turn: What's Next?

Aja Watkins (symposium organizer)

Caleb Hazelwood
University of Wyoming
caleb.hazelwood@uwyo.edu

Morgan Thompson
Northwestern University
morgan.thompson@northwestern.edu

Jacqueline Mae Wallis
University of Pennsylvania
jacqwa@sas.upenn.edu

The so-called "practice turn" in philosophy of science both provided the impetus for the formation of the Society for Philosophy of Science in Practice (SPSP) and was ushered in by its growing membership. By providing a much-needed corrective to a philosophy of science done from the armchair, the mandate to make our work accountable to both historical and contemporary scientific practice has been tremendously generative for our field. Time and time again, evidence from scientific practice has provided support for views like pragmatism, coherentism, perspectivalism, and pluralism, and now a whole generation of fresh scholars have been raised to apply and extend these approaches by engaging with history of science, science and technology studies (STS), and practicing scientists to an unparalleled extent. However, the claim that animates this symposium proposal is that having undergone the practice turn is not enough; substantive, impactful questions remain about both method and theory in a practice-based philosophy of science. The symposium presenters will label and exemplify several open paths forward. First, Caleb Hazelwood and Aja Watkins will distinguish between two questions about the practice turn — those about its purview and those about its fruitfulness — and will set up the content of the subsequent talks by labeling their contributions as (partial) answers to these questions. Next, using a biomedical case study, Morgan Thompson will levy an important critique against those who think that practice turn advocates can remain ambivalent about the value-ladenness of science. And, finally, Jacqueline Wallis argues for the inadequacy of the "adequacy-for-purpose" view of model evaluation as applied to animal studies in preclinical research.
In all, these talks demonstrate that there is controversial, substantive work to be done beyond the practice turn: in general philosophy of science, work on science and values, and philosophy of measurement and modeling.

(1) "The Purview and Fruitfulness of Practice-Oriented Philosophy of Science"
Caleb Hazelwood (presenting author)
Assistant Professor, Department of Philosophy and Religious Studies
University of Wyoming
caleb.hazelwood@uwyo.edu
Aja Watkins
Assistant Professor, Department of Philosophy
University of Wisconsin-Madison

aja.watkins@wisc.edu

There are at least two sets of questions about the "practice turn" in philosophy of science that need to be disentangled.
The first set concerns the purview of the practice turn: what it is and how we should do it. These include questions like: what, exactly, does the practice turn commit us to doing? How should we decide what kinds of engagement with scientific practice are appropriate, and by what standards do we evaluate the relative success of these methods? Is there a single correct way of doing philosophy of science in practice, or should we be pluralists? Is there some feature of our methodologies that would unite the various branches under the practice turn? Why, or for how long, should we take different approaches in tandem? Should professional structures and incentives be revised in order to promote these methodologies and approaches? We contend that, despite widespread and increasing acceptance of practice-oriented methodology, there is actually very little consensus on what it means and what it requires of us.
A second set of questions about the practice turn's fruitfulness can be pursued alongside these. These questions emerge from applying a single instantiation or understanding of the practice turn to new problems. For example, some researchers have taken on board practice-inspired insights such as an "adequacy-for-purpose" view of scientific tools (model, data, measurement, etc.) but have generated important amendments to this view by applying it to, and letting it be further complicated by, scientific practice. Others have productively put practice-centered philosophy of science in contact with work in values in science or feminist philosophy of science, revealing that the practice turn cannot give us all the tools we need to address important questions about science and society, political philosophy of science, and more. In such cases, when an individual applies the practice turn to new problems, this application reflects the individual's understanding of the turn, its principles, and its borders. We argue that there is untapped power in distinguishing these sets of questions. We are convinced of an iterative relationship between them akin to the "looping effect" that Hacking, an early advocate of the practice turn, once identified. That is, when our taxonomies of practice-oriented methodology differ, we consequently differ in how we apply its principles in our analyses of science; this application then contributes to the community's understanding of what the practice turn is and what it can or cannot do. Philosophers with different ideas about the practice turn's purview will arrive at different answers concerning its fruitfulness (about whether, for example, the practice turn is sufficient to solve a problem), and, likewise, their verdicts on fruitfulness will inform commitments regarding purview. We argue that a range of disagreements within practice-oriented philosophy of science right now can be diagnosed by disentangling and tracing the relationship between these two kinds of questions.

(2) "Limitations of Parker's Epistemic Projection Account of Legitimate Values in Science"
Morgan Thompson
Assistant Professor, Department of Philosophy
Northwestern University
morgan.thompson@northwestern.edu

Parker's Epistemic Projection view aims to provide a rapprochement to proponents and critics of the value-free ideal (2024; see also Lusk and Elliott 2022). Epistemic projection involves explicitly reasoning from one's nonepistemic aims/values to "a set of epistemic goals, preferences, and constraints" that guide research choices during the project (Parker 2024, 19). The basic insight is that nonepistemic aims/values are "screened off from the research stage" by referencing only those epistemic goals and desiderata in making and

justifying decisions, including fitness-for-purpose judgments (Parker 2024, 27). The value-free ideal (VFI) is maintained for decision-making during "internal" scientific activity, but nonepistemic values can guide the selection of epistemic goals and desiderata.
However, I argue that the Epistemic Projection view is inadequate for rapprochement because it raises new problems in social communities with a diversity of nonepistemic values. These problems are: (1) the problem of value capture, and (2) it is too demanding in requiring multiple lines of reasoning to the same epistemic goals and desiderata. These problems are better resolved by critics of the VFI.
First, there are potential harms in implementing epistemic projection at the level of the research community. When a research problem is specified and sufficiently fit-for-purpose, this may lead others in the research community to adopt the research problem without having seriously considered their own nonepistemic values. In some cases, values will be "captured" by the structure of data or other aspects of the epistemic projection (Nguyen 2024). Some members may not have deliberated or substantially reflected on their social/political values. Seeing a fit-for-purpose epistemic research problem may lead their values to be captured in the sense that they may consider only the set(s) of values that allow one to reason to those epistemic goal(s) and desiderata.
Second, Parker's Epistemic Projection view attempts to separate the debates about what values are pre-supposed in research from questions of justification. This move results in Parker's Epistemic Projection account being too demanding. It requires, at least, two complete lines of reasoning (one appealing to certain nonepistemic values/aims and the other not) to support the same epistemic research problem. Parker overestimates the frequency of convergent lines of reasoning from different nonepistemic aims/values.
As a case study for my argument, I analyze ChoGlueck's (2022) work on decisions about safety and effectiveness of male contraception. Depending on whether researchers view preventing pregnancy as an individual or shared responsibility, their epistemic desiderata will differ and so, lead to different decisions about when to end a clinical trial.

Works Cited

ChoGlueck, C. (2022). Still no pill for men? Double standards & demarcating values in biomedical research. Studies in History and Philosophy of Science, 91, 66-76.
Lusk, G., & Elliott, K. C. (2022). Non-epistemic values and scientific assessment: an adequacy-for-purpose view. European Journal for Philosophy of Science, 12(2), 35.
Nguyen, C. T. (2024). Value capture. J. Ethics & Soc. Phil., 27, 469.
Parker, W. S. (2024). The epistemic projection approach to values in science. Philosophy of Science, 91(1), 18-36.

(4) "Making Flies, Organoids, Simulations, and Mice Play Nice: Evaluating experimental ensembles in translational biomedical research"
Jacqueline Mae Wallis
PhD candidate, Department of Philosophy
University of Pennsylvania
jacqwa@sas.upenn.edu

In this talk I argue that an "adequacy-for-purpose" view of model evaluation is inadequate for biomedical research when ensembles of experimental model systems are used in preclinical studies aimed at translating results to humans.
Roughly 90% of the drugs and medical interventions which succeed in animal studies eventually fail to be safe and effective in humans. Some call this the "translatability crisis."

To help address the crisis, the US National Institutes of Health recently announced major changes to their funding structures: they will no longer develop funding calls for animal-only research (NIH 2025a,b). Instead, the focus will be projects which use human-focused methods (sometimes called new approach methods [NAMs]) either solely or in combination with animal models. Human-focused methods include using human cell lines, organoids, or organs-on-a-chip, and AI/ML simulations, among other NAMs.

In the new funding landscape, preclinical researchers will need normative guidance for evaluating potential experiments, especially when they are combining older, more familiar animal models with NAMs. A popular view advocates researchers select the model system which is adequate for (their) purpose (Parker 2020). This adequacy-for-purpose view has gotten support from philosophers of science and seems to match practice. In a working group on enhancing rigor, transparency, and translatability in animal research, for instance, the NIH reported on the need to "help investigators to select the most appropriate model for a given question" (Wold & Tabak 2021, emphasis original). I will show, however, that in translational research, approaching model evaluation according to adequacy considerations can be a mistake.

Such an approach commits what I here call modeling atomism, where experimental model systems are evaluated narrowly, one-by-one, and according to a circumscribed purpose. To establish this, I will first show how current accounts in the philosophy of science fail to capture the nature of modeling systems in translational research. I propose a new concept to help diagnose problems with modeling atomism: experimental system ensemble (ESE). An ESE is the set of experimental systems used for research on a particular target. In translational research, the target is typically a human disease or other condition: Research on Alzheimer's disease, for example, employs a suite of animal models (like flies and mice), various in vitro systems such as human induced pluripotent stem cells and organoids, and in silico techniques (Trujillo-Estrada et al. 2022). I address some worries and challenges about the ESE concept before using it to show why adequacy-for-purpose is inadequate for providing scientists with the needed normative guidance. Because preclinical research is characterized by the ensemble strategy, model evaluation cannot proceed atomistically, model+purpose by model+purpose. Instead, such research must be effective qua its ensemble nature; studies in a set of animals only partially achieve this by making use of whole organisms. Thus the failing record of unreliable animal research may have been due to problems with ensembles, rather than problems with species differences. This result means both that the NIH has de-prioritized animals for the wrong reason and that hopes for NAMs may go unrealized.

Works Cited:

National Institutes of Health. 2025a. "NIH Funding Announcements to Align with NIH Initiative to Prioritize Human-Based Research." National Institutes of Health News Releases, July 10. https://grants.nih.gov/news-events/nih-extramural-nexus-news/2025/07/nih-funding-announcements-to-align-with-nih-initiative-to-prioritize-human-based-research

National Institutes of Health. 2025b. "NIH to Prioritize Human-Based Research Technologies." National Institutes of Health News Releases, April 29. https://www.nih.gov/news-events/news-releases/nih-prioritize-human-based-research-technologies

Parker, Wendy S. 2020. "Model Evaluation: An Adequacy-for-Purpose View." Philosophy of Science 87 (3): 457–77. https://doi.org/10.1086/708691.

Trujillo-Estrada, Laura, Elisabeth Sanchez-Mejias, Raquel Sanchez-Varo, et al. 2022. "Animal and Cellular Models of Alzheimer's Disease: Progress, Promise, and Future

Approaches." The Neuroscientist 28 (6): 572–93.
https://doi.org/10.1177/10738584211001753.
Wold, Barbara, and Lawrence A. Tabak. 2021. ACD Working Group on Enhancing Rigor, Transparency, and Translatability in Animal Research. Final report. NIH.
https://acd.od.nih.gov/documents/presentations/06112021_ACD_WorkingGroup_Presentation.pdf

Pragmatic understanding in scientific practice

Oscar Westerblad
University of Iceland
owesterblad@hi.is

Mike Stuart
University of York
mike.stuart.post@gmail.com

Sebastian de Haro with Eline De Jong
University of Amsterdam
e.l.dejong@uva.nl

André Curtis-Trudel
University of Cincinnati
curtisa4@ucmail.uc.edu

Siyu Yao
University of Cincinnati

20 years ago Henk de Regt and Dennis Dieks (2005) published their seminal paper developing a contextual account of scientific understanding. The account characterises understanding as having an ineliminably pragmatic element, namely, understanding necessarily involves the context and skills or abilities of the understander. On their development of this idea, pragmatic understanding, or understanding how to do something, involves the skill and judgement necessary to produce explanations of phenomena, in other words, to produce explanatory understanding. Since this paper was published, much ink has been spilt on the nature of scientific understanding. While much philosophical work has been done on explanatory (e.g., Strevens 2013; De Regt 2017; Khalifa 2017; Potochnik 2017) and objectual (e.g., Elgin 2017; Wilkenfeld 2017; Baumberger 2019; Dellsén 2020) forms of understanding, comparatively little work has been done on the nature of pragmatic understanding.

Some philosophers of science have taken seriously the skill required for pragmatic understanding and how it might go beyond its role in the use of theory, practices of modelling, and explanation (Lenhard 2006; Chang 2009; Leonelli 2009; Bengson 2020; Delarivière & Van Kerkhove 2021; Stuart 2025; Westerblad and de Regt 2025). What they show is that not all scientific achievements are explanatory in nature, nor about the application of theory. However, much work remains to be done in exploring pragmatic understanding's relationship to explanation, the extent to which it can account for various non-theoretical aspects of scientific practice, and its relationship to other key notions in philosophy of science.

Instead of defending an abstract philosophical account of pragmatic understanding, each speaker in this symposium begins by drawing attention to particular scientific practices that do not straightforwardly fit with dominant thinking about understanding as necessarily explanatory or theoretical, drawing on detailed cases studies and interviews with scientists. Each speaker develops these themes by attending to the details of cases of scientific and technological practices where theory and explanation are not at the fore. Instead, what these cases reveal, is that scientists rely on a host of epistemic skills and abilities to achieve a variety of practical and epistemic aims: from the building and use of robots (talk 2) and

technologies (talk 3) to the development of methods (talk 1) and the evaluation of pursuitworthiness (talk 4).
By focusing on these cases, we bring to the fore the more general need to recognise pragmatic reasoning, pragmatic understanding, and cognitive skill as essential promoters and constituents of epistemic and practical achievements in the sciences. By focusing on overlaps between science, technology, and engineering, we also demonstrate how these achievements arise in cases that are not so often discussed in the literatures on scientific understanding.
What unities these talks and the work they build on, is the recognition that pragmatic understanding involves aim-oriented epistemic skills, like de Regt and Dieks argued, and that such skills can constitute epistemic achievements on their own, even if they are not used for the development of explanation, contra de Regt and Dieks.

Representative bibliography

Baumberger, Christoph 2019. Explicating Objectual Understanding: Taking Degrees Seriously, Journal of General Philosophy of Science, 50: 367-388.
Bengson, John 2021. Practical Understanding: Skill as Grasp of Method. In Concepts in Thought, Action, and Emotion: New Essays, edited by Christoph Demmerling and Dirk Schröder, pp. 215-235. New York and London: Routledge.
Chang, Hasok 2009. Ontological Principles and the Intelligibility of Epistemic Activities. In Scientific Understanding: Philosophical Perspectives, edited by Henk de Regt, Sabina Leonelli, and Kai Eigner, pp. 64-82. Pittsburgh: University of Pittsburgh Press.
Currie, Adrian 2020. Bottled understanding: The role of lab work in ecology. The British Journal for the Philosophy of Science, 71(3), 905-932.
Delarivière, Sven and Bart Van Kerkhove 2021. The mark of understanding: In defense of an ability account. Axiomathes, 31(5), 619-648.
Dellsén, Finnur 2020. Beyond Explanation: Understanding as Dependency Modelling. The British Journal for the Philosophy of Science: 71 (4): 1261-1286.
De Regt, Henk W. 2017. Understanding Scientific Understanding. Oxford: Oxford University Press.
De Regt, Henk W. and Dennis Dieks 2005. A Contextual Approach to Scientific Understanding. Synthese, 144 (1): 137-170.
Elgin, Catherine Z. 2017. True Enough. Cambridge, MA: MIT Press.
Khalifa, Kareem 2017. Understanding, Explanation, and Scientific Knowledge. Cambridge: Cambridge University Press.
Lenhard, Johannes 2006. Surprised by a Nanowire: Simulation, Control, and Understanding. Philosophy of Science, 73 (5): 605-616.
Leonelli, Sabina 2009. Understanding in biology: The impure nature of biological knowledge. In Scientific Understanding: Philosophical Perspectives, edited by Henk de Regt, Sabina Leonelli, and Kai Eigner, pp. 189-209. Pittsburgh: University of Pittsburgh Press.
Potochnik, Angela 2017. Idealization and the Aims of Science. Chicago: The University of Chicago Press.
Strevens, Michael 2013. No understanding without explanation. Studies in History and Philosophy of Science Part A, 44 (3): 510-515.
Stuart, Michael T. 2025. A New Account of Pragmatic Understanding, Applied to the Case of AI-Assisted Science. Philosophical Studies. https://doi.org/10.1007/s11098-025-02336-6
Westerblad, Oscar and Henk W. de Regt 2025. Scientific Understanding Beyond Representing: Lessons From Ian Hacking's Work. The Monist, 108 (4): 353-371.
Wilkenfeld, Daniel 2017. MUDdy Understanding. Synthese, 194: 1273–1293.

Talk 1: Engineering methods for better understanding (Oscar Westerblad)

In his attempts to re-engineer philosophy for limited beings, William Wimsatt (2007) develops what he terms an "engineering perspective". This perspective encourages philosophers to look at what scientists do, especially in cases of concrete application and problem-solving. While many accounts of scientific understanding focus on the explanatory and theoretical upshots of scientific inquiry, relatively little has been said about if--and, if so, how--understanding might arise in more practical problem-oriented and design-oriented domains, like in engineering. In this talk, I develop the epistemology of the "engineering perspective" by suggesting that methods supply the means appropriate to specific ends in our pragmatic reasoning, which produces pragmatic understanding consisting in the general ability to achieve intended aims by solving problems. Methods--procedures, techniques, and heuristics--emerge organically as the solutions to specific problems, only later becoming systematised, if at all. Better and more systematic methods allow us to solve more problems by identifying previously disparate problems as similar in kind (Kuhn 1977), thus increasing our pragmatic understanding. These solutions also provide further standards for evaluating problem-solutions (Dewey 1938; Chalmers 2013, ch. 11), thus undergirding the kinds of judgement identified as necessary for pragmatic understanding, according to de Regt (2009).
To show how methods originate and evolve, I draw on the rich historical cases of aeronautical engineering in Walter Vincenti's (1990) work. I use his case of the development of control-volume analysis (Vincenti 1990, ch. 4) to demonstrate how methods evolved and produced understanding. The control-volume analysis started out as a series of disparate and ad hoc solutions to particular problems. Only later in its development did it become a unified and systematic method for identifying and solving problems in fluid mechanics. While engineers had some understanding of how to solve problems of fluid flow, I suggest that their understanding was improved once the control volume analysis allowed them to identify different problems as of the same type, by allowing them to employ the same method. Importantly, this method works around the kind of physical explanations necessary for explanatory understanding of those phenomena; yet, we should say that the engineers and scientists working with the control volume analysis exhibited pragmatic understanding. Pragmatic understanding involves the ability to identify and solve problems in a systematic fashion, which depends on the development and employment of methods.
While it is easy to dismiss these engineering sciences as of only practical worth that do not produce deep or genuine scientific understanding, this would miss out on the potential benefits of reflecting more carefully on the epistemology of the engineering sciences. In particular, many contemporary sciences demonstrate engineering methodologies and the "engineering perspective"; for example neuroscience (Chirimuuta 2024), biology (Leonelli 2009; Lazebnik 2002), and work in artificial intelligence (Stuart 2025). At the end of the talk, I therefore conclude with some more general lessons for the epistemology of science, especially on the importance of epistemic agency and means-ends reasoning (Chang 2022; Woodward 2023).

Talk 2: Pragmatic Understanding and Scientific Robotics (Michael T. Stuart)

There are many areas in science that could not exist without robotics. For those interested in the epistemology of science, there are some especially interesting cases in which the robotic elements are conceived, designed, and implemented with scientific aims in mind from the start. For example, much research in biomedicine, chemistry, pharmacy and space science revolves around certain robotic instruments that impose constraints on what may be considered as data and what is possible to do. In addition, the robotic components of the

research often produce new data and play a key role in the justification of scientific conclusions (e.g., Ross King's robotic scientists Adam, Eve, and Genesis). In these cases, design decisions are also epistemological decisions.
This talk attempts to get clearer on the epistemic aims of scientific robotics in cases like these, by looking closely at the history and practice of robotics. I claim that one main aim of scientific robotics is to extend human abilities, and this is an epistemic aim when those abilities are epistemic, that is, when they extend abilities that allow us to achieve epistemic aims like learning more about a system. I argue that the best way to understand this epistemic aim is to characterize it as a kind of understanding, specifically, pragmatic understanding.
The resources of pragmatic understanding can help us make sense of the claim that robotics is making (more or less) linear progress without paradigm shifts, as well as to explain which advances in robotics researchers single out as the most important in the history of their field, and how and when roboticists praise each other in terms of being imaginative, intuitive, and creative.
It might be intuitive, even obvious, to characterize robotics in this way. But doing so can still be helpful at least because it makes it easy to state (and perhaps also address) a worry that many have had about this and related disciplines: While it might be right to characterize the primary aim of robotics as one of increasing pragmatic understanding, that might be precisely what is wrong with fields that prize ability first and truth second. Namely, such fields end up (explicitly or implicitly) licensing corner-cutting, encouraging means-end thinking without responsibility, and enabling a generally toxic atmosphere.
In response, I sketch a critical feminist, decolonial, anti-oppressive reply, building on the radical empiricist feminism first developed by Elizabeth Anderson and Sharyn Clough. On their view, the aim of pragmatic understanding can be shown to itself suggest an answer. Specifically, pragmatic understanding should be applied not just to developments in the robots themselves, but to the entire field, including its methods, values, and social structures. Thus, the values championed by critical approaches must be adopted insofar as these are likely to increase skills and abilities across the discipline. Increasing (epistemic and ethical) responsibility requires increasing skills which themselves allow for greater robotically-extended scientific abilities. This will be illustrated with a case study from recent space science.

Talk 3: A Contextual Approach to Technological Understanding (Sebastian de Haro and Eline De Jong)

What does it mean to understand technology? While understanding has been extensively analysed in philosophy of science, it remains comparatively underexplored in relation to technology. Drawing on the concept of scientific understanding (De Regt, 2017), De Jong and De Haro (forthcoming) conceive technological understanding as the cognitive skill involved in engaging—practically or conceptually—with technologies. More specifically, it is the ability to realise an aim by using a technological artefact. The notion of technological understanding thus foregrounds the epistemic dimension of engaging with technologies: every successful use of a technology involves some form of understanding. Yet what counts as 'use', and which specific cognitive skill it requires, depends crucially on context.
This paper develops this insight into a systematic, contextual account of technological understanding. I argue that technological understanding is not a singular concept, but a family of context-relative cognitive skills that admit of degrees. A taxi driver may 'understand' their car well enough to operate it, yet this understanding may prove inadequate when the car breaks down. Similarly, a shipyard owner may have a deep understanding of their vessels' design and capabilities, but lack the understanding required to sail them. Both the taxi driver

and the shipyard owner thus possess the ability to ‘use’ their technologies, yet in importantly different ways. Treating their understanding as points on a single scale obscures the qualitative differences between their forms of expertise and their relevance in different contexts.
To capture this, I distinguish three typical contexts of technological engagement: design, where artefacts are constructed; operation, where they are practically used; and innovation, where new applications are devised. In each context, technological understanding takes the form of a specific cognitive skill: the ability to design an artefact, to operate it, or to envisage and evaluate its possible uses. Correspondingly, artefacts must be intelligible in different ways: in terms of their inner workings (design), their practical performance (operation), and their functional possibilities and limitations (innovation).
I further argue that these context-relative forms of understanding can be analysed in terms of counterfactual reasoning. To have technological understanding, on this view, is to be able to answer relevant ‘what-if’ questions: what if the artefact’s structure were changed (design), what if conditions of use were to vary (operation), and what if the artefact were deployed in new circumstances (innovation). This counterfactual account highlights the pragmatic character of technological understanding and provides a systematic basis for assessing this skill across different contexts of engagement.
Finally, I show how this contextual framework not only does justice to the different ways in which we engage with technological artefacts and the distinct types of expertise that can be attained, but also helps to better target efforts to improve technological understanding. For instance, when anticipating the societal impact of emerging technologies, a ‘functional’ type of understanding oriented towards potential uses may be more pertinent than a ‘technical’ understanding focused on inner workings. Yet many current public engagement initiatives still prioritise the latter, potentially misaligning the kind of understanding they cultivate with the aims they pursue.

Talk 4: Deep Learning in Deep Space: How AI guides pursuit in frontier space science (André Curtis-Trudel and Siyu Yao)

Recent achievements by deep learning (DL) systems have reinvigorated interest in automated scientific discovery (Choudhary et al. 2023). The prospect is that DL methods might facilitate the discovery of phenomena, hypotheses, or models more efficiently than traditional approaches—or even uncover results that would otherwise remain epistemically inaccessible. Drawing on a series of scientist interviews conducted between Spring 2023 – Fall 2025 in Belgium, the UK, and the USA, this paper examines how DL is actually being deployed in frontier space science, focusing on gravitational-wave astronomy (GWA) and cosmological inference as representative case studies. We argue that while DL rarely delivers fully autonomous discovery in these domains, it is poised to play a crucial epistemic role in justifying pursuit of promising hypotheses.
In GWA, noise-dominated detector outputs make it difficult to infer properties of GW-emitting events. In cosmological inference, researchers must constrain cosmological parameters by identifying sensitive sources and extracting information inaccessible to traditional statistical methods. Researchers in both domains have increasingly turned to DL to address these challenges. In GWA, DL is now being used to circumvent otherwise costly Bayesian parameter-estimation pipelines, while in cosmology, DL enables simulation-based inference (SBI), which tackles situations where explicit likelihood functions are unavailable due to stochastic or nonlinear processes.
Across both domains, we argue that DL's primary epistemic contribution is not to automate discovery per se, but rather to guide scientists towards promising avenues of inquiry.

Drawing on recent economic models of pursuitworthiness (e.g., Duerr & Fischer 2024; Curtis-Trudel et al. ms), we argue that DL guides pursuit by freeing up resources, reducing adjustment costs, and increasing researcher bandwidth. With DL, scientists can more quickly flag anomalous signals, reveal exotic sources, and highlight parameters whose investigation promises epistemic gain.

This analysis has two broader implications. First, it refocuses attention on how DL can facilitate pragmatic understanding in space science (cf. Kieval & Westerblad, ms). This is because the amount of pragmatic understanding had by a scientist plausibly impacts which ideas are pursuitworthy for them. This focus contrasts with analyses that foreground the role of explanatory understanding (e.g., Meskhidze 2021). Second, it highlights the pursuitworthiness of DL itself as a scientific tool. A tool can be pursuitworthy not only when it directly advances specific research projects, but also when its exploratory use provides information for evaluating and comparing multiple potential projects. This second aspect justifies the general development and application of DL even when it does not promise immediate novel discoveries—a practical consideration particularly relevant for resource-intensive fields like astrophysics.

Evidential Pluralism in Practice

Michael Wilde
Kent and Medway Medical School
michael.wilde@kmms.ac.uk

Alexandra Trofimov
University of Manchester
alexandra.trofimov@manchester.ac.uk

Jon Williamson
University of Manchester
jon.williamson@manchester.ac.uk

Evidential Pluralism is a theory about the epistemology of causal claims. It maintains that, in order to establish a causal claim, it is typically necessary to establish two other claims: (i) an "association claim", that the cause is appropriately associated with the effect; (ii) a "mechanism claim", that the cause is linked to the effect by an appropriate mechanism (Williamson, 2019). Evidential Pluralism maintains also that, as a result, different types of study are typically required to establish a causal claim, namely, association studies and mechanistic studies (Shan and Williamson, 2023). The idea is that the different studies are complementary: the strengths of one type of study may address the weaknesses of another type of study, and vice versa (Illari, 2011).

Evidential Pluralism has come in for some criticism. For example, Mattia Andreoletti and David Teira (2019) object that Evidential Pluralism has not been shown to be better in practice to alternative approaches to causal inquiry; instead, Evidential Pluralism has been supported only by 'Platonic' philosophical arguments. Similarly, Bennett Holman (2019) argues that Evidential Pluralism depends upon a 'friction-free epistemology', abstracting away real practical obstacles to causal knowledge, for example, the effects of industry. In practice, the epistemology of causation is full of friction.

In this symposium, we apply Evidential Pluralism to three distinct areas of practice: (1) Medical Education; (2) Evidence-Based Law; (3) Evidence Review. We hope to address the question: Is Evidential Pluralism helpful in practice?

Evidential Pluralism and evidence review

What is the best way to evaluate whether an intervention works? Orthodox 'evidence-based' evaluation methods review certain kinds of studies (especially randomised trials) but exclude other relevant evidence (such as mechanistic studies and views of stakeholders) by claiming that such evidence is of inherently low quality.

Evidential Pluralism maintains that mechanistic studies are not of inherently low quality and that they should be scrutinised alongside the association studies that are the focus of orthodox methods. The same applies to stakeholder testimony, at least insofar as this testimony constitutes evidence of mechanisms and their contextual variation.

The question then arises as to how practical methods for evidence review should be modified in order to properly accommodate mechanistic evidence. This paper puts forward a general framework for mechanism-informed evidence review (MIER). This is a flexible approach to evidence review that can be tailored to a variety of kinds of research questions (including questions about average effect, variation in effect, extrapolation, particularisation to a subpopulation, explanation and optimal policy); a variety of types of review (including

systematic review, qualitative evidence synthesis, mixed methods review, narrative review, meta-narrative review, realist review, integrative review and umbrella review); and a variety of modes of review (standard review, rapid review, living review).
While MIER fits with a variety of review questions, types and modes, it does require changes to the review structure, because it requires reviewing mechanistic studies alongside the association studies that are the focus of orthodox review. The paper provides an account of which mechanistic considerations are relevant to each kind of review question, and thus an account of how the structure of evidence review needs to change in each case.
We then turn to a further structural change. A systematic review will often incorporate a meta-analysis, which generates a confidence interval estimate of effect size from all studies included in a review. This is often criticised as relatively uninformative due to the heterogeneity of the included studies. Mechanistic evidence can tell us which studies should be included in a meta-analysis in order to more reliably estimate an effect size.
Mechanistic evidence can also inform the ways in which individual studies are assessed. The paper discusses three such considerations. (i) It is important to assess whether the study design is coherent in the light of what is understood about mechanisms. (ii) Mechanistic evidence can also help us to predict direction of bias and compensate for bias, rather than simply discount studies that may be susceptible to bias. (iii) Mechanistic evidence can help in the assessment of the relevance (aka 'directness') of a study: a study is more relevant if the study population is mechanistically similar to the population of interest to the review question.
We then show how MIER can be applied in practice with three very different interventions as case studies: face-mask mandates during COVID-19, educating jurors about rape myths, and an antibody-drug conjugate cancer medicine.

Evidential Pluralism and Medical Education

Medical students in the UK are often taught the methods of Evidence-Based Medicine (EBM). EBM emphasises the importance of identifying and evaluating association studies, such as randomised controlled trials (RCTs), for medical decision-making; this emphasis is revealed in its hierarchy of evidence (cf. Howick, 2011). For example, to decide whether patients with arthritis should be given a specific drug, benoxaprofen, relevant RCTs should be identified and evaluated.
Evidential Pluralists instead emphasise the importance of identifying and evaluating association studies alongside mechanistic studies. For example, to decide whether patients with arthritis should be given benoxaprofen, relevant RCTs should be identified and evaluated alongside relevant mechanistic studies. For instance, studies into the mechanisms of benoxaprofen and arthritis may have suggested that the results of the RCTs were not representative of the elderly population (Worrall, 2011). Evidential Pluralists call this the EBM+ approach (Wilde, 2023).
Trisha Greenhalgh and colleagues have recently argued in favour of an EBM+ approach (Greenhalgh et al., 2022). Her famous textbook on EBM now includes a chapter on the EBM+ approach (Greenhalgh and Dijkstra, 2024). However, the EBM+ approach is still not explicitly taught in medical schools. One reason is that the approach was developed in another discipline, namely, in the philosophy of science. As a result, medical educators may be unfamiliar with the approach and its benefits.
In this paper, I argue the EBM+ approach is beneficial to medical educators not just because it is the proper way to make medical decisions, but also because it provides a framework for teaching students the basics of medical methodology more effectively. In particular, there is an overwhelming variety of medical methods: quantitative and qualitative studies,

experimental and observational studies, studies involving animal and computational models, and so on. Compared to an approach focused on association studies, the EBM+ approach provides a more comprehensive and systematic framework for thinking about the roles of the different studies, and their different strengths and weaknesses. In particular, the strengths and weaknesses of each study type can be understood in terms of how well that study type contributes to establishing an association claim or a mechanism claim.

Evidence-Based Law

The aim of this talk is to demonstrate the benefits of applying Evidential Pluralism to evidence-based law.

Evidence-based law (EBL) is an emerging approach to law that seeks to make use of the best available evidence to evaluate and justify laws and inform legal decision-making. EBL is increasingly advocated as a best practice to improve lawmaking and achieve 'better' law (OECD, 2021; UK Government, 2023; EU Commission, 2021 and 2023). The implementation of EBL, however, remains limited.

A key obstacle to the implementation of EBL is methodological. The standard evidence-based approach of evidence-based medicine (EBM) and evidence-based policy (EBP) involves evaluation methods and evidence hierarchies that place high value on certain kinds of evidence, especially RCTs (Luján, 2023).

Limitations of the standard evidence-based approach have been highlighted in both medicine (Parkkinen et al., 2018; Greenhalgh et al., 2022) and policy (Cartwright and Hardie, 2012; Lujan, 2023; Shan and Williamson, 2023). When we turn to law, however, the limitations are even more significant (Davis, 2019; Santos and Magnani, 2024; Rantala et al., 2023). The main reason to think that EBL needs to look beyond the standard approach of EBM and EBP is that it can be practically impossible to produce RCTs of high enough quality to establish the relevant causal claim.

In this talk, I introduce a new approach to evidence-based law, called EBL+, based on the principles of Evidential Pluralism (Trofimov and Williamson, 2025).

The EBL+ methodology and its benefits will be illustrated through an application to the issue of online fake news. Numerous jurisdictions around the world have proposed or implemented legal interventions to mitigate the negative impact of online fake news. Until recently, however, little evidence has been provided to support concerns that online fake news has a detrimental effect. This is problematic for the justification of legal interventions. Legal interventions must be necessary and proportionate, which requires demonstrating that online fake news does have a negative impact.

The lack of evidence to support concerns that online fake news has a negative impact is due, at least in part, to difficulties involved in isolating and assessing the impact of online fake news. Robust RCTs could provide good evidence. However, given concerns over the detrimental effects of fake news, purposefully exposing a large, representative cohort to fake news in the real-world social media environment would be unethical. RCTs conducted in controlled study environments can overcome ethical concerns but they fail to capture crucial aspects of the real-world environment, reducing their external validity. Observational studies can provide evidence that captures the real-world context. However, given the complexity of human behaviour, it can be difficult to control all variables in observational studies and therefore difficult to establish causation.

These difficulties, however, can be overcome. I will show that, by combining evidence of mechanisms alongside evidence of correlation, EBL+ enables robust overall evidence of the effects of online fake news to be provided. This, in turn, enables an evidence-based justification for legal interventions aimed at tackling online fake news.

References

Andreoletti M and Teira D. Rules versus standards: What are the costs of epistemic norms in drug regulation? Science, Technology, and Human Values, 44(6): 1093-1115. 2019.
Cartwright N and Hardie J. Evidence-Based Policy: A Practical Guide, Oxford University Press. 2012.
Davis KE. 'The Limits of Evidence-Based Regulation: The Case of Anti-Bribery Law', SSRN. 2019.
EU Commission. Better Regulation Guidelines. Available at: Better regulation: guidelines and toolbox. 2021.
EU Commission. Better Regulation Toolbox. Available at: Better regulation: guidelines and toolbox. 2023.
Greenhalgh T and Dijkstra P. How to Read a Paper. Wiley Blackwell, 7th edition. 2024.
Greenhalgh T, Fisman D, Cane DJ, et al. Adapt or die: how the pandemic made the shift from EBM to EBM+ more urgent. BMJ Evidence-Based Medicine, 27:253-260. 2022.
Holman B. Philosophers on drugs. Synthese, 196(11): 4363-4390. 2019.
Howick J. The Philosophy of Evidence-Based Medicine. BMJ Books. 2011.
Illari P. Mechanistic evidence: Disambiguating the Russo-Williamson thesis. International Studies in the Philosophy of Science, 225(2): 139-157. 2011.
Lujan LJ. Evidence-Based Policies: Lessons from Regulatory Science, Politics and Policy. 2023.
Organization for Economic Co-operation and Development (OECD). Regulatory Policy Outlook. Available at: OECD Regulatory Policy Outlook 2021 | OECD. 2021.
Parkkinen V-P, Wallmann C, Wilde M, Clarke B, Illari P, Kelly MP, Norell, C, Russo, F, Shaw, B, and Williamson, J. Evaluating Evidence of Mechanisms in Medicine: Principles and Procedures, Springer, Cham, Switzerland. 2018.
Rachlinski JJ. Evidence-Based Law, Cornell Law Review, 96(4): 901-924. 2011.
Rantala K, Alasuutari N, and Kuorikoski, J. The Logic of Regulatory Impact Assessment: From Evidence to Evidential Reasoning, Regulation and Governance. 2023.
Russo F. and Williamson J. Interpreting Causality in the Health Sciences, International Studies in the Philosophy of Science, 21(2): 157-170. 2007.
Santos D.H.M. and Magnani A.B. Evidence-Based Law: A New Approach to Legal Practice Under the Scope of the Pragmatic Methodologies of Evidence-Based Medicine, Beijing Law Review, 15. 2024.
Shan Y and Williamson J. Evidential Pluralism in the Social Sciences. Routledge. 2023.
Trofimov A. and Williamson J. Applying Evidential Pluralism to Evidence-Based Law: EBL+. Jurisprudence. 2025.
UK Government. Better Regulation Framework. Available at: Better Regulation Framework - GOV.UK. 2023.
Wilde M. The EBM+ Movement. The International Journal of Biostatistics, 19(2): 283-293. 2023.
Williamson, J. Establishing causal claims in medicine. International Studies in the Philosophy of Science, 32(1): 33-61. 2019.
Worrall J. Evidence in Medicine and Evidence-Based Medicine. Philosophy Compass, 2: 981-1022. 2011.

Race, Genetic Diversity, & Health

Lauren Wilson
University of Exeter
l.wilson4@exeter.ac.uk

Abigial Nieves Delgado
Utrecht University
a.nievesdelgado@uu.nl

Carolina Mayes
University of Edinburgh
carolina.mayes@ed.ac.uk

Matthew Maxwell
University of Wisconsin-Madison
mjmaxwell3@wisc.edu

Scientists recognize that human genomics has a diversity problem (Sirugo et al. 2019). These research projects promise to revolutionize medicine with more precise diagnoses and treatments of health conditions. Yet, most genomic studies are conducted on populations of European ancestry. This is not just an issue of representation; the findings of some genomic studies (like results from GWAS) are known to be less accurate the more the study population differs from the target population. In response, scientists are working to make health genomics more inclusive of human diversity. However, this focus on human diversity runs the risk of perpetuating harm. For example, projects which prioritize diversity often operate in a top-down way; potentially promoting extractive practices and research that is divorced from the needs of the often-marginalized communities it is meant to benefit (Lee et al., 2022; Shim et al., 2022). There is also the risk of incorrectly attributing poor health outcomes to marginalized racial or ethnic groups (via attributions of genetic ancestry) based on inheritance patterns for genomic regions implicated in poor health outcomes. This could in turn lead to blaming individuals (or racial or ethnic groups) for the health inequality they experience (Duster 2003; Erdmann et al. 2021).

This symposium weaves together approaches from philosophy of science, sociology and ethnography to investigates the question: what epistemic and ethical issues arise in the genomic study of human health once scientists recognize the need for more inclusive research? And how can understanding scientific practices improve our ability to mitigate the risk of racism and social harm in human health genomics?

The panel comprises senior scholar Abigial Nieves Delgado, historian and philosopher of science. Her project analyzes the epistemic role the logic of underrepresentation plays in genomics and microbiome research in the Global South by looking closely to two important projects: Latinbiota and AWI-Gen Phase 2. In doing so, she focuses on how real inclusion in health requires examining the relationship between Global North and Global South. In turn, Carolina Mayes uses sociological and ethnographic methods to demonstrate how frontline workers navigate the process of community brokerage, balancing dissatisfaction with diversity and inclusion practices against anxiety that nonparticipation will jeopardize their community's capacity to benefit from genomic findings. She argues that the over-reliance on quantified representations of inclusion undermines the ethics of research projects even as it also compels participation. Lauren Wilson demonstrates how identifying points where concerns of diversity shape methodological choices in GWAS of health can illuminate new

areas for mitigating potential harms. Matthew Maxwell interrogates the idea that there exists a correct subdivision of the human population in genomic studies of human health. In doing so he shows that different population subdivisions are useful for different purposes and that finding one privileged subdivision does not aid in avoiding spurious correlations. This prospective symposium showcases just a few of the many diverse and impactful way's philosophy of science, social sciences, and bioethics can work together to uncover the ways race and genetic diversity inform scientific practice in genomics.

1. Global health and the logic of underrepresentation: "Filling the gap" with Global South samples

Abigail Nieves Delgado

Population descriptors used in biomedicine like race, ethnicity, nation, culture, religion should represent a "concept of difference" and reveal meaningful biological and health differences. These descriptors are useful to diversify datasets and potentially make biomedical applications more inclusive. Calls for inclusion usually highlight the need to "fill the gap" in databases – they defend, what I call, a logic of underrepresentation. This logic, historically and today, promotes the development of specific research programs and the funding of international benchmark projects. Prominent examples are the Human Genome Diversity Project, the HapMap and All of Us projects. While diversity and inclusion in databases is a worthy goal to pursue and it is justifiable to address underrepresentation, there are epistemic and ethical problems that are included when following this goal. Epistemically, more data doesn't necessarily always leads to better explanations and more data isn't always translatable into health interventions. Moreover, several authors have pointed to the ethical problems many diversity-focused project have, like racialization, racism, commercialism, exploitation, and cultural imperialism. At the core of these projects lies an understanding of human diversity in terms of race and a colonial representation of this difference. In their efforts to sample all populations around the world, these projects impose a racializing way of understanding human diversity assuming that racial categories are generalizable categories. This is problematic given the historical and geographical contingency of racial labels. In addition, these studies create Indigeneity following a colonial logic of representation and deny the possibility of self-recognition.
In this paper I rely on an integrated approach to history and philosophy of science and values to (1) uncover the origins of the logic of underrepresentation in the historical development of global health studies. I discuss how this framework was implemented to pursue health internationally and how an historically grounded idea of "global health" has led to the powerful logic of underrepresentation and data inclusion in biomedicine. This historical analysis contributes to (2) understanding the development of local sampling initiatives (of meta-genomic samples) with the purpose of inclusion and, explains why this logic continues to fuel the creation of research projects around the world. As a case study, (3) I analyze the epistemic role the logic of underrepresentation plays in genomics and microbiome research in the Global South by looking closely to two important projects: Latinbiota and AWI-Gen Phase 2. These two initiatives, funded by the National Institute of Health and the Wellcome Trust, have the objective to collect microbiome samples to solve the problem of underrepresentation of Latin America and South Africa. With this analysis I aim to show how global health ideas shape local research. That is, how global and local power relationships counteract local research independence and the possibility of local teams to create health solutions for their communities. Ultimately, I argue that it is fundamental to consider the

North-South economic and technological dependencies shaping international science if we want to achieve real inclusion in global health.

2. Diversity and disempowerment: Community brokers, epistemic injustice, and the compulsion of genomic representation
Carolina Mayes

Much has been written concerning the diversity problem in human genomics research and its uncomfortable intersection with both the perpetuation of health inequities and the reification of biological race realism. This dualism reflects the entrenched salience of racial and ethnic categories in biomedical and academic research (Adetiba & Reich, 2025; Epstein, 2007), a pattern that has made diversity and inclusion ubiquitous but insufficient concepts for remedying persistent forms of inequality and disadvantage (Ahmed, 2007; Ferryman, 2023). As a recent body of scholarship describes, diversity initiatives in genomics are commonly top-down interventions, preoccupied with meeting recruitment targets and quantifying proportions of "underrepresented" populations included in their data (Lee et al., 2022; Shim et al., 2022). This instrumentalist approach to diversity and inclusion is additionally reflected in the operationalization of participatory research practices like community engagement, described by one study as a "pantomime" performed in the service of data collection (Shim et al., 2023).
The shortcomings of genomic diversity initiatives place particular strain on their frontline staff, who are often tasked with leveraging their membership with a minority community to facilitate recruitment and retention (Liu & Meyers, 2020; Vasquez et al., 2023). This position charges staff with multiple and conflicting forms of responsibility, demanding fulfillment of their obligations to the research project while also urging attention to the needs and circumstances of the community. Drawing on the sociological concept of brokerage (Stovel & Shaw, 2012), in this paper I situate how genomic diversity initiatives place frontline research staff in a uniquely divided role, in which the promise of genomic inclusion becomes both a community benefit and a threat. Using interviews conducted with staff at a high-profile US-based precision medicine research project, I demonstrate how frontline workers navigate the process of community brokerage, balancing dissatisfaction with diversity and inclusion practices against anxiety that nonparticipation will jeopardize their community's capacity to benefit from genomic findings. This analysis draws attention to the consequences of impoverished conceptualizations of diversity in human genomics, pointing to how the over-reliance on quantified representations of inclusion undermines the ethics of research projects even as it also compels participation. Finally, in light of the policy changes of the new US administration, this paper calls for further study of the aftermath of diversity initiatives, and of the particular need to consider the impact of research programs on community partners, who trade their social networks in furtherance of biomedical goals.

3. How do considerations of race inform scientific practice in genomic research on human health?
Lauren Wilson

This talk investigates the question "how do concerns about racial and genetic diversity structure scientific practice in genomic studies of human health?". Over the last decade there has been an alarm raised, most genomic studies are performed on white populations of European ancestry. To prevent disparity in access to the potential benefits of genomic studies of health, there have been continued calls to diversify the populations studied, with emphasis placed on increasing racial diversity, a proxy for genetic ancestry (Wojcik et al. 2019).

However, little attention has been paid to how concerns about racial and genetic diversity structure scientific practice outside of the formation of the sample population. Failing to recognize the diversity of way race and genetic diversity can shape research can allow serious harms to persist, such as interpreting health disparities between racial groups to genetic difference rather than environmental differences. Continuing to be blind to the way race and genetic diversity shape other aspects of the research process restricts our ability to mitigate potential harms.

To move beyond this focus on sample populations, I analyze the ways race and genetic diversity shape scientific practices in genomic studies of health. I identify choice points in the workflow of scientists in which concerns about race and genetic diversity shape their scientific investigation. I show how sources of funding shape the way research formulate their study design. I also explore how the intended audience of the findings, where it is basic research or results that will be used by clinicians in specific community health systems shape the refinement of results. In doing so, I demonstrate that scientists are not only concerned with racial and genetic diversity of their starting sample population but along the whole trajectory of scientific research. Put simply, how scientists think about race and genetic diversity can inform their choices from the initial formulation and funding of a study to how results are used in clinical environments. Recognizing the various points in a research project where concerns about race shape scientists' choices can help address questions about these studies potential to perpetuate or mitigate racial inequalities in health even if the diversity of study populations increased.

However, given the great heterogeneity of genomic research projects on human health, my analysis offers a kind of model of what an investigation like this could look like. I focus my analysis on how concerns about race and genetic diversity shape research choices of studies utilizing GWAS methods (including those working on multi-ancestry populations) originating in the United States and the United Kingdon. I focus on these studies because they are widely used to study health disparities. I conclude illustrating how several other areas such as research on rare diseases and rare variants could benefit from similar analysis.

4. Population Subdivison and the Search for Disease-Linked Loci: Must K Be Arbitrary?
Matthew Maxwell

Discovering the number of population subdivisions (K) is an important research goal in the genetic analysis of human populations—one held to have implications for health and medical research. Genetic case-control studies—an important tool for finding find disease-linked loci—require controlling for population structure in order to avoid spurious correlations (Pritchard and Donnelly 2001; Satten, Flanders, and Yang 2001; Evanno, Regnaut, and Goudet 2005). For example, one early study (Knowler et al. 1988) found an association between some Gm haplotypes and non-insulin-dependent diabetes in Pima and Papago Native American populations. This association disappeared, however, when admixture with Caucasian populations was controlled for. Failure to control for population structure, therefore, risks harms to populations with diverse genetic backgrounds.

Studies of human population structure, however, have been strongly criticised. One criticism is that K is either chosen arbitrarily or else based on unjustified and biased assumptions about the population structure (Weiss and Fullerton 2005). Indeed, it may seem from what is said in some widely-cited population human genetics papers, such as Rosenberg et al. (2002), that the algorithms used cannot support such an inference and that finding the correct number of subdivisions is not a goal. Many have pointed out that these algorithms are "exploratory" (as opposed to confirmatory) and that K is an input to the algorithm, not an output (Griesemer and Barragán 2022).

It is true that the authors of widely-cited population genetics papers often do little to justify the inclusion of various values for K. Rosenberg et al. (2002), for example, simply evaluate the highest probability subdivision for the global population given each K from two to six. They make no claim that the much criticized K = 5 set of continental populations is the correct, or even the best-supported, population subdivision. In supplementary materials, Rosenberg et al. (2002) suggest that analysis of their data at K > 6 was not carried out because multiple clustering solutions appeared at that level. In spite of this, citing Pritchard, Stephens, and Donnelly (2000), the authors discuss that finding the K that maximizes the posterior probability of the data could be used to infer the number of subdivisions in the population, however the inference seems to have been considered unreliable and so not reported. Other estimators are suggested by Pritchard, Stephens, and Donnelly (2000) but are either unstable 1 or have a high-dimensional parameter space, making computation infeasible.

In this talk I examine two questions. First, could the suggested methods in fact support a value for K? Multiple clustering solutions with a high posteriors, do not obviously disqualify a value for K. Moreover, larger values for K may provide a better fit with the data, but this risks over-fitting the data (Evanno, Regnaut, and Goudet 2005). Second, what assumptions does this research question depend on? Most obviously, finding the correct K assumes that there exists a correct subdivision of the human population—an assumption that critics of race research challenge (Weiss and Fullerton 2005; Winsberg 2022). It also requires the assumption that finding these subdivisions will aid research, but it is plausible that different population subdivisions are useful for different purposes and that finding one privileged subdivision does not aid in avoiding spurious correlations.

Symposium References:

Adetiba, E., & Reich, A. (2025). The Science of Blackness: Biology and Social Science in Black-Focused Scholarship, 1900–2020. Sociology of Race and Ethnicity, 00(0), 1–21. https://doi.org/10.1177/23326492251384914

Ahmed, S. (2007). The language of diversity. Ethnic and Racial Studies, 30(2), 235–256. https://doi.org/10.1080/01419870601143927

Duster, Troy (2003). Backdoor to Eugenics. Taylor & Francis. 2nd ed., Routledge.

Epstein, S. (2007). Inclusion: The Politics of Difference in Medical Research. University of Chicago Press.

Erdmann, A., Rehmann-Sutter, C., & Bozzaro, C. (2021). Patients' and professionals' views related to ethical issues in precision medicine: a mixed research synthesis. BMC Medical Ethics, 22(1), Article 116.

Evanno, Guillaume, Sebastien Regnaut, and Jérôme Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." Molecular Ecology 14 (8): 2611–20.

Ferryman, K. (2023). Bounded Justice, Inclusion, and the Hyper/Invisibility of Race in Precision Medicine. The American Journal of Bioethics, 23(7), 27–33. https://doi.org/10.1080/15265161.2023.2207515

Fullwiley, D. (2008). The Biologistical Construction of Race: `Admixture' Technology and the New Genetic Medicine: `Admixture' Technology and the New Genetic Medicine. Social Studies of Science, 38(5), 695-735.

Griesemer, James, and Carlos Andrés Barragán. 2022. "Re-Situations of Scientific Knowledge: A Case Study of a Skirmish over Clusters Vs Clines in Human Population Genomics." History and Philosophy of the Life Sciences 44 (2): 16.

Knowler, William C, RC Williams, DJ Pettitt, and A Steinberg. 1988. “Gm3; 5, 13, 14 and Type 2 Diabetes Mellitus: An Association in American Indians with Genetic Admixture.” American Journal of Human Genetics 43 (4): 520.
Lee, S. S.-J., Fullerton, S. M., McMahon, C. E., Bentz, M., Jeske, M., Vasquez, E., Foti, N., Saco, L., & Shim, J. K. (2022). Targeting Representation: Interpreting Calls for Diversity in Precision Medicine Research. Yale Journal of Biology and Medicine, 95(3), 317–326.
Liu, C., & Meyers, K. (2020). Beyond clinical trials: Social outcomes of structured stakeholder engagement in biomedical HIV prevention trials in China. Culture, Health and Sexuality, 22(12), 1365–1381. https://doi.org/10.1080/13691058.2019.1683230
Pritchard, Jonathan K, and Peter Donnelly. 2001. “Case–Control Studies of Association in Structured or Admixed Populations.” Theoretical Population Biology 60 (3): 227–37.
Pritchard, Jonathan K, Matthew Stephens, and Peter Donnelly. 2000. “Inference of Population Structure Using Multilocus Genotype Data.” Genetics 155 (2): 945–59.
Rosenberg, Noah A, Jonathan K Pritchard, James L Weber, Howard M Cann, Kenneth K Kidd, Lev A Zhivotovsky, and Marcus W Feldman. 2002. “Genetic Structure of Human Populations.” Science 298 (5602): 2381–85.
Satten, Glen A, W Dana Flanders, and Quanhe Yang. 2001. “Accounting for Unmeasured Population Substructure in Case-Control Studies of Genetic Association Using a Novel Latent-Class Model.” The American Journal of Human Genetics 68 (2): 466–77.
Shim, J. K., Bentz, M., Vasquez, E., Jeske, M., Saperstein, A., Fullerton, S. M., Foti, N., McMahon, C., & Lee, S. S.-J. (2022). Strategies of inclusion: The tradeoffs of pursuing “baked in” diversity through place-based recruitment. Social Science & Medicine, 306, 115132. https://doi.org/10.1016/j.socscimed.2022.115132
Shim, J. K., Foti, N., Vasquez, E., Fullerton, S. M., Bentz, M., Jeske, M., & Lee, S. S.-J. (2023). Community Engagement in Precision Medicine Research: Organizational Practices and Their Impacts for Equity. AJOB Empirical Bioethics, 14(4), 185–196. https://doi.org/10.1080/23294515.2023.2201478
Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. Cell. 2019 Mar 21;177(1):26-31. doi: 10.1016/j.cell.2019.02.048. Erratum in: Cell. 2019 May 2;177(4):1080.
Stovel, K., & Shaw, L. (2012). Brokerage. Annual Review of Sociology, 38(1), 139–158. https://doi.org/10.1146/annurev-soc-081309-150054
Vasquez, E. E., Foti, N., McMahon, C. E., Jeske, M., Bentz, M., Fullerton, S., Shim, J. K., & Lee, S. S.-J. (2023). Rethinking Benefit and Responsibility in the Context of Diversity: Perspectives from the Front Lines of Precision Medicine Research. Public Health Genomics, 103–112. https://doi.org/10.1159/000531656
Weiss, Kenneth M, and Stephanie M Fullerton. 2005. “Racing Around, Getting Nowhere.” Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews 14 (5): 165–69.
Winsberg, Eric. 2022. “Putting Races on the Ontological Map: A Close Look at Spencer’s ‘New Biologism’ of Race.” Biology & Philosophy 37 (6): 46.