BSPS 2019 Annual Conference

17-19 July, Durham University

Wednesday 17th July

14.00 – 15.00	Registration desk opens – *Chemistry Atrium*
15.00 – 16.30	Parallel Sessions 1
16.30 – 17.00	Coffee Break
17.00 – 18.30	Plenary 1: Jenann Ismael (Columbia) – “Rethinking Time and Determinism”
18.30 – 19.30	OUP Drinks reception – *Calman Learning Centre top floor*

Thursday 18th July

09.30 – 11.00	Parallel Sessions 2
11.00 – 11.30	Coffee Break
11.30 – 13.00	Plenary 2: Robin Hendry (Durham) and Paul Needham (Stockholm) – “Chemical Substances”
13.00 – 14.00	Lunch Break / Poster Session
13.25 – 13.55	BJPS Meet the Editors – *Chemistry 218*
14.00 – 15.30	Parallel Sessions 3
15.30 – 16.00	Coffee Break
16.00 – 18.00	Parallel Sessions 4
19.30 – 22.00	Conference dinner – *Zizzi’s Restaurant, 43-44 Saddler St, Durham DH1 3NU*

Friday 19th July

09.30 – 11.00	Parallel Sessions 5
11.00 – 11.30	Coffee Break
11.30 – 13.30	Parallel Sessions 6
13.30 – 14.30	Lunch Break
13.40 – 14.25	BSPS Annual General Meeting – *Calman Learning Centre 202*
14.30 – 16.00	Plenary 3: Nicholas Shea (London) – “Syntactic and Semantic Inferences in the Representational Theory of Mind”

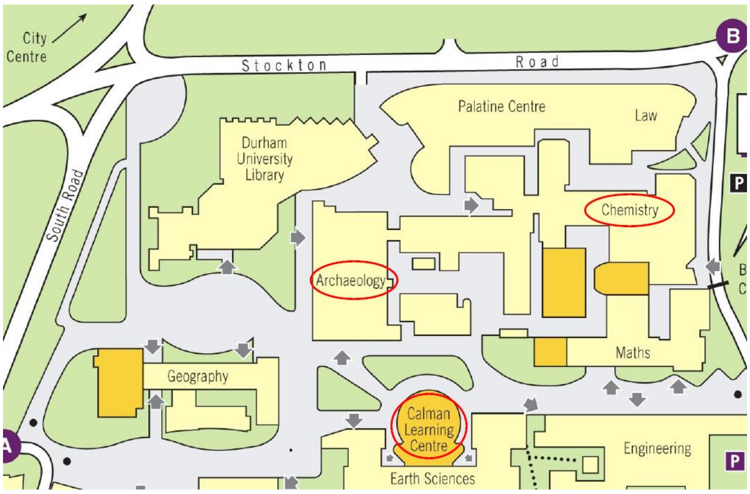
All plenary sessions will take place in the Rosemary Cramp lecture theatre in the Calman Learning Centre, Room 202.

For **parallel sessions**, see schedule below.

Coffee breaks and lunches will be held in the Chemistry Atrium and Café.

The **poster session** will be located in the Chemistry Atrium; a list of posters is below.

The **map** on the right shows conference buildings; arrows indicate entrances.



Parallel Sessions

Parallel Sessions 1: Wednesday 15.00 – 16.30

Symposium Calman 202	**Symposium** Chemistry 060	**Symposium** Chemistry 218	**Symposium** Archaeology 210
Biases in the Sciences and Science-Based Policy *Lorenzo Casini* *Bennett Holman* *Saana Jukola* *Juergen Landes*	Cosmology in Silico *Helen Meskhidze* *Marie Gueguen*	Multiple Realisability in the Sciences *Alexander Franklin* *Tuomas E. Tahko* *Marion Godman*	Moving Past the Naturalism-Normativism Dichotomy in Philosophy of Medicine *Frances Fairbairn* *Brandon Conley* *Shane Glackin*

Parallel Sessions 2: Thursday 09.30 – 11.00

Symposium Calman 202	**Symposium** Chemistry 060	**General** Chair: *Sophie Ritson* Chemistry 218	**Life** Chair: *Elselijn Kingma* Archaeology 210
The New Reduction: Formal, Conceptual, and Physical Perspectives *Neil Dewar* *Samuel C. Fletcher* *Laurenz Hudetz* *Katie Robertson*	The impact of the replication crisis on philosophy: two case studies *Suilin Lavelle* *Richard Morey* *Hugh Rabagliati*	*Enno Fischer* Causation, Intervention and Responsibility	*Jonathan Grose* Disease, Sex, Senescence and Pregnancy. Who's normal?
		Ulrich Stegmann Quantifying causal specificity comes up short	*Sam Fellowes* Why symptom-based approaches are not enough: the value of psychiatric diagnoses
		Margherita Harris Confidence: A New Dimension of Scientific Knowledge?	*Michael Wilde* Evidence in cancer epidemiology at IARC

Parallel Sessions 3: Thursday 14.00 – 15.30

General Chair: *William Peden* Calman 202	**Physical** Chair: *Alastair Wilson* Chemistry 060	**Life** Chair: *Riana Betzler* Chemistry 218	**Mixed** Chair: *Katie Robertson* Archaeology 210
Harry Lewendon-Evans Socially Extended Scientific Understanding	*Henrique Gomes* Gauge and boundary: a complicated relationship	*Joe Dewhurst* Perspectival realism about mechanistic functions	*Chloé de Canson* Salience and the Sure-Thing Principle
Katherine Furman Going it Alone (Epistemically)	*Caspar Jacobs* Inequivalent Representations & the Coalesced Structures Approach: Non-Radically Unpristine	*Geoff Keeling/Niall Paterson* Proper Functions: Etiology Without Typehood	*Jean Baccelli/Rush Stewart* Support for Geometric Pooling
Franklin Jacoby What are we pluralist about?	*John Dougherty* Fields, loops, and the Strong CP problem	*Martin Zach* Revisiting abstraction and idealization in molecular biology	*David Glass/Jonah Schupbach* Conjunctive Explanations

Parallel Sessions 4: Thursday 16.00 – 18.00

General Chair: *Katherine Furman* Calman 202	**Physical** Chair: *Lina Jansson* Chemistry 060	**Life** Chair: *Suilin Lavelle* Chemistry 218	**Mixed** Chair: *Beth Hannon* Archaeology 210
Robert Northcott Prediction markets and extrapolation	*Peter Evans/Sam Baron* What's So Spatial About Time Anyway?	*Catherine Greene* Predictable behaviour and intentional action: Disentangling the two	*Ana-Maria Crețu* Natural Kinds as Real Patterns
Donal Khosrowi What is (successful) extrapolation?	*Lucy James* Time, Cauchy Problems and Physical Modality	*Domi Dessaix* What does it take to be a psychological primitive? Separating innateness from foundationalism	*Riana Betzler* Stability and the Looping Effects of Human Kinds
Olav Vassend Justifying the Norms of Inductive Inference	*Matt Farr* Indeterminism and the C theory	*Manolo Martinez* Direct Perception and Computation	*Tiziano Ferrando* The ontology of patterns
William Peden Direct Inference in the Material Theory of Induction	*Baptiste Le Bihan* Spacetime Emergence and Functional Realization	*Mihnea Capraru* Drawing the semantics–pragmatics distinction in animal communication	*[talk cancelled]*

Parallel Sessions 5: Friday 09.30 – 11.00

Symposium Calman 202	**Symposium** Chemistry 060	**General** Chair: *H.Lewendon-Evans* Chemistry 218	**Modelling** Chair: *Wendy Parker* Archaeology 210
Effective Field Theories: Top-down and bottom-up *Richard Dawid* *Michael Stöltzner* *Porter Williams* *Martin King*	Structure and Composition in Chemistry *Karoliina Pulkkinen (by pre-recorded video)* *Vanessa A. Seifert* *Geoffrey Blumenthal (deceased; paper read by V.Seifert)*	*Arthur Harris* Quinean Realism and a New Defence of Antirealism	*Karen Crowther/Niels Linnemann/Christian Wüthrich* What we cannot learn from analogue experiments
		Rune Nyrup Explanatory Pragmatism as a Philosophy for the Science of Explainable Artificial Intelligence	*Alexander Gebharter/ Christian J. Feldbacher-Escamilla* Confirmation Based on Analogical Inference: Bayes meets Jeffrey
		Chris Dorst Predictive Infelicities and the Neo-Humean Conception of Laws	*Philippos Papayannopoulos* Computing and Modelling: Analog vs. Analogue

Parallel Sessions 6: Friday 11.30 – 13.30

Physical Chair: *Karen Crowther* Calman 202	**Mixed** Chair: *Ana-Maria Crețu* Chemistry 060	**General** Chair: *Peter Vickers* Chemistry 218	**Modelling** Chair: *Juha Saatsi* Archaeology 210
Javier Anta Is physically significant the analogy between Shannon's information and mechanical statistical entropies?	*Alice Murphy* The Literary Form of Scientific Thought Experiments	*Joseph D. Martin/Agnes Bolinska* Negotiating History: Contingency, Canonicity, and Case Studies	*Atoosa Kasirzadeh* Levels and a new role for mathematics in empirical sciences
James Wills Gibbs' Solution of Gibbs' paradox	*Sophie Ritson* Probing Novelty at the Large Hadron Collider: Heuristic appraisal of disruptive experimentation	*Luca Tambolo* Multiple discoveries, multiple errors, and the inevitability of science	*Michele Lubrano* Difference-making and explanation in mathematics
Niels Linneman Quantisation as a method of discovery	*Roberto Fumagalli* On the Individuation of Choice Options		*Hannah Tomczyk* Descriptions don't always close the gap in the mapping account
Ronnie Hermens Sufficiently Real? A Critical Review of the Theorems by Colbeck and Renner			*Colin McCullough-Benner* Heaviside's Operational Calculus and the Application of Unrigorous Mathematics

Posters

Thursday 13.00 – 14.00 in the Chemistry Atrium

Antonis Antoniou
A pragmatic approach on the ontology of models

Niels Linnemann and Kian Salimkhani
The constructivist's programme and the problem of pregeometry

BSPS 2019 Plenary Talks

Rethinking Time and Determinism

Prof Jenann Ismael (Columbia University)

I discuss the openness of the future in a relativistic setting in which there are deterministic laws. I argue against many kinds of common wisdom.

Chemical Substances

Defending Microstructural Essentialism

Prof Robin Hendry (Durham University)

Microstructural essentialism in chemistry is the thesis that a particular chemical substance is the one that it is (and not another) in virtue of its microstructure. In this paper I articulate this thesis by saying what a chemical substance is, and what a microstructure is. (Each task is surprisingly non trivial.) I also defend microstructural essentialism.

Water and Macroscopic Concepts

Prof Paul Needham (Stockholm University)

Water's compositional formula characterises the substance without involving any assumptions about microstructure. The microstructure of water raises the question of why it should be regarded as a single substance. I suggest that the phase rule provides a suitable criterion. This shows how macroscopic concepts bear on the characterisation of substances.

Syntactic and Semantic Inferences in the Representational Theory of Mind

Prof Nicholas Shea (Institute of Philosophy, University of London)

A long-standing datum in cognitive science is that people make semantic inferences, which draw on the meaning of concepts, as well as purely syntactic inferences, which don't. That contrast is puzzling since the representational theory of mind (RTM) assumes that all inferences are a matter of causal transitions between representational vehicles in virtue of non-semantic properties. Semantic inferences used to be picked out as those that draw on the internal structure of a concept. However, experimental work on concepts has produced a near-consensus that, for a typical lexical concept, there is no single representational structure which is always involved in thinking with that concept. At the same time, the recent conspicuous success of deep convolutional neural networks in modelling various categorisation tasks suggests that much of the information we draw on when using a concept is not conceptually or explicitly represented at all, but is instead implicit in dispositions to apply the concept on the basis of non-conceptual representations. This new landscape has many attractions, however the old contrast between syntactic and semantic inferences seems to have been squeezed out. Can we still explain the contrast, within the strictures of RTM? This paper argues that we can, not by appealing to conceptual structure, but by making a novel distinction between two types of representational processing in which concepts are involved.

BSPS 2019 Symposia

Biases in the Sciences and Science-Based Policy

Lorenzo Casini (Université de Genève), Bennett Holman (Yonsei University), Saana Jukola (Bielefeld University), Jürgen Landes (Munich Center for Mathematical Philosophy)

Summary

Incentive structures in and outside the sciences which are not conducive to truth-directed research have resulted in biased research and researchers. While there has been much recent work on biases in (the philosophy of) medicine, biases in nutrition science have received scant attention and are addressed by the first two presentations. Jukola discusses the different standards for evidence in nutrition research and medicine and concludes that the recommendation to include sustainability as a goal of the Dietary Guidelines for Americans can be taken to be science-based and not representing political bias. Next, Holman explores the role of the sugar industry in dietary guidelines and extends the discussion of the role of values to the level of the social structure of inquiry and shows that incentive structures can bias an epistemic community without corrupting (illegitimately influencing) any of the individual community members.

The remaining presentations focus on the consequences of biased research in general. Landes presents a seemingly paradoxical fact in social epistemology regarding the assessed bias of groups of scientists. He explains the “paradox" and argues that science is in need of good PR. Casini discusses three strategies for teasing out confirmation from meta-analyses of biased randomised controlled trials, and argues that, in practice, a combination of these strategies must be used.

Meta-analyses and Conflicts of Interest

Lorenzo Casini & Jan Sprenger

In medical research, meta-analyses over multiple randomized controlled trials (RCTs) are praised for mitigating the problem of confounding due to the small sample size of individual RCTs (Worrall 2002). Meta-analyses promise to eliminate confounding by pooling together multiple RCTs designed to test the efficacy of a given intervention, and by recalculating the effect size over a larger sample. In spite of this apparent advantage, some authors have observed that meta-analyses, too, have limitations (Stegenga 2011, Ioannidis 2016). This paper focusses on an important but under-appreciated limitation: many of the RCTs pooled together by meta-analyses (according to a recent study by Roseman et al. [2011], almost 70%) are subject to conflicts of interest. The problem is not explicitly addressed by current protocols on how to perform a meta-analysis (see, e.g., Higgins & Green [2011], x10.4), with the result that current reviews tend to omit any reference to conflicts of interest, let alone solve the problems due to them. Intuitively, the two considerations-increasing the sample size and avoiding biases due to conflicts of interest-pull in different directions, namely including vs excluding biased RCTs from meta-analyses. Unfortunately, current protocols for producing meta-analyses do not specify how to deal with conflicts of interests. In this paper, we argue that-contrary to what one might at first think-evidence from biased studies may be useful to more accurately calculate effect sizes, conditional on using an appropriate discounting procedure.

To explain, we assume that the distribution of effect sizes across the various RCTs is multimodal: it is the result of unbiased and biased data generating processes, where the biased processes correspond to the presence of conflict of interest. We do not know the extent of the bias, but we know its direction: RCTs subject to conflicts of interest overestimate, if at all, the effect size of the intervention. If possible, one would like to exploit this information by appropriately discounting, rather than ignoring, the biased estimates. How?

Assume for simplicity that there are just two processes at work, an unbiased one and a biased one. Our proposal is that the pooling procedure should proceed in two steps. First, one should pool separately RCTs without a conflict of interest and RCTs with a conflict of interest. Next, 6one should then compute the overall effect size by using both average effect sizes, respectively, each being weighed by its aggregate sample size, and the aggregate biased sample size being weighed in addition by a discount factor. The crucial question is, of course, about the magnitude of the discount factor, and our paper explores several discounting procedures that could outperform the simple pooling function of unbiased RCTs. Three strategies come to mind: the first based on the usefulness of biased information; the second based on the reliability of potentially biased information; and the third based on the empirical record of previous meta-analyses where a sufficient amount of unbiased studies exists. We argue that, in practice, a combination of these three strategies must be used, dependent on the particular nature of the meta-analysis, and the research question of the underlying RCTs.

References

Higgins, J. and S. Green (Eds.) (2011). Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0 ed.). The Cochrane Collaboration, 2011.

Ioannidis, J. P. (2016). The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. The Milbank Quarterly 94(3), 485-514.

Roseman, M., K. Milette, L. Bero, J. Coyne, J. Lexchin, E. Turner, and B.

Thombs (2011). Reporting of conflicts of interest in meta-analyses of trials of pharmacological treatments. JAMA 305(10), 1008-17.

Stegenga, J. (2011). Is meta-analysis the platinum standard of evidence?

Stud Hist Philos Biol Biomed Sci 42(4), 497-507.

Worrall, J. (2002). What evidence in evidence-based medicine? Philosophy of Science 69, 316-30

Bias without Corruption: An analysis of the influence of Big Sugar on dietary research

Bennett Holman & William Berger, Aaron Bramson, Patrick Grim and Daniel J. Singer

Sugar is quickly being added to the pantheon of industrial sectors which are preceded by the epithet “big". What unites such diverse fields (e.g., tobacco, oil, pharma) and prompts public derision, is the formation of syndicates by erstwhile competitors to promote their sector-specific interests at the expense of the public good. The publication of the “cigarette papers" served as a landmark event that showed that these efforts went beyond political lobbying and reached deep into shaping the scientific record upon which political discourse was based (Glantz et al. 1996). Since then careful scholarship has shown that such efforts have long been a part of other commercial sectors such as pharmaceuticals (Krimsky 2003), energy (Oreskes & Conway 2010), and chemicals (Douglas 2009; Elliott 2011).

Recently, a number of authors have claimed that research funding from Big Sugar led to a distortion of nutritional guidelines for over a generation (e.g. Taubes 2017). According to this account, by funding health researchers at Harvard's nutrition science department, the sugar industry was able to persuade the medical community that high-fat diets were at fault for chronic illnesses and thus deflect attention away from the ills of sugar. However, in a recent article in Science, historians of science have argued that the charge that there was untoward behavior stems from an idealized notion of science (Johns & Oppenheimer, 2018). They claim that researchers were independent, unbiased and simply following the evidence, which at the time seemed to indicate that fat was primarily to blame.

By separating the structure of inquiry from the character of individual researchers, we will show that these view are compatible. In section 2, we examine this historical case in detail and assess the extent to which the research funded by the sugar industry did in fact bias scientific research. In section 3 we step from the particulars of this case and explore a formalized model of scientific research. Here we show that it is possible for rational agents engaged in inquiry to become systematically biased without being corrupted. In section 4, we compare these results to previous work in the philosophy of industry-funded science (Holman & Bruner, 2017) and argue that while these results pick out a new phenomenon, they are a species of the "industrial selection effect."

References

Douglas, H. (2009). Science, Policy, and the Value-Free Ideal. Pittsburgh: University of Pittsburgh Press.

Elliott, K. (2011). Direct and Indirect Roles for Values in Science. Philosophy of Science, 78, 303-324.

Glantz, S., Slade, J., Bero, L., Hanaur, P., & Barnes, D. (1996). The Cigarette Papers. Berkeley: University of California Press.

Johns, D. M., & Oppenheimer, G. M. (2018). Was there ever really a "sugar conspiracy"? Science, 359(6377), 747-750.

Holman, B. & Bruner, J. (2017). Experimentation by Industrial Selection. Philosophy of Science, 84, 1008-1019.

Krimsky, S. (2003). Science in the private interest: Has the lure of profits corrupted medical research? Landham, MD: Rowman and Littlefield

McGandy RB, Hegsted DM, Stare FJ. (1967). Dietary fats, carbohydrates and atherosclerotic vascular disease. New England Journal of Medicine, 277(5):245247.

Oreskes, N. & Conway, E. (2010). Merchants of Doubt. New York: Bloomsbury Press.

Taubes, G. (2017). The case against sugar. Anchor Books.

Political Bias in Nutrition Guidelines - The Case of Sustainability, Standards of Evidence, and Concepts of Health

Saana Jukola

In evidence-based nutrition policymaking, science converges with economic and political concerns and cultural habits. Debates on nutrition policy often focus on the recommendations that official, government-issued nutrition guidelines give. These recommendations are not immutable or universal, and the fact that guidelines have different emphases in different countries is sometimes interpreted as a sign of the lack of objectivity. Consequently, discussions surrounding healthy eating and nutrition policy often become heated, and accusations of politicization and bias are not rare (e.g., Freidberg 2016; Rucker & Rucker 2016).

This paper focuses on one such dispute, namely the debate that took place in the United States in 2015 when the American Dietary Guidance Advisory Committee (DGAC) suggested that the Dietary Guidelines for Americans should promote economically and environmentally sustainable diets, which, in practice, means promoting plant-based diets. The critics (e.g., Kuttner 2014) denounced the recommendation as unscientific and politically motivated. According to them, the DGAC transgressed its statutory bounds by introducing sustainability as an integral part of dietary guidelines.

The aim of the paper is to make explicit and critically evaluate some of the background assumptions that underlie the argumentation of the proponents and the critics of including sustainability concerns into governmental nutrition guidelines. I will focus on two, partly intertwined, issues: the concepts of 'health' that are implicitly invoked in the discussion and the ideals concerning standards of evidence that are referred to while questioning sustainability recommendations. The US dietary guidelines state that "its recommendations are ultimately intended to help individuals improve and maintain overall health" (US Department of Health and Human Services 2015). According to the critics, the agenda of including sustainability does not relate to this goal and, thus, undermines the justification of the guidelines. Moreover, the critics claim, the evidence from computer models the DGAC referred to while recommending plant-based diets was weak: unlike clinical randomized trials and quasi-experiments, computer models cannot deliver evidence on causal relationships. Consequently, it was argued that including sustainability concerns into the Dietary Guidelines for Americans would undermine their scientific basis (Kuttner 2014).

By referring to recent philosophy of science literature (e.g., Valles 2018), I will show that labelling sustainability concerns irrelevant with respect to health presupposes a very narrow concept of health and excessively strict understanding of determinants of health. Moreover, the standards of evidence that the critics appeal to are based on ideals that originate from the context of evidence-based medicine (EBM). However, transferring standards of evidence from one context to another has been disputed. As Glasziou, Vandenbroucke & Chalmers (2004, 39) put it, "different types of questions require different types of evidence". I will argue that despite what the critics of the DGAC suggestion claimed, the recommendation to include sustainability as a goal of The Dietary Guidelines for Americans can be taken to be science-based and not representing political bias.

References

Freidberg, S. (2016). Wicked nutrition: The controversial greening of official dietary guidance. Gastronomica: The Journal of Critical Food Studies, 16(2), 69-80.

Glasziou, P., Vandenbroucke, J., & Chalmers, I. (2004). Assessing the quality of research. Bmj, 328(7430), 39-41.

Kuttner, H. (2014). How to sustain sound dietary guidelines for Americans: Mission creep within the federal Dietary Guidelines Advisory Committee threatens Americans health and well-being. Hudson institute, Washington

Rucker, R. B., & Rucker, M. R. (2016). Nutrition: ethical issues and challenges. Nutrition Research, 36(11), 1183-1192.

US Department of Health and Human Services. (2015). 20152020 Dietary guidelines for Americans. Washington (DC): USDA.

Valles, S. A. (2018). Philosophy of Population Health: Philosophy for a New Public Health Era. London: Routledge

On the Assessed Strength of Agents' Bias

Jürgen Landes & Barbara Osimani

We tackle the question of how to assess a group of scientists providing testimony vis-à-vis a single scientist providing testimony. Unlike previous work (e.g., Zollman 2013) comparing different communication structures on the same group of agents (N vs. N comparison), we study how a group of agents compares to a single agent (N vs. 1 comparison).

The fallible agents considered here are either reliable inquirers or (sponsorship) biased inquirers. Intuitively, we are less likely to believe that a group of N independent agents each reporting a finding are all biased than we are to believe that one single agent providing these same N reports is biased, ceteris paribus. Why? Prior to obtaining evidence, the prior probability of a single agent being biased is equal to some value, $\bar{\rho}$ say. The ceteris paribus clause then entails that the probability of any one of N agents is biased with probability $\bar{\rho}$. The independence judgement requires that the prior probability for all N agents being biased is $\overline{\rho^N}$, clearly $\bar{\rho} >> \overline{\rho^N}$. One feels that the inequality continues to hold after receiving evidence.

We here show that this intuitive probability judgement is not always true and explain why.

From the scientists' perspective it is not ideal that the entire group is perceived more strongly to be biased than a single agent. One wonders, is there nothing the group can do to overcome this unfortunate state of affairs? The short answer is: no. There is nothing to be done. Once the prior probabilities are set, Bayesian updating kicks in and finishes the job. The group of agents falls victim to the following slogan which is borne out in our models: "A group of liars will not be believed even when they all speak the truth."

This means that the only road to salvage the standing of the group of agents is a more favourable assessment prior to reporting. This can be achieved by either a more favourable assessment of the strength of bias or by a more favourable assessment of being reliable (greater ρ). This then highlights - once again - the importance of appearances and of the choice of a subjective prior probability function in Bayesian epistemology. For science and scientists it highlights the need of good PR.

Finally, we remark that while sponsorship bias provided the motivation for our model of a biased agent, our analysis applies to all other biases (or other cognitive states) which make Type II errors more likely and Type I errors less likely. Since the list of biases is rather large (Hahn & Harris 2014), the analysis presented here may prove relevant for a variety of strands of research.

References

Ulrike Hahn and Adam J. L. Harris. What Does It Mean to be Biased: Motivated Reasoning and Rationality. Psychology of Learning and Motivation, 61:41102, 2014.

Kevin J.S. Zollman. Network Epistemology: Communication in Epistemic Communities. Philosophy Compass, 8(1):1527, 2013

Cosmology in Silico

Helen Meskhidze (University of California), Marie Gueguen (University of Western Ontario), Chris Smeenk (University of Western Ontario)

Summary
Computer simulations have played an ineliminable role in building confidence in the Cold Dark Matter (hereafter CDM) model. Since the 1970s, cosmologists have simulated ever larger cosmological volumes with higher resolution, hoping to reproduce the large scale successes of the model at smaller scales. Yet, determining whether a simulation appropriately represents its target system depends on what role the simulation plays. Similarity between the model and the target does not guarantee success; rather, the factors involved with drawing successful inferences depend on whether simulations are used to extract predictions from the model, to supplement sparse observations, or to rule out alternatives to a model. Drawing on recent literature on these issues in philosophy, symposiasts will argue that recognizing and clarifying the distinct roles played by simulations is a necessary precondition to assessing when simulations lead to valid conclusions. In addition to contributing to ongoing philosophical debates, they will challenge the idea that increasing the resolution should be the primary goal for the future of cosmological simulations.

The Explanatory and Methodological Roles of New Simulation Methods in Cosmology
Helen Meskhidze

The increasing precision of cosmological observations of the large-scale structure of the Universe has created a problem for simulators: running the N-body simulations necessary to interpret these observations has become impractical. Specifically, the parameter spaces the simulations investigate are enormous and the simulations themselves exhibit non-linear behavior. To address this difficulty, simulators have turned to machine learning (ML) algorithms that offer substantial reductions in computational cost. Though ML algorithms decrease computational expense, one might be worried about the use of ML for scientific investigations: how can algorithms that have repeatedly been described as black-boxes (even by their developers) deliver scientific understanding? While the pragmatic benefits to using ML are clear, the worry remains that in employing ML, investigators are sacrificing scientific understanding for the sake of reductions in computational expense and this worry warrants serious consideration. In this presentation, I investigate the use of two machine learning algorithms in the cosmological context. By drawing on a distinction between black-boxes themselves and black-boxing as a methodology, or the method of ignoration as I will call it, I claim that these ML algorithms ought not to be considered black-boxes but as part of a larger methodology that uses the method of ignoration. To understand how ML algorithms can fulfil this methodological role, it will be essential to first clearly outline the roles of the underlying N-body simulations. After doing so, I will claim that machine learning algorithms allow cosmologists to use the method of ignoration on their underlying simulations in order to explore statistical relationships between the responses of the simulations to various inputs. Ultimately, I argue that understanding the scientific context in which ML algorithms are employed is essential to understanding whether they ought to be understood as black-boxes

Fuzzy Modularity and Crucial Simulations

Marie Gueguen (presenter) and Chris Smeenk

One of the main challenges in assessing the reliability of simulations is their lack of modularity. A simulation is modular to the extent that it consists of quasi-isolated components, whose causal contributions to the overall outcomes can be determined. For a modular system, the praise or blame that results from comparisons with observations can be distributed to individual components rather than to the simulation as a whole. But, as many others have argued (e.g., Lenhard and Winsberg 2010), few modern simulations are modular in this sense. “Errors" are often introduced in one part of the model to avoid specific computational or numerical problems, but the effect of these changes on other components is often not transparent. “Tuning" simulations to match some set of data often introduces compensating errors, with the effect of blurring the distinctive contributions of different aspects of the simulation. N-Body simulations, for instance, approximate the density of real stellar systems by simulating fewer, but more massive, particles. Force softening is then needed to prevent divergences in the gravitational force when massive bodies get close to each other. Smoothing the gravitational potential to avoid these numerical errors, however, tends to artificially enhance matter disruption (Van den Bosch, 2017), and so to reduce the amount of substructure predicted in dark matter halos. As this example illustrates, it is difficult to determine precisely how a particular feature of simulation results depends upon the basic physics, and to separate genuine results from numerical artifacts.

Here we will evaluate one method that cosmologists have pursued, which we will call the method of “crucial simulations. Crucial experiments are meant to be cases where the contrast between competing theories can be made particularly sharp, with enough agreement on background assumptions 4to counter worries about holism. Similarly, crucial simulations focus on idealized situations where the holistic challenges associated with modularity can be minimized. In particular, with simulations it is possible to construct situations that serve to isolate and clarify the impact of very specific aspects of the simulation e.g., such that one physical factor is minimized, or by exploring enormously long time scales, or by combinations of these. Van den Bosch et al. (2017) appealed to such a scenario to test whether the amount of satellite galaxies predicted in dark matter haloes is due to matter being stripped away from satellite galaxies by tidal forces, or instead reflects the presence of numerical artifacts. The crucial simulation treated an idealized scenario such that only tidal stripping could cause matter disruption. Since the simulated outcomes were sensitive to the systematic increase of force softening, the physical hypothesis could be ruled out, confirming the artificial nature of most of the disruption. We will argue that crucial simulations can help to overcome the “fuzzy modularity" of simulations (Lenhard and Winsberg, 2010) with respect to physical components, for their effects can be ruled out one by one in these simplified scenarii where auxiliary physical hypothesis can always be turned off.

Multiple Realisability in the Sciences

Alexander Franklin (King’s College London and University of Bristol), Tuomas E. Tahko (The University of Bristol), Marion Godman (The University of Helsinki and Copenhagen University)

Summary

Multiply realised phenomena are those which are instantiated in multiple different kinds of system. Such phenomena are the subject of extensive philosophical analysis, and yet little work has been

done to compare and contrast the examples taken from the different sciences. This symposium would seek to remedy that lacuna by addressing two salient and inter-connected philosophical questions in the context of examples drawn from the different sciences.

First, in order to make sense of and define multiple realisability, one needs an account of kinds – there is multiple realisability only if the higher-level kind individuation is robust and objective. This prompts the question: what criteria need to be satisfied for kind status to be appropriate at the higher level?

Second, multiple realisability has, since the coining of the term, had implications for reduction. The autonomy of higher-level multiply realised kinds seems to preclude bottom-up explanation. And yet, multiple realisability is found not only in biology but in chemistry and physics in contexts where reduction seems otherwise assured. Thus we ask: does any acceptable characterisation of multiple realisability allow for compatibility with reduction?

Multiple Realisability and Reduction Reconciled

Alexander Franklin

Multiple realisability principally prompts the question: how is it that multiple systems all exhibit the same phenomena despite their different underlying properties? I will argue that this is a serious and difficult question, and that it poses a challenge for putative reductions. If the question cannot be answered then reduction is undermined. I argue that the account defended by Polger and Shapiro (2016) rules out any interesting answers to this question. As such, their account trivialises the multiple realisability challenge to reduction.

In this talk, I define reduction such that the principal question can be answered; this allows me to show, firstly, that multiple realisability is fairly commonplace, even within physics, and secondly that multiple realisability poses a case by case challenge to reduction within science.

I proceed by defining multiple realisability as a form of autonomy: multiply realised phenomena are autonomous with respect to changes in the lower-level constitution of the realiser systems. This is cashed out with respect to the example of electrical conductivity in metals. Potassium, Lithium, Sodium etc. all have identical electrical conductivity properties; moreover, these can all be described by the same quantitative theories of electrical conductivity. However, these metals all have different lower-level physical properties.

I claim that electrical conductivity is multiply realised across these metals, and that this poses a serious challenge to the reductionist: how is electrical conductivity thus multiply realised despite the differences between the metals? This question is addressed by adducing commonalities between the metals at the lower level, and demonstrating processes through which the features which distinguish the metals are irrelevant – that is, processes which secure the higher-level autonomy are identified. This provides an instance of a reductive strategy which, it is claimed, is applicable to multiply realised phenomena more generally.

References:

Polger, Thomas W., and Lawrence A. Shapiro. The multiple realization book. Oxford University Press, 2016.

Multiple Realisability and Higher-Level Kinds

Tuomas E. Tahko

The conventional wisdom not just in philosophy of mind but also in philosophy of science is that the multiple realisability of higher-level kinds undermines ontological reductionism once and for all, thus vindicating the existence of genuine higher-level kinds and special science laws. Most of us are familiar with the traditional arguments from the classic papers by Jerry Fodor (1974, 1997) and Jaegwon Kim (e.g., 1992), where Fodor represents the ontological anti-reductionist and Kim the ontological reductionist.

The traditional argument relies on the phenomenon of multiple realisability. If a higher-level kind cannot be reduced to a single lower-level kind – except perhaps a “wildly disjunctive” one, then it would seem to be genuine. However, there are reasons to think that it’s not really the phenomenon of multiple realisability itself that matters here (as argued by Antony 2003). Rather, what matters is whether we have some principled reasons to think that the higher-level kinds are “really there”, as Antony (2003: 8) puts it. These principled reasons, so the argument goes, are provided by the special sciences, namely, the laws and causal explanations that the higher-level kinds participate in. But for the ontological reductionist, there is something deeply unsatisfying about this response, and the reason for this was made clear already by Fodor: laws and kinds go hand in hand, so if we disagree about whether there are genuine higher-level kinds, we probably also disagree about whether there are genuine special science laws (Fodor 1974: 102). Given this, it’s striking that we can see the same anti-reductionist strategy still being repeated, over 40 years after Fodor’s version. Surprisingly, there are still aspects of this debate that have been overlooked. These can aptly be illustrated with the Fodor-Kim debate regarding the case of jade, which, as I will argue by engaging with the chemistry of minerals like jadeite and nephrite, ultimately turns on purely conventional matters. Accordingly, it will be suggested that multiple realisability as it is applied in Fodor’s arguments does not give us good reasons to reify higher-level kinds.

References:

Antony, L. (2003) ‘Who’s Afraid of Disjunctive Properties’, Philosophical Issues 13: 1–21.

Fodor, J. (1974) ‘Special Sciences (Or: The Disunity of Science as a Working Hypothesis)’, Synthese 28: 77–115.

Fodor, J. (1997) ‘Special Sciences: Still Autonomous after All These Years’, Philosophical Perspectives 11: 149–163. - Kim, J. (1992) ‘Multiple Realization and the Metaphysics

Multiple Realization in Biology: Replacing Shared Function with Homology

Marion Godman

Multiple realisability is often understood to be a mandatory commitment if we want to defend the integrity and realism of a non-fundamental science such as biology. Less ambitiously perhaps, one might think that multiple realization is an important possibility for scientific investigations of biological kinds. On either view we should be interested in what a test for multiple realisability should look like. When advocates try to defend the thesis of multiple realisation, they usually defer to the functional role of certain kinds (e.g. Putnam 1967). Even Lawrence Shapiro and Thomas Polger, prominent critics of multiple realization, propose a functionalist test for identifying multiply realised kinds. The upshot is that only artefacts like watches and corkscrews pass the test. They therefore conclude that ‘multiple realization is much less common in naturally occurring systems’ (2016, p. 73).

But one might instead wonder if their test's clear fit with artificial or technological systems instead means it's so much the worse for the test of multiple realization in natural systems!

An alternative conclusion could therefore be that one needs to go beyond the reliance on functionally specified kinds or artefacts as models if one wants to find important naturally occurring cases of multiple realizability. In lieu of a functional test for multiply realized kinds in biology, I instead propose a test via the role an instance plays as member of a homology. In biology, homologous traits which have common ancestry are recognized as a scientifically superior categorization compared to their analogous counterparts, which merely share a functional role (see e.g. Ereshefsky 2012) and so if we can find a test for homologous traits that are multiply realized will also be defending robust and objective kinds. After sketching a test, I briefly evaluate the prospect and scope for multiply realized biological homologues.

References

Ereshefsky, Marc. (2012) "Homology thinking." Biology & Philosophy 27.3: 381-400.

Polger, T. W., & Shapiro, L. A. (2016). The multiple realization book. Oxford University Press.

Putnam, Hilary, 1967. "Psychological Predicates," in W.H. Capitan and D.D. Merrill (eds.), Art, Mind, and Religion, Pittsburgh: University of Pittsburgh Press, 37–48.

Moving Past the Naturalism-Normativism Dichotomy in Philosophy of Medicine

Francis Fairbairn (Cornell University), Brandon Conley (Virginia Commonwealth University), and Shane Glackin (University of Exeter)

Summary

Since at least the 1970's, debate concerning philosophical accounts of disease, and related concepts, has typically been framed around a choice between naturalism and normativism, with naturalism in this context usually understood as the denial of the normativist claim that applying the concept of disease in part consists in making a normative judgment. However, some have argued that the naturalism-normativism dichotomy is misleading and rules out a range of viable positions. (Notable examples include Ereshefsky 2009, Simon 2007, Kingma 2014, and Broadbent 2018.) Our aim in this session is to build on critiques of the naturalism-normativism dichotomy in philosophy of medicine, to explore alternative ways of framing the debate, and to defend concrete philosophical accounts of medical concepts that do not fit neatly into the standard dichotomy because they are either both naturalist and normativist (Glackin and Conley) or neither (Fairbairn).

Coding Groups in the Mental Illness Literature

Francis Fairbairn

The question 'what is mental illness?' has generally been taken to have two possible answers: either it is a socially constructed phenomenon, or it is a natural one. In this literature, the term 'natural' is taken to code for a group of concepts including 'natural,' 'fundamental,' 'real,' and 'non-normative.' Similarly, the term 'socially constructed' is taken to code for concepts such as 'social,' 'non-fundamental,' and 'reducible.' My claim is that these 'coding relationships' (or 'coding groups' in my language) are such that:

1. They are especially invisible or hard to track.
2. They are often inherited from previous assumptions or views.
3. They inhibit research programs by foreclosing discussion in certain areas.
4. They perpetuate real social and epistemological harms.

On my picture, concepts within coding groups code for each other so subtly and so strongly that we tend to use them interchangeably without realizing. And yet, when we do interchange them in this way, it changes the flavor of the dialectic, sometimes radically. For example, if we end up thinking that:

1. In order to be 'natural' (as opposed to socially constructed) mental illnesses must be value-free, objective, non-social … etc and

2. The natural is more robustly 'real' than socially constructed

and we also think that:

1. It is ethically important to reify mental illnesses so that they are appreciated as genuine, sometimes debilitating, conditions…

...then one's desire to reify mental illness as part of a project to make mental illnesses less stigmatized will lead one to argue that (e.g.) mental illnesses are non-social.

The upshot of my account is that mental illness should not be assessed against the categories 'natural' vs 'social' because these categories represent problematic coding groups. To establish this, I tie together historical analysis and conceptual analysis; the history of the debate shows the effect of social and political ideology on the requirements of success for analyses and even on the meaning of the question at stake. In this way, the debate has not just 'switched' via a clean break but rather the underlying inferences/inferential ideologies are still there. Our dialectics inherit the sins of their fathers.

How to be a Naturalist and a Social Constructivist About Disease, Part 1

Brandon Conley

The apparent conceptual connection between disease and dysfunction has been used to develop naturalistic accounts of disease, as opposed to social-constructivist and normative accounts, by serving as a tool for providing explications of the disease concept which are either non-normative, or normative in a reductive sense. However, this strategy presumes that normative judgements, including social ones, are not already part of the scientific practice of making dysfunction attributions.

I present an independently attractive framework, and some simple formal machinery, for understanding the role dysfunction attributions play in scientific practice. On this view, any arbitrary norm can serve as a descriptive point of reference for expressing causally relevant information about the system under scrutiny; however, the goals of a given discipline or research program will make some norms especially useful for doing scientific work. Given that the scientific work could, in principle, be done by any arbitrary norm, it is possible that the norm arises from social judgments. Beyond mere possibility, I argue that the goal of some sciences, including medicine, to control and manipulate, in addition to describing and explaining, is in fact best served by the kinds of social-normative judgments emphasized by normativist and social-constructivist accounts of disease. Importantly, a social constructivist account developed along the lines I suggest would count as naturalistic in the broad sense in which the label 'naturalist' is used in philosophy more generally, and which motivates naturalistic accounts of disease, namely that the account is methodologically, ontologically, and epistemically continuous with the natural sciences, and the disease concept does not reduce to simply "that to which we apply the term 'disease'." For an account of disease to be naturalistic in the important sense, does not imply the concept is non-normative or non-socially grounded.

Demonstrating the full potential of this brand of naturalistic social-constructivism requires developing specific proposals about the normative judgements underlying attributions of disease, a task Shane Glackin will undertake in a companion presentation. However, I close by outlining one argument that Shane will develop in more detail, and which will serve as a base for addressing various problem cases in the literature: Our brand of naturalistic social-constructivism enjoys an advantage over rival views because it can capture the apparent explanatory power of both the selected-effect (Griffiths and Matthewson 2016) and biostatistical (Boorse 1975) accounts, but has additional explanatory resources because it includes a role for social-normative judgements. The intuitions supporting rival views can be explained by conceptions of innateness facilitating a move from knowledge about evolutionary history or statistical norms to judgements of social acceptability, or at least non-manipulability. However, showcasing the additional explanatory work social norms can.

How to be a Naturalist and a Social Constructivist About Disease, Part 2

Shane Glackin

By way of a "proof of concept" for the claim that a broadly naturalistic account of dysfunction not only makes space for, but positively encourages, a role for social norms, I start by outlining a simple set of socially evaluative criteria for the attribution of disease-status: a biological or behavioural state is judged to be a disease, briefly, just in case it is regarded:

1. as not representing a tolerable state of affairs; but
2. neither as representing a moral failing of the individual concerned.
3. as not being worth reorganising society so as to fully neutralise the relative impairment caused by the state; but
4. as being nevertheless worthwhile to divert resources to "correct" and/or ameliorate it.

This allows us to supplement the naturalistic account of dysfunction given by Brandon Conley in the preceding talk with a set of normative social grounds for selecting the particular subset of the broad class thus defined which are to count as diseases, in a way that accords with common intuitions about the disease-concept's extension.

This demonstrates the compatibility of naturalism and social constructivism. To show that this is not merely a coherent position, however, but an attractive one, we need to go further. I therefore elaborate and extend Brandon's closing argument. Canonical accounts of the disease concept such as Boorse's (1975) Biostatistical Theory, Griffiths & Matthewson's (2016) Selected Effect Account, and Wakefield's (1992) Harmful Dysfunction account look particularly plausible as applied to veterinary diseases, which are only rarely as controversial as human cases can be; conversely, social constructivist accounts can be difficult to apply outside the context of human societies. Our approach explains the social evaluative judgements that underlie intuitions about disease-attributions, as well as how evaluative differences underlie the clash of intuitions in problem cases. We can therefore go beyond Boorse, Wakefield, and Griffiths & Matthewson by explaining in terms of our own theory why those accounts produce intuitively plausible results, especially in "natural" and non-socialised or pre-social cases; the evolutionary and statistical phenomena they invoke do not themselves directly play a role in a proper account of the disease concept, but they do influence the way the social evaluative judgements which play a central role in our theory are made. We can also go beyond them in another way; by explaining the disease-status of conditions affecting non-functional body-parts, such as appendicitis, in the same way as other diseases, rather than by extension or disjunctive courtesy.

The New Reduction: Formal, Conceptual, and Physical Perspectives

Neil Dewar (Munich Center for Mathematical Philosophy), Samuel C. Fletcher (University of Minnesota), Laurenz Hudetz (London School of Economics), Katie Robertson (University of Birmingham)

Summary

This symposium continues and unites two traditions of approaching intertheoretic reduction: the Nagelian and the limiting-case approach. We provide novel contributions to both strands of research and show how these can be combined to make progress. Fletcher shows how the idea of limiting-case reduction can be made precise by endowing classes of models with extra topological (or topologically inspired) structure that encodes similarity relations between models. Robertson adopts an account of reduction-as-construction in the Nagelian tradition and analyzes the role that abstractions and changes of variables play in reductions. To illustrate this, she looks at coarse-graining in statistical mechanics. Dewar shows that abstraction and the introduction of higher-level variables for entities composed of lower-level objects can be handled even in the framework of first-order logic by generalising the Nagelian definition in a way suggested by results from the recent debate about theoretical equivalence. Hudetz argues that the two traditions can be fruitfully united by an explication acknowledging that a reduction may involve steps of different kinds (including both definitional reconstructions along the lines of Nagel and taking limits). The proposed explication makes Robertson's reduction-as-construction account more precise and shows how definitional constructions can be combined with Fletcher's limiting-case relations.

Reduction, Construction, and (Generalised) Translation

Neil Dewar

The most natural way to make precise Nagel (1961)'s classic definition of reduction is as follows: T_1 reduces to T_2 just in case there exists a translation from T_1 to T_2 (i.e. an association of simple L_1-terms with complex L_2-terms, in such a way that every theorem of T_1 is mapped to a theorem of T_2). As is well-known, there are various problems with this characterisation of reduction, such as the absence of any mention of approximation, the assumption that our theories are presented as collections of sentences, and so on.

In this paper, the particular problem that I want to investigate is that, as stated, reduction is incompatible with differences in ontology between the two theories. For, if there is a translation from T_1 to T_2, then any model of T_2 is associated with a (unique) model of T_1. This pair of models is naturally interpreted as representing a pair of possible worlds, the latter of them supervening on the former. But if we are working with the standard notion of translation, as stated for single-sorted logic, then the two models have the same domain. So, on the face of it, this account of reduction is unable to handle cases where the objects dealt with by a theory at one level are different to those dealt with by a theory at another. But such cases are entirely typical of putative cases of reduction: for instance, statistical mechanics concerns particles whereas thermodynamics deals with objects composed of such particles (e.g. gases); psychology deals with brains, but neuroscience of neurons; etc. (Even if we think that thermodynamics will ultimately turn out to be about particles, that result should be a consequence of successful reduction, not a presupposition for it.)

Fortunately, however, recent work has extended the relevant notions of definability to multi-sorted logic, in such a way that new sorts (hence, new objects) can be defined: the relevant theory is known as the theory of “generalised definitions" (Andréka et al., 2008) or “Morita extensions" (Barrett and Halvorson, 2016). In particular, this apparatus offers a means of extending a theory of individual objects to a theory that treats mereological sums of those objects (see Halvorson (2016)); so we are able to treat the notion of so-called micro-reduction, i.e., of cases of reduction in which one theory deals with the parts of the objects dealt with by the other (Oppenheim and Putnam, 1958)

The Topology of Intertheoretic Reduction

Samuel C. Fletcher

Nickles (1973) first introduced into the philosophical literature a distinction between two types of intertheoretic reduction. The first, more familiar to philosophers, involves the tools of logic and proof theory: “A reduction is effected when the experimental laws of the secondary science (and if it has an adequate theory, its theory as well) are shown to be the logical consequences of the theoretical assumptions (inclusive of the coordinating definitions) of the primary science" (Nagel, 1961, p. 352). The second, more familiar to physicists, involved the notion of a limit applied to a primary equation (representing a law) or theory. The result is a secondary equation or theory. The use of this notion, and the subsequent distinction between so-called “regular" and “singular" limits, has played a role in understanding the prospects for reductionism, its compatibility (or lack thereof) with emergence, the limits of explanation, and the roles of idealization in physics (Batterman, 1995; Butterfield, 2011).

Despite all this debate, there has been surprisingly no systematic account of what this second, limitbased type of reduction is supposed to be. This paper provides such an account. In particular, I argue for a negative and a positive thesis. The negative thesis is that, contrary to the suggestion by Nickles (1973) and the literature following him, limits are at best misleadingly conceived as syntactic operators applied to equations. Besides not meshing with mathematical practice, the obvious ways to implement such a conception are not invariant under substitution of logical equivalents.

The positive thesis is that one can understand limiting-type reductions as relations between classes of models endowed with extra, topological (or topologically inspired) structure that encodes formally how those models are relevantly similar to one another. In a word, theory T reduces T’ when the models of T’ are arbitrarily similar to models of T - they lie in the topological closure the models of T. Not only does this avoid the problems with syntactically focused account of limits and clarify the use of limits in the aforementioned debates, it also reveals an unnoticed point of philosophical interest, that the models of a theory themselves do not determine how they are relevantly similar: that must be provided from outside the formal apparatus of the theory, according to the context of investigation. I stress in conclusion that justifying why a notion of similarity is appropriate to a given context is crucial, as it may perform much of the work in demonstrating a particular reduction's success or failure.

I illustrate both negative and positive theses with the elementary case of the simple harmonic oscillator, gesturing towards their applicability to more complex theories, such as general relativity and other spacetime theories (Fletcher, 2014, 2016).

Complex Reductions

Laurenz Hudetz

The basic idea of intertheoretic reduction is that to reduce a theory T to a theory T' is to recover T from T'. There are two prominent accounts of what it is to recover T from T': a logical account due to Nagel (1961) and a limit-based account noted by Nickles (1973). According to Nagel's account, it means that the laws of T are derivable from the laws of T' by means of coordinating definitions (and auxiliary assumptions). According to the limit-based account, it means that T is a limiting case of T'. These accounts gave rise to two separate traditions of research on reduction, each with its own proponents and critics.

In this talk I argue that the divide between the two traditions can be overcome. Arguing over which one of the two accounts is the right one rests on a false dilemma. It overlooks that many reductions are actually complex. By a 'complex reduction' I mean a reduction that involves several reductive steps of potentially different types. On the proposed account, both taking limits and logically reconstructing one theory from another (along the lines of Nagel) are legitimate reductive steps. I claim that, in many cases, steps of both types are necessary. In particular, logical reconstructions play a role when theories are formulated in different formal frameworks. For instance, the work of Ehlers (1981, 1986) has shown that a rigorous analysis of the relationship between the Newtonian theory of gravitation (NG) and general relativity (GR) involves both a reconstruction and taking a limit. One first reconstructs NG in a formal framework that also contains the models of GR. This yields the geometrized Newton-Cartan theory of gravitation (GNG), which takes curved spacetime as primitive. Since GR and GNG lie in a common background framework, one can directly compare their models and make precise what it means that sequences of models of GR converge to models of GNG by defining a topology on the class of framework structures.

To be able to check whether there is a complex reduction relation between given theories, we need clear criteria for the basic reductive steps. While Fletcher provides an explication of limiting case relations between theories in his talk, I propose an explication of the concept of theory reconstruction. Roughly speaking, a reconstruction of a theory T from a theory T' is given by a reconstruction functor from (a definable subcategory of) the category of models of T' to the category of models of T. The notion of a reconstruction functor makes precise what it is to construct models of T from models of T' in a uniform way. An important similarity between this explication and Fletcher's account of limits is that it also concerns classes of models endowed with extra structure. This makes it relatively easy to combine theory reconstructions with limiting case relations in Fletcher's sense to form complex reductions. Moreover, since the proposed explication takes morphisms between models into account, it establishes a link between questions of reducibility and recent work on categorical equivalences between theories (e.g., Weatherall, 2016).

The Relationship between Reduction and Abstraction

Katie Robertson

List (2017) has offered a category-theoretic framework which precisifies talk of "levels of description" implicit in much of the debate around inter-theoretic reduction. Central to this framework is the notion of abstraction: an abstraction map σ is a surjective function which will generally map an equivalence class of lower-level states, s_1, s_2, . . . to a single higher-level state, s'. The informal gloss on this abstraction map is that abstraction throws away irrelevant details (Strevens, 2008). Whether the system's state is s_1, s_2, . . . is irrelevant for the higher-level description in terms of s': it matters which equivalence class represents the system, but not which element of that equivalence class does. Such a case of abstraction is demonstrated by coarse-graining in statistical mechanics. More generally, changing variables-such as moving to a collective variable by summation-as described by Knox (2016), can be irreversible and is an example of abstraction.

In this talk, I claim that this idea of abstraction connects to reduction: changing variables is key part of constructing a higher-level description from a lower theory. This, I claim, is key to my account of reduction: reduction-as-construction. According to this account, to reduce a higher-level theory Tt to a lower-level theory Tb, the equations and quantities of Tt must be constructed from the equations and quantities of Tb.

This new account of reduction will allow me to shed light on two traditional problems:

1. *The classic case study of thermal physics*: (i) it allows us to reconcile the lower-level time-symmetry with the higher-level time-asymmetry in statistical mechanics, and (ii) it allows us to explain how equilibrium thermodynamics proceeded in ignorance of the nature of matter and yet is true in virtue of more fundamental theories, such as quantum statistical mechanics.
2. *Autonomy*: When discussing the relationships between different scientific levels of description, the concern has often been to establish how the higher-level, or special, sciences are underpinned by more fundamental theories, but nonetheless "float-free" to a certain extent (Fodor, 1997). In other words, higher-level theories retain a degree of "autonomy". For example, the success of thermodynamics in spite of ignorance about the nature of matter suggests a certain epistemic autonomy-even if thermodynamics is subsequently reduced. But there are many definitions and degrees of autonomy-e.g., epistemic, methodological, and explanatory. I will argue that these forms of autonomy can be explained within my account of reduction. If the higher-level dynamics fulfil the mathematical condition of "autonomy" familiar from ordinary differential equations (Robinson, 2004), they will have no explicit dependence on (i) the lower-level details or (ii) time (and thus no "covert" dependence on these lower-level details). Subsequently, these lower-level details do not make a difference for the dynamical evolution of the higher-level variable. This explains why knowing these details didn't matter for the higher-level theory (epistemic autonomy) nor for higher-level explanations (explanatory autonomy). During reduction, the higher-level equations are constructed from the lower-level equations, thus demonstrating which lower level details are dynamically irrelevant and so the degree to which higher-level theory is autonomous.

The Impact of the Replication Crisis on Philosophy: Two Case Studies

Dr. Suilin Lavelle (The University of Edinburgh), Dr. Richard Morey (Cardiff University), Dr. Hugh Rabagliati (The University of Edinburgh)

Summary

Contemporary philosophy of mind is closely integrated with the empirical sciences. Many of the most respected theories of vision, thought, consciousness and free will draw on findings from neuroscience and psychology. But what happens when the psychological findings upon which this work is based do not replicate? How are the philosophical theories affected?

This symposium brings together a philosopher and two psychologists who are involved in large-scale replication studies that question data central to key areas in philosophy of mind. All three presenters examine the role of replication in advancing science, through the lens of work that is of direct relevance to philosophers.

Dr. Richard Morey presents a large scale replication attempt of the 'Action Sentence Compatibility Effect', which is often cited in favour of embodied cognition; and Dr. Hugh Rabagliati will discuss on-going complications in infant cognition replications, which impact key philosophical theories about nativism and concept acquisition. Dr. Suilin Lavelle is a philosopher whose work is directly affected by the on-going replication crisis in infant cognition, and will discuss (a) how Dr. Morey's and Dr. Rabagliati's work impacts the related philosophical theories, and (b) how the replication debates are a living example of scientists arguing about what constitutes 'data', bringing forward aspects of theory-choice so famously developed by Kuhn (1962).

A Large-Scale, Multi-Lab Test of the Action-Sentence Compatibility Effect: Results and Implications

Dr Richard Morey (Cardiff University)

The Action-Sentence Compatibility Effect (ACE; Glenberg and Kaschak, 2002) is a speeding of response times to evaluate sentences when the actions described in those sentences are congruent with the actions necessary for a response. For instance, if a participant must evaluate the meaningfulness of the sentence "You handed Meghan the book," and the correct response is away from the participant's body, the described movement (handing to) is compatible with the response. The ACE effect is presumed to be due to strong links between cognitive systems for understanding language about motor actions and motor systems themselves. Although the ACE has been described in dozens papers, Papesh (2015) recently reported a number of failures to replicate the effect concluding that the effect may not be as robust as previously believed. A large-scale, multilab attempt to replicate the effect designed by the proponents of ACE did not yield evidence for the effect. I will discuss what the null replication teaches us about the value of large-scale replications for scientific discovery, for how we should view "solutions" for the replication crisis.

Interpreting Research on Infant Cognition in light of the Replication Crisis

Dr. Hugh Rabagliati

Experiments on the cognitive and social abilities of human infants have provided some of the most vivid illustrations of how psychological experiments can contribute toward answering key questions in philosophy of mind. These experiments, for example, have suggested that infants possess innate knowledge (e.g., so-called "core knowledge" of physical objects, Spelke, 1990), have provided evidence for symbolic mental representations (Marcus et al., 1999), and suggested precocious skills at learning to represent the beliefs of others (Onishi & Baillargeon, 2004).

But now, against a background of tumult in the social and biological sciences, the collective evidence provided by studies of infant cognition is being re-evaluated. Recent work in disciplines such as medicine (Ioannidis, 2005), psychology (Open Science Foundation, 2015), and economics (Camerer et al., 2018), have suggested limits to the global scientific record: Statistical evaluations suggest that many (perhaps most) research findings may be overstated, direct replications of prior work frequently fail to duplicate the previously-found results, and it has become painfully clear that, in many fields, incentives such as publication bias had caused unintentional deviations from scientific best practices.

This paper will provide an overview of how these issues have impacted, and will continue to impact, research into infant cognition. Part 1 describes the challenges of measuring cognition in infancy, with reference to the dominant experimental paradigm, i.e., the analysis of infant looking times. It will discuss how looking times (e.g., to novel, unfamiliar, or unexpected stimuli) are typically interpreted, and the mechanics of collecting looking time data, focusing on aspects of data collection that may not be obvious to those looking in from outside the field.

Part 2 describes reasons for caution when evaluating infant cognition research. These include concerns about publication bias, about researcher degrees of freedom (e.g., how researchers make decisions when processing data, such as excluding participants), and about so-called HARKing

(hypothesising after the results are known) in which researchers, when writing up the results of an experiment, retrofit their supposed hypotheses to dress up post-hoc interpretations as predictions.

Part 3 describes solutions to these problems. Solutions include 1) Large-scale, multi-lab replication studies, in which groups of laboratories work together to standardize an experimental protocol and collect far larger datasets than ever before; 2) Meta-analytic aggregation of prior work, which allow scientists to draw firmer conclusions by combining the results of many different published studies, and adjusting for publication bias; 3) Improving the psychometric validity of experimental paradigms, which is to say, showing that the paradigm measures what it was intended to measure.

Thus, this paper hopes to provide a primer allowing non-experts to give a more informed, nuanced and critical evaluation of research into infant cognition.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., ... & Altmejd, A. (2018). Nature Human Behaviour, 2(9), 637.

Ioannidis, J. P. (2005). Why most published research findings are false. PLoS medicine, 2(8), e124.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Science, 283(5398), 77-80.

Onishi, K. H., & Baillargeon, R. (2005) Science, 308(5719), 255-258.

Open Science Collaboration. (2015). Science, 349(6251), aac4716.

Spelke, E. S. (1990). Cognitive science, 14(1), 29-56.

Replication, Variation and Theory Choice

Dr. Suilin Lavelle

Many philosophers draw on work from the sciences to support their theories. In the philosophy of mind alone, findings in neuroscience, developmental psychology, cognitive ethology and anthropology inform respected theories of cognition and consciousness. The apparent crisis facing the psychological sciences as well-cited findings to replicate should therefore be of paramount concern to those philosophers who rely on them in developing their positions. This paper expands on themes presented by my colleagues working in embodied cognition and infant cognition, to better understand how philosophers whose work is directly affected by replication failure should react to the crisis.

The complexities of working with infants, and the messy data emerging from large-scale replication attempts of infant work, provides rich material for philosophers of science. From the perspective of replication there are important questions about the epistemic value of failed replications: while this has been discussed in science more generally (Leonelli 2018), these insights have yet to be applied to developmental psychology. There are critical debates to be faced about what should be considered 'data': the infants whose looking time suggests they can attribute beliefs to others, or those infants who do not look in the 'right' way? This question has a direct impact on philosophical theories. Those of a nativist leaning tend to focus on infants' (and childrens') successes on false belief tasks, taking it as evidence of an innate BELIEF concept. Those who prefer empiricist, learning theories, find childrens' errors to be key data, taking it to evidence a child trying out new hypotheses, sometimes successfully and sometimes not. These questions, about what constitutes data and what the core explanandum of theories of cognition should be, contribute to on-going, live examples of theory choice, and reflect core questions in the philosophy of science.

The second theme raised by these talks is that of variation. Advocates of embodied cognition cite the ACE as evidence for their view, but as Dr. Morey argues, the ACE may not be a stable effect. While one failed replication does not cast doubt on the entire embodied cognition project, there are important questions to be considered about the reasons for the failed replication. The most pressing of these concerns variation in embodied cognition. Perhaps the failed replication can be understood in terms of variation across individuals' propensity to engage in embodied thought processes. Or perhaps the variation is within individuals, with subtle cues priming one towards or against embodied thought processes. If variation is the key, then advocates of embodied cognition need to accommodate this within their theories. Alternatively, if the ACE simply does not exist at all, then this also needs to be accounted for.

Effective Field Theories: Top-down and bottom-up

*Richard Dawid (**University of Stockholm**), Michael Stöltzner **(University of South Carolina)**, Porter Williams (University of Southern California), Martin King (University of Bonn)*

Summary

In recent years, effective field theories have attracted the interest of philosophers of science. They are not only a traditional technique of physics and elementary particle physics, but they also invite general considerations about scientific realism, models, and explanation. An effective theory can be embedded within an existing, or purported, more fundamental theory; for instance, the Standard Model of elementary particle physics is a top-down effective theory within any unified theory of the fundamental forces. But the same techniques can also be applied bottom-up as a tool to search for such a fundamental theory by postulating higher order processes and comparing such calculations with the available precision data. The aim of this symposium is to discuss the relationship between both kinds of effective field theories and to what extent this duplicity within the same theoretical

framework of quantum field theory affects the suitability of effective field theories for scientific realism and model explanations.

Effective realism: of fundamental concern?

Porter Williams

In recent years, effective field theories (EFTs) have received increased philosophical attention. Particular focus has been directed toward the possibility of, and strategies for, extracting realist commitments from EFTs. Several philosophers, myself included, have presented their explorations of "Effective Realism" as a turn away from an approach to scientific realism according to which the relevant interpretive question is, "If this theory were literally true, what would the fundamental structure of the world be like?" in favor of a scientific realism that seeks to answer the question, "Given that this theory is approximately true over a certain range of scales, what is the world approximately like at those scales?"

This raises a worry that Effective Realism devalues, or even sits in tension with, the pursuit of more fundamental theoretical frameworks than quantum field theory. This is one of the core pursuits of contemporary high energy physics; insofar as one of the virtues of Effective Realism is supposed to be that aligns better with physical practice, sitting in tension with that practice would put the Effective Realism in a particularly awkward spot.

In this talk, I argue that this worry is misguided. I highlight several methodological virtues of Effective Realism for the pursuit of more fundamental theoretical frameworks and argue that Effective Realism, in fact, better coheres with this aspect of contemporary high energy physics practice than a scientific realism that emphasizes fundamental ontology.

Real Beyond Effective

Richard Dawid, Michael Stöltzner

In a recent paper, Porter Williams argues against the position that a physical theory's interpretation spells out the realist import the given theory would have if it were true. He proposes a non-fundamentalist take on theory interpretation that, unlike the former position, allows for a realist view of well-established effective theories in high-energy physics. The present paper argues for keeping separate two aspects of fundamentalism that go together in Williams' reasoning. The position we call counterfactual fundamentalism amounts to the assertion that an interpretation of a theory requires imagining that theory as a fundamental theory about the world. The position we call reductionist fundamentalism amounts to the assertion that the true fundamental theory is the only judge over the question whether an effective theory is approximately true. While we fully concur with Williams' rejection of counterfactual fundamentalism, we argue that rejecting reductionist fundamentalism abandons a core element of scientific realism. We present three arguments in support of this claim.

First, the two forms of fundamentalism are conceptually independent from each other. It is possible to endorse each of them without the other. Axiomatic approaches implicitly assume counterfactual fundamentalism – even though they might consider the choice of the axioms not as an ontological commitment. The approximate truth of the effective theory, on the other hand, may be linked to the fundamental theory without insisting on the full consistency of the effective scheme.

Second, the degree of ontological or structural similarity between effective and fundamental theory constitutes an important issue in judging the stability of physical conceptualization. In the past, it has played an important role in understanding the extent to which a physical theory could be called approximately true. Even if progress in physical theory building (for example based on the increasing relevance of duality relations in fundamental physics) led to a situation where issues of structural or ontological similarity could not be addressed anymore in a meaningful way, we argue that this fact would itself amount to an important message for scientific realism that would be lost if one chose an entirely non-fundamentalist view.

Third, discussing realism only in its effective form would amount to committing a similar mistake as counterfactual fundamentalism. It would ignore the importance of the prospects of theory building for the way physicists understand the status of their current theories. That importance is closely related to the generally reductionist character of physical reasoning.

None of the three presented points devaluates Williams' observation that the issue of robustness plays an important role in the physicist's understanding of the status and explanatory value of scientific claims. In this light, we suggest that effective scientific realism and fundamental scientific realism (in the reductionist sense) both have their roles to play and should not be pitted against each other. Reductionist fundamentalism leads to a stronger form of realism than its effective cousin. It raises important questions of physical and philosophical understanding, which reach out beyond what can be addressed from an entirely non-fundamentalist perspective.

Are Standard Model Effective Field Theories Models?

Cristin Chall, Martin King, Peter Mättig, Michael Stöltzner

Model-independent searches for new physics have become increasingly popular in recent years as LHC data continues to push beyond the existing Standard Model (BSM) models to the corners of their parameter spaces. These approaches have gotten recent philosophical attention by McCoy & Massimi who cogently argue that simplified models exhibit the four main features of the models-as-mediators framework by generalizing the representational requirements. A first aim of our paper is to see how far this framework for models can be extended to another recently popular model-independent approach called the Standard Model effective field theory (SMEFT). More generally, this case also allows us to study a nuanced search process and address questions about representation, the notion of a target system, and at what point something ceases to be a model.

SMEFT is a bottom-up rather than top-down effective field theory, like a simplified model, and as such it is not obvious that it also qualifies as a model merely by virtue of its being embedded into a more fundamental theory. SMEFTs apply the theoretical machinery of quantum field theory in a different way than top-down approaches. In SMEFT one does not simplify a Lagrangian, but begins with a SM Lagrangian and expands it with additional operators in order to parameterize the effects of potential new physics and, by constantly confronting the SMEFT with experimental data, to obtain constraints on or clues about it.

In order to adequately answer the question as to whether SMEFTs are models even in a generalized sense, we identify various stages of the SMEFT in the discovery process.

1. At the earliest stage, there are very few assumptions and SMEFT has on the order of 2500 additional dimension-6 operators and a great deal more of dimension-8 and higher. Here, it has no specific new physics target, cannot make any predictions, and does not allow for any ontological commitments to new physics objects.

2. If one introduces assumptions to narrow the focus, the SMEFT has only a few operators and can make some falsifiable predictions.

3. However, we argue that it is only after experimental evidence has been discovered that the representational nature of SMEFT becomes clearer. At this third stage, actual experimental deviations

constrain the operators and the basis in which they are formulated. Here, SMEFT is modelling new physics and the difference in ontological commitments towards new physics between simplified models and the SMEFT largely fade out.

Since representation of a target system is a key feature of models, we take our examination of the various stages to indicate an interesting philosophical result, namely, that whether or not a SMEFT is a model is not a function of its theoretical characteristics, but dependent on the status and specificity of experimental knowledge.

Explanation and Effective Field Theories

Martin King

Since the Higgs boson discovery in 2012, there have been no indications of physics beyond the standard model (BSM). Concrete BSM models have been pushed to the edges of their parameter spaces and as a result model-independent approaches, such as effective field theories (EFTs), have become increasingly popular in particle physics. The EFTs employed in new physics searches at the Large Hadron Collider (LHC) are what are known as bottom-up EFTs and are quite distinct from the top-down ETFs that have been more thoroughly treated in the philosophical literature. The aim of the paper is to examine the role of bottom-up EFTs in potentially explaining new physics.

The paper proceeds by first arguing that top-down EFTs can be understood as abstract and idealised versions of higher-energy (or UV-complete) theories. Similar points have been argued in philosophical work on EFTs and renormalisation group equations, such as (Batterman 2002), (Batterman and Rice, 2014), (Bain, 2013), and others. I will briefly present the Fermi theory of beta decay and make the case that claims about its being explanatory can be supported by an abstraction and idealisation process from the SM.

The paper then contrasts this in three ways with a bottom-up EFT, in particular the Standard Model EFT (SMEFT). For the SMEFT, the UV-complete theory is not known and it is not known where the theory will break down and new physics will become relevant. And so the first distinction is that there is no guarantee about the predictive, and hence explanatory, power of a bottom-up EFT.

The second distinction is that the SMEFT is not an abstraction or a idealisation of the SM, and cannot borrow its explanatory power. One constructs the SMEFT by expanding the SM Lagrangian with an infinite series of effective operators that parameterise the effects of BSM physics. Physicists make certain assumptions about UV physics in order to reduce the number of operators, but which operators are actually relevant is not yet known. The SMEFT is not optimised, per (Strevens, 2008), as it contains many irrelevant operators and cannot highlight explanatorily relevant features.

A third difference is that the SMEFT plays a very different role in the eventual explanation of new physics, namely, it is only a stepping stone on the way to an explanation. This can be seen in how it is used in LHC searches. Indications of new physics will result in non-zero coefficients for some set of the operators, which physicists can then use to constrain the structure of a BSM model that may explain the physics that underlies the deviation.

While the SM serves as the UV-complete theory that grants Fermi theory its ability to explain beta decay, the SMEFT is probative, tentative, and used to constrain the structure of future BSM models.Thus, EFTs can differ significantly with respect to their ability to explain, depending on whether they are top-down or bottom-up.

Structure and Composition in Chemistry

Karoliina Pulkkinen (The University of Cambridge), Vanessa A. Seifert (The University of Bristol), Geoffrey Blumenthal (The University of Bristol)

Summary

Our proposed symposium considers two of the most important topics in the philosophy of chemistry: structure and composition. Structure can refer either to the organisation and bonding of atoms in molecules or to the structuring of information in chemistry. This symposium includes a paper on each topic. The first paper examines the nature and reality of molecular structure by considering an understanding of chemical bonds as real patterns. The second examines the role of values in constructing periodic systems of elements. The third paper focuses on the notion of composition; it examines how chemists arrived at views on chemical composition which became standard, and why some experiments were taken as being crucial by some researchers.

Values and the Periodic System

Karoliina Pulkkinen

The periodic system is a representation that structures all of the chemical elements in a manner that effectively displays similarities between the elements. This paper examines the role of values in constructing the periodic system. In contrast to the existing accounts that examine periodic system through the lens of values, I focus on the role of values in the context of developing the systematisations rather than the context of their justification or reception (e.g. for predictive accuracy, see e.g. Lipton, 1990; Barnes, 2008; for coherence, see Schindler, 2014),

This paper demonstrates how three of the chemists who competed for the priority of the discovery of the periodic systems elevated different values during the development of their systems. Where the English chemist John Newlands emphasised what he called a "simple relation," the German chemist Julius Lothar Meyer stressed the importance of carefulness in assessing the quality of observations used as a basis for his system. Unlike his colleagues in England and Germany, the Russian chemist Dmitrii Ivanovich Mendeleev did not emphasise simplicity or carefulness as strongly. Instead, Mendeleev was also more permissive in including more dubious findings to his system, as long as the system was as complete. I will argue that Mendeleev especially emphasised the importance of completeness in establishing a systematisation of elements

After demonstrating how values guided the construction of the periodic systems, I argue that emphasising specific values also importantly influenced the chemists' uses of those systems. I argue that Meyer's emphasis of carefulness allowed him to use the periodic system a theoretical tool for identifying errors in experimental results. With Mendeleev, I show that his emphasis on completeness in including all of the elements (and conveying many of their properties) supported predicting the properties of undiscovered chemical elements. In other words, where Meyer's valuing of carefulness paved the way for using his system to identify errors in experimental results, Mendeleev's valuing of completeness supported making predictions.

The examples of Meyer, Mendeleev, and Newlands shows that shifting our focus from the context of justification to the context of development offers new avenues for studying values in science. In this particular case, moving to the context of development allows identifying relationships between values emphasised during the construction of scientific representations and their further uses.

Barnes, E. C. (2008). The Paradox of Predictivism. Cambridge: Cambridge University Press.

Lipton, P. (1990). Prediction and prejudice. International Studies in the Philosophy of Science, 4(1), 51–65. http://doi.org/10.1080/02698599008573345

Schindler, S. (2014). Novelty, coherence, and Mendeleev's periodic table. Studies in History and Philosophy of Science Part A, 45, 62–69.

http://doi.org/https://doi.org/10.1016/j.shpsa.2013.10.007

The Chemical Bond as a Real Pattern

Vanessa A. Seifert

A central concept which is invoked in chemistry and in quantum chemistry in order to describe the structure of a molecule is the chemical bond. Given this, a pressing philosophical question is whether the chemical bond exists and what sort of thing it is. This question is primarily discussed in the context of Hendry's distinction between the structural and the energetic conception of the chemical bond.

The structural conception takes chemical bonds to be 'material parts of the molecule that are responsible for spatially localized submolecular relationships between individual atomic centers' (Hendry 2006: 917). The structural conception is taken as supporting an understanding of chemical bonds as entities. The energetic conception takes 'chemical bonding' to signify 'facts about energy changes between molecular or supermolecular states' (Hendry 2006: 919). The energetic conception remains agnostic as to whether the chemical bond is an entity (or as to whether it even exists) and it is consistent with an understanding of chemical bonds as properties of a molecule. The metaphysical interpretation of each conception allegedly creates a tension between the two conceptions because the former is consistent with an understanding of chemical bonds as entities, whereas the latter is consistent with an understanding of chemical bonds as either fictional entities, or real properties of molecules.

I argue that this tension can be resolved in a manner that supports the reality of chemical bonds. Specifically, if one takes the two conceptions as representing distinct yet incomplete intensions of the same referent (i.e. the chemical bond), then both conceptions can be invoked to mutually support an understanding of chemical bonds as patterns within a molecule. Such an understanding of chemical bonds is also supported by how chemistry and quantum chemistry each describe and pictorially represent chemical bonds.

Several questions need to be addressed in order to sufficiently support the reality of chemical bonds as patterns. First, if a chemical bond refers to a pattern within molecules, then what is it a pattern of? Secondly, assuming that chemical bonds are patterns, what is the respective 'noise' in the chemical and quantum chemical descriptions of a chemical bond, and what is the role of 'noise' in predicting a molecule's structure? Thirdly, is there sufficient empirical evidence to support that the elements of this pattern are real and not merely apparent? I examine these questions in light of the literature on real patterns and briefly outline the advantages of understanding chemical bonds as real patterns. Examining the nature and reality of chemical bonds in the context of the literature on real patterns provides a novel perspective through which one can understand the nature of the chemical bond, but also through which one can reevaluate the tenability of structural realist accounts in the philosophy of science.

References

Hendry R.F., 2006, 'Two Conceptions of the Chemical Bond', Philosophy of Science, Vol. 75, No. 5, pp. 909-920.

On the Nature of apparently Crucial Experiments in Chemistry

Geoffrey Blumenthal

This article follows two accounts analysing the detailed difficulties of the crucial experiment account in biology (Weber 2009; Baetu 2017). It analyses equivalent issues concerning two experiments that proved to be decisive in chemistry. These experiments are the reduction of mercury calx by Hermbstädt and Klaproth in 1793, and the large-scale experiment on the synthesis of water in 1790.

Several factors underlaid the extent to which these experiments were taken to be crucial. Each of these experiments was the last in a long chain of published versions of the type of experiment, during which the difficulties of the type of experiment had been worked out in detail. In each case, the factors such as impurities which affected the experiment were understood in considerable detail due to previous experiments. In each case, the precautions needed to minimise the effect of impurities had been specified before the experiment. No further versions of the experiments were published in detail. In the first case, the tested hypotheses were mutually exclusive and no other hypotheses seemed possible. In each case, the final experiment had been designed to reduce the fragility of the results as far as was then practicable.

Each of these experiments was a cause of the public change of view by prominent researchers, including Gren (1794) and Westrumb in the first case and Kirwan (1791) in the second. Gren and Kirwan each stated that their changes of mind were based on not arguing against the "truth". Yet in each case, the decision did not relate solely to the experiment, but also to a cluster of other associated factors, including other experiments and publications. In effect, each experiment acted as a focus for a cluster of concerns, and was selected as being a crucial indicator, rather than actually being fully decisive on its own. Accordingly, neither experiment was taken as being crucial by all parties. Yet the result in both cases was the focus of a considerable shift of opinion among contemporary chemists.

Baetu, Tudor. 2017. "On the Possibility of Crucial Experiments in Biology." British Journal for the Philosophy of Science (1 Sept 2017):1-23.

Fourcroy, Antoine-François, Louis Vauquelin, and Armand Séguin. 1791. "Mémoire sur la Combustion du Gaz Hydrogène dans des Vaisseaux clos". Annales de Chimie, 8:230-307.

Gren, Friedrich A. C. 1794. Letter to van Mons. Journal der Physik 8:14-18.

Hermbstädt, Sigismund. 1793. "Neue Bestätigung einer chemischen Grundwahrheit, den Gehalt des Sauerstoffs, im wasserfreyen Quecksilberkalke betreffend". Chemische Annalen, 1793.1:303-314.

Kirwan, Richard. 1791. Letter to Crell, Chemische Annalen, 1791.1:425-426.

Séguin, A. 1791. Suite du Mémoire sur la Combustion du Gaz Hydrogène dans des Vaisseaux clos. Annales de Chimie, 9, 30-50.

Weber, Marcel. 2009. "The Crux of Crucial Experiments: Duhem's Problems and Inference to the Best Explanation". British Journal for the Philosophy of Science, 60:19-49.

BSPS Open Session Abstracts

The abstracts below are listed in alphabetical order by TITLE

Causation, Intervention, and Responsibility

Enno Fischer
Leibniz Universität Hannover

According to interventionist theories of causality, we are interested in causal claims because they enable us to interact effectively with the world. Interventionists have also tried to explain the function of more specific causal claims that concern actual causation. They argue that while causal claims generally tell us where we could intervene in order to change the effect, claims of actual causation tell us where we should intervene (Hitchcock and Knobe 2009). Interventionist accounts of the function of actual causation have been criticised widely, in particular, by philosophers and psychologists who see a close relation between causal judgement and the ascription of responsibility. These opponents argue that an exclusive focus on interventions is neither inherently plausible nor does it fit the data (e.g. Alicke et al. 2011).

In this talk I will present a novel taxonomy of causal claims that is based upon distinguishing three senses in which contributors to the debate have been using the term “actual causation”. Based on the taxonomy I will provide a more fine-grained analysis of the function of actual causation. In particular, I will indicate which kinds of causal claims are better explained in terms of responsibility than in terms of interventions. This sheds new light on the limitations of the interventionist theory, which has often been considered particularly plausible in explaining the reasons for our interest in causation.

(1) I will begin with distinguishing three senses in which the term "actual causation" has been used. First, actual causation (AC1) refers to claims about sequences of token-events as opposed to claims about types of events. Second, actual causes (AC2) are contrasted with potential causes. Merely potential causes are factors that can bring about a certain effect but in contrast to actual causes they do not bring the effect about, for example, because they are pre-empted. Third, actual causation (AC3) is understood as referring to factors that are salient in bringing about an effect and, therefore, are to be distinguished from background conditions.

(2) I will then develop a taxonomy of actual causal claims by clarifying the relation between the three senses of actual causation. First, I will argue that AC1 describes a proper subset of AC2. Token causal claims like "c caused e" entail that c not only could but also did bring about e. By contrast, AC2 claims do not need to be singular. For example, we can make type level claims about pre-emption such as those that concern redundancy in biological systems. Second, I will argue that AC1 is largely independent of AC3. Our causal claims tend to be selective no matter whether they concern tokens or types of events. For example, the claim that a short-circuit caused a fire is selective because a number of background conditions had to be in place for the fire to occur. Likewise the general claim that short-circuits cause fires only holds under certain background conditions. Third, I will use simple examples of graphical causal models to illustrate that claims of AC2 are based on claims of AC3.

(3) I will then use the taxonomy in order to provide a more fine-grained analysis of the function of actual causation.

(3.1) First, interventionists argue that we identify those factors as actual causes (AC3) that violate some norm and that such a preference for norm-violating factors is rational (e.g. Hitchcock and

Knobe's 2009). From an interventionist perspective norm-violating factors are the most important factors because they represent the best targets for intervention. According to my taxonomy, AC3 claims can be either singular (AC1) or not. So far interventionists have implicitly focused on type claims of AC3. Yet, as I will show, interventionists face difficulties in accounting for a particular class of singular AC3 claims. These are claims that concern the past and require very specific background conditions. It is sometimes not clear how these causal claims are to be generalized in order to be exploitable for future interventions. By contrast, such claims play an important role when we ascribe responsibility, or so I will argue.

(3.2) Second, interventionists address type-level claims of AC2. Hitchcock (2017) argues that such claims help us design goal-directed strategies in contexts with complex causal structure. I will object that some such claims are more plausibly related to responsibility. In particular, it is difficult to motivate the distinction between cases of late pre-emption and symmetric overdetermination from an interventionist perspective. Consider late pre-emption first. Suzy and Billy both throw stones at a bottle. Suzy throws a little earlier, hits the bottle and thereby pre-empts Billy's hitting. In the corresponding case of overdetermination Suzy and Billy throw at the same time and hit at the same time. An intervener who is interested in preventing the bottle from being hit needs to intervene on both Suzy and Billy, no matter whether she deals with a case of overdetermination or of late pre-emption and, thus, no matter whether Billy is an actual cause or not. By contrast, for the ascription of responsibility the difference between late pre-emption and overdetermination is relevant. In the overdetermination case Billy is more likely to be made responsible for the bottle's shattering than in the late-preemption case, or so I will argue.

(4) I conclude that a closer look at the various forms of actual causal claims reveals limitations of the interventionist's functional account of causation. This is an important result because explaining the function(s) of causal reasoning has often been considered one of the major strengths of interventionist approaches (Woodward 2014).

Hitchcock, Christopher and Knobe, Joshua. Cause and norm. The Journal of Philosophy, 106(11):587–612, 2009.

Alicke, Mark D., Rose, David, and Bloom, Dori. Causation, norm violation, and culpable control. The Journal of Philosophy, 108(12): 670–696, 2011.

Hitchcock, Christopher. Actual causation: What's the use? In Making a Difference, pages 116–131. OUP, 2017.

Woodward, James. A functional account of causation. Philosophy of Science 81(5):691-713, 2014.

Computing and Modelling: Analog Vs. Analogue

Philippos Papayannopoulos
Hebrew University of Jerusalem

Overview: The talk examines aspects of the interplay between computing and scientific practice, with some greater focus on analog computing. Although very much neglected today, analog computing has been the main computational paradigm used in science up until the 1980s, when only replaced completely by digital computing. The motivation for this study is that investigating what people have been doing in scientific practice for centuries and calling it "computation" can give us a new perspective on the nature of computing per se, as well as on epistemic methods in science such as analogical reasoning; perspectives that are often missed insofar as computing is solely considered in its modern classical (silicon-based) form. In this talk, we are particularly concerned with gained insights into the following three matters: (a) the nature of computation as an epistemic process and the role of representation in it (b) the analog/digital dichotomy, and (c) the interrelationships between analog computational modelling, analogue (physical) modelling, and analogical reasoning.

Details: We first look very briefly at some paradigmatic cases of analog computing in history; namely, certain examples of mechanical computing devices as well as electronic analog computers. We propose that the long-standing role of computation in scientific practice suggests a conception of it as an epistemic process relative to agents, wherein representation has an indispensable role. The agents are always aided by some kind of machinery or instrumentation (possibly just pen and paper). The basic idea is that “representation” is a key underlying feature of every case of computing and, hence, it can be very helpful in understanding the scope and limits of various forms of computation (from classical to non-classical, such as digital, analog, quantum, etc.). More precisely, we argue that each computational process actually involves two, conceptually distinct, types of representation: one related to how the input and output of the calculated problem are represented (encoded and decoded) in the computing machinery, and one related to the evolution of the computing machinery through time (which, crucially, doesn’t have to be a strictly accurate description of the actual workings of the physical system but only accurate enough for the level of analysis that is relevant to the computing process; idealizations also have a role to play here).

We then go over different existing accounts of the distinction between "analog" and "digital" computation, and classify them in four groups, based on the amount of emphasis they put on representation, and on whether they require the existence of 'analogies' between the computing and the computed systems. We then put forward a new account, largely inspired by Goodman (1976). The motivation for that is that we need a precise characterisation of “analog” in order to be able to compare analog computational modelling with analogue modelling and analogue reasoning in science. The proposed account in this work is based on the semantic and syntactic properties of the involved representations in computations. But, following Goodman, we maintain a view that the relevant properties are not agent-independent but hinge on the intended reading of the representation by the computing agent, thereby adopting a pragmatic approach to the analog/digital dichotomy. We put the proposed account to the test by showing its consistency with how we pre-theoretically characterize paradigmatic cases of analog and digital computing. To this end, we apply the account to theoretical models, such as Turing machines and General-Purpose Analog Computers, as well as actual computing devices, such as PCs, slide rules, and Electronic Analog Computers. Besides accounting for these paradigmatic cases, we also show how the proposed account is able to explain universality, a property apparent only in those systems that are commonly characterised as "digital".

Now, the developed account provides us with a framework for comparing the contributions of analog computational models with those of analogue-physical models. We distinguish between two kinds of the latter: models that serve only demonstrative purposes, such as the famous model of DNA constructed by Watson and Crick, and scale models used in similitude theory, such as the scale model of a bridge in a wind tunnel. We then examine foundational aspects of both analog computing and analogue modelling (the latter as grounded in dimensional analysis and similitude theory). Our working example is modelling the airflow around an aeroplane wing. We consider analog computer models of that system (on an analog electronic computer made by Operational Amplifiers) and scale models of it in a wind tunnel. The outcome is that although analog computational modelling relies on constructing a computational system that is governed by the same differential equations as the target system, it is however not an instance of analogical reasoning, contrary to analogue modelling. The two practices also differ with respect to (a) their mathematical foundations (Shannon's theory of analog computation vs. dimensional analysis), (b) the kind of knowledge required about the target system in advance, (c) the kind of epistemic functions the two practices fulfil. Thus, despite their apparent similarities, analog computational modelling and analogue modelling are in fact orthogonal practices for studying a certain target system.

Reference:

Goodman, N. (1976). Languages of Art (2nd ed.). Hackett Publishing

---

Confidence: A New Dimension of Scientific Knowledge?

Margherita Harris
London school of Economics and Political science

The Intergovernmental Panel on Climate Change (IPCC) has long recognized that the question of how best to characterize and communicate uncertainties is far from a trivial matter. Through the joint and iterative effort of the many experts that have participated in the various IPCC assessments, an important distinction between two types of uncertainty has emerged. The first type is expressed with a likelihood metric and the second with a confidence metric. Although the assessment of both likelihood and confidence involves expert judgments of some kind, it is the latter that dominates the expert judgment scene. The aim of this talk is to understand what the confidence metric is today and what it might look like tomorrow.

In the first part of the talk I will discuss what I consider the two most problematic aspects involved in the characterization of the confidence metric in the most recent uncertainty guide. The first has to do with the puzzling bifurcation of evidence and agreement in the evaluation of confidence; I will argue that the apparent tension arising from this bifurcation is not unresolvable per se, but that any attempt to resolve this tension would have to begin with giving an explicit and satisfactory answer to the following two questions:

(i) by what criteria are the evidence and the agreement metrics evaluated so that they are clearly untied/independent from one another?

(ii) with respect to what should the evidence and the agreement metrics be judged?

The second issue with the characterization of confidence has to do with the ambiguity surrounding the relationship between the likelihood and confidence metrics. In particular I will discuss the AR5's practice of downgrading likelihood and thereby upgrading confidence, which although is evidently intended to account for sources of uncertainty not adequately addressed in the formal analyses, it is unclear (due to the lack of transparency surrounding this practice) the extent to which they are accounted for (and indeed how they should be accounted for in the first place).

In the second part of the talk, I will further argue that the IPCC's current practice of reporting findings expressed in terms of imprecise probabilities, qualified by qualitative confidence judgements does not seem to provide information that can be sensibly integrated into any account of decision making and that this is a substantial problem. The main, if not only, reason an institution such as the IPCC is in place is to give relevant and useful information regarding the state of knowledge in studies of climate change to agents that will ultimately want to make decisions based on this information. If it is not at all clear how one should interpret this information so as to make rational decisions based on it, then the project seems to have (at least partly) failed. So the question I shall be concerned with is the following: what role, if any, can these qualitative confidence judgments play in decision making? After arguing that they can theoretically play a workable role in decision making, I will consider and evaluate two recent proposals for a new uncertainty framework for future IPCC assessments that seem to address the above problem, but in very different ways. In Mach et al.'s (2017) proposal, the confidence metric once and for all leaves the scene, making way for a whole new likelihood metric, one that would be always explicitly based on subjective probabilistic assessments and that would reflect all sources of uncertainty. In Helgeson et al.'s (2018) proposal, the confidence metric stays on the scene, but the IPCC's current practice of reporting findings at only one confidence level is called into question. I will argue that one should approach with caution Mach et al.'s proposal to effectively remove a dimension of knowledge (i.e. confidence). However, I will further argue that it is only through a better understanding of robustness analysis in climate science that one can adequately assess whether Helgeson et al.'s proposal to keep the confidence metric (but make better use of it) is feasible in practice.

References:

Mach, K. J., M. D. Mastrandrea, P. T. Freeman, and C. B. Field. 2017. "Unleashing Expert Judgment in Assessment." Global Environmental Change 44: 1–14.

Helgeson, C., R. Bradley and B. Hill. 2018. "Combining Probability with Qualitative Degree-of-Certainty Metrics in Assessment." Climatic Change 149, nos 3–4: 517–25.

Confirmation Based on Analogical Inference: Bayes Meets Jeffrey

Alexander Gebharter and Christian J. Feldbacher-Escamilla
University of Groningen (Gebharter), Duesseldorf Center for Logic and Philosophy of Science (DCLPS), University of Duesseldorf (Feldbacher-Escamilla)

Sometimes evidence for a hypothesis cannot be directly observed. This might be the case if the evidence is inaccessible for theoretical reasons. An example would be evidence for certain hypotheses about the dynamics of black holes (Winsberg, 2009; Dardashti, Thébault, & Winsberg, 2015). But even if observing the evidence for a certain hypothesis is theoretically possible, we still might not possess the know-how or the right tools to measure it. Alternatively, the costs to produce the evidence or to build the tools required to measure it might be too high. In such cases, evidence cannot be accessed for different practical reasons. The existence of widely recognized moral reservations might also make it impossible to observe evidence. Producing evidence to directly confirm a certain psychological hypothesis might, for example, require surgical interventions on the brains of subjects.

Cases in which a hypothesis H cannot be directly confirmed by observing evidence E obviously cause trouble for scientists. Though such a hypothesis cannot be directly confirmed, it might make perfectly reasonable true or false claims about the world. So is there really no way to confirm (or disconfirm) such a hypothesis? One possible option consists in trying to find systems s’ that are similar (or analogous) enough to the systems s about which H claims this and that. One could then formulate a corresponding hypothesis H’ for these similar enough systems s’. Contrary to the systems s, these systems s’ might produce evidence E’ that can be observed directly. Now the hope is that our original hypothesis H can somehow be confirmed on the basis of observing E’. After all, H’ makes a claim about systems s’ that is analogous to what H claims about systems s. If E’ can somehow be used to confirm H, then it seems that there is a possibility to empirically assess hypotheses whose corresponding evidence cannot be observed (for whatever reasons).

Some kind of confirmation on the basis of analogical reasoning is clearly applied in sciences such as biology, climate science, economics, medicine and pharmacology, etc. However, whether evidence E’ that directly confirms a hypothesis H’ can be used to confirm an analogous hypothesis H, is, in some sense highly controversial (see., e.g., the critique of Duhem, 1991, pp. 97ff or Bartha, 2010, sec. 1.9). A recent approach put forward by Dardashti, Hartmann, et al. (2015) seems to support the view that confirmation based on analogical inference is quite reasonable. They propose a Bayesian analysis of confirmation on the basis of analogical reasoning. In particular, they argue that if the systems described by H and H’ (at least partially) share the same structural features, there might be a connection between H and H’ that establishes probability flow between evidence E’ and hypothesis H. This seems to be everything required for E’ to (indirectly) confirm H Bayesian style. So confirmation based on analogical inference would simply be a certain kind of Bayesian confirmation according to Dardashti, Hartmann, et al.’s approach.

In this talk we take up Dardashti, Hartmann, et al.’s (2015) idea to make sense of confirmation based on analogical inference in a Bayesian framework. We first identify three more or less classical types of analogical inference. We then introduce and illustrate Dardashti, Hartmann, et al.’s approach by means of a simple toy example. We argue that their approach—in its original version—covers only one of the types of analogical inference and show that it can be expanded in such a way that it also covers a second type. We then generalize their approach to scenarios in which common causes play the same role shared structures (or analogies) play in their account. This move will turn out to be quite straightforward, since from a formal point of view, common causes work exactly like shared structures. We also highlight several possible problems with the view that evidence E’ for a hypothesis H’ can confirm another hypothesis H making a claim about a totally different system. We finally develop a model for the missing type of confirmation by analogy and suggest to supplement Bayesian update by Jeffrey conditionalization for cases in which direct evidence for the hypothesis of interest is unavailable.

References

Bartha, P. (2010). By parallel reasoning. the construction and evaluation of analogical arguments. Oxford: Oxford University Press.

Dardashti, R., Hartmann, S., Thébault, K., & Winsberg, E. (2015). Confirmation via analogue simulation: A Bayesian analysis. Retrieved from http://philsci-archive.pitt.edu/12221/

Dardashti, R., Thébault, K., & Winsberg, E. (2015). Confirmation via analogue simulation: What dumb holes could tell us about gravity. British Journal for the Philosophy of Science, 68(1), 55-89.

Duhem, P. M. M. (1991). The aim and structure of physical theory. Princeton: Princeton University Press.

Winsberg, E. (2009). A tale of two methods. Synthese, 169(3), 575-592.

Conjunctive Explanations

David Glass and Jonah N. Schupbach
University of Ulster (Glass), University of Utah (Schupbach)

Sometimes two explanations are better than one. This may happen, for example, in cases of "explanatory pluralism" when theories each do qualitatively different explanatory work. An object's existence can be explained either by referring to its causes or its function (Wright, 1976). One hypothesis may explain an event by telling us a causal-mechanical story leading up to the event, while another may explain the same event by referring to a nomic regularity that the event instantiates—as in Salmon's (2001) "friendly physicist" example. In such cases, accepting a plurality of explanations provides us with a richer understanding of the explanandum. But sometimes even distinct explanations of the same type can be better together than apart. In general, several explanations are better than one just when the explanatory benefits of accepting them all outweigh the costs (in complexity and otherwise). In such cases, we will say that the distinct potential explanations in question are "conjunctive", and we will refer to the above observation as the phenomenon of "conjunctive explanation."

This talk explores the logic and epistemology of conjunctive explanations, as they occur in scientific practice. We attempt a formal investigation into the precise conditions under which the phenomenon of conjunctive explanation arises. A prima facie account asserts that any potential explanations of a phenomenon that are consistent can constitute conjunctive explanations. If potential explanations are compatible, why not accept them all? But we argue that this plausible idea is incorrect. We highlight cases in which consistent explanations nonetheless may compete strongly with one another—in the sense of (Schupbach and Glass, 2017). In such cases, explanatory considerations may compel us to choose between candidate explanations instead of accepting them all (despite their consistency). This point motivates a second intuitive idea: that conjunctive explanations are those proffered by non-competing hypotheses. This account fails too, however, due to cases in which hypotheses that compete nonetheless provide conjunctive explanations.

Comparative approaches to evidential support provide another way to pursue an account of conjunctive explanation. We can say that the conjunctive explanation offered by h1&h2 for evidence e is to be preferred to the explanation put forward by h1 alone if e supports h1&h2 over h1 (for some discussion along similar lines, see Crupi et al., 2008; Atkinson et al., 2009). It is well-known that an absolute approach to evidential support is problematic since it fails to take probabilistic relevance into account, but it is particularly weak in the current context since it is always the case that $P(h1\&h2|e) \le P(h1|e)$ and so the conjunctive explanation could never be preferred. According to an alternative approach based on the law of likelihood (or equivalently on the ratio measure of evidential support), e would support h1&h2 over h1 if $P(e|h1\&h2) > P(e|h1)$. In contrast to the previous case, this approach faces the opposite problem of making it too easy for the conjunctive explanation to be preferred since it focuses only on the likelihoods and has no cost associated with the inclusion of an additional hypothesis. A more reasonable approach would require an appropriate trade-off to be made between increased likelihood and increased complexity.

To address this issue, we explore the potential of other Bayesian measures of evidential support. In particular, we investigate the difference, likelihood ratio and relative distance measures and show that all of these measures provide a more adequate approach to conjunctive explanation than the law of likelihood. In particular, we show that while the inequality $P(e|h1\&h2) > P(e|h1)$ must be satisfied if the conjunctive explanation is to be preferred by these measures, this is no longer a sufficient condition for doing so. Instead, a more demanding constraint, which differs according to the measure, must be met if the conjunctive explanation is to be preferred. We will explore some parallels with the problem of irrelevant conjunction (or tacking problem) for measures of evidential support. A popular response to this problem is to show that various measures result in a lower degree of support when an irrelevant hypothesis is conjoined to a relevant one (though this response does not work for the ratio measure). Hence, we might think of conjunctive explanation as presenting a parallel problem of relevant conjunction for the ratio measure.

Finally, we consider a more direct approach to accounting for conjunctive explanation. Since conjunctive explanations are those that are explanatorily better together than apart, this approach builds directly upon criteria of explanatory goodness; i.e., this approach explores what it takes for distinct explanations to be explanatorily better together than apart. We mine the budding formal literature explicating proposed explanatory virtues, including simplicity (Forster and Sober, 1994), unification (Myrvold, 2003), consilience (McGrew, 2003), coherence (Shogenji, 1999; Glass, 2018), and power (Schupbach and Sprenger, 2011). And we consider various senses in which net explanatory virtue could be effectively gained by accepting multiple potential explanations of the same explanandum. The results clarify probabilistic conditions under which certain tradeoffs in distinct virtues result in conjunctive explanation scenarios.

Throughout the talk, we demonstrate the relevance and potential fruitfulness of this formal work by showing how it applies to actual instances of conjunctive explanation in the history of scientific thought. We conclude with some philosophical speculation to do with a common approach to theoretical, explanatory reasoning. Theoretical scientists and philosophers of science could be served, we suggest, by more often considering candidate hypotheses as possible conjunctive explanations—as opposed to considering them, by default, as epistemic competitors.

Atkinson, D., Peijnenburg, J., Kuipers, T.(2009). Philosophy of Science, 76:1–21.

Crupi, V., Fitelson, B., Tentori, K.(2008). Thinking and Reasoning, 14(2):182–199.

Forster, M., Sober, E.(1994). British Journal for the Philosophy of Science, 45:1–35.

Glass, D.(2018). British Journal for the Philosophy of Science, axy063.

McGrew, T.(2003). British Journal for the Philosophy of Science, 54:553–567.

Myrvold, W.(2003). Philosophy of Science, 70:399–423.

Salmon, W.(2001). Hon, G., Rakover, S., editors, Explanation: Theoretical Approaches and Applications, 61–91. Kluwer Academic, Dordrecht.

Schupbach, J., Glass, D.(2017). Philosophy of Science, 84(5):810–824.

Schupbach, J., Sprenger, J.(2011). Philosophy of Science, 78(1):105–127.

Shogenji, T. (1999). Analysis, 59(4):338–345.

Wright, L.(1976). Teleological Explanations. University of California Press, Berkeley.

Descriptions Don't Always Close the Gap in the Mapping Account

Hannah Tomczyk
University of Cambridge

Mathematics seems to be very different from the physical world. And yet, scientists use mathematics to explain and predict the behaviour of physical systems. Why does that work? In virtue of what can mathematics be successfully applied to the physical world? The `mapping account' gives an answer to this question. It is, roughly, that we can successfully represent a part of the world with a mathematical formalism if the structures of the two are similar. That they are similar means that there is an isomorphic/homomorphic mapping between them. This answer seems intuitive enough, but it comes with the presupposition that parts of the physical world have a structure that can be isomorphic/homomorphic to another structure. But `structure', as it is used in this context, is a term defined in mathematical set-theory. A structure is therefore an abstract object that is -- at least prima facie -- not located in the physical world (e.g. Suárez (2003), Frigg (2006)). The challenge for proponents of the mapping account is therefore to explain how to connect abstract structures to physical systems. Otherwise, there is a gap in the mapping account between structures and the world, and the mapping account is incomplete.

As a response to this problem, Nguyen and Frigg (2017) have put forward their `extensional abstraction account', according to which descriptions have to be included in the mapping account. They act as the bridge between concrete parts of the world and abstract structures. The authors suggest that for a successful application of mathematics, we have to i) give a description of a part of the world, ii) abstract from the physical nature of what we described, thereby obtaining a set of abstract objects and relations, iii) define a structure that contains those abstract objects and relations, and iv) show that this is similar to the mathematical structure in question. I believe the authors have succeeded in clarifying how abstract structures can be connected to parts of the world. However, I argue that there are many cases of applied mathematics in science for which the account does not work. That is because including descriptions in the suggested way closes the gap in the mapping account only if the descriptions are true. If a description is false, the structure derived from it might have nothing to do with what the system is actually like. And then, a similarity between that structure and the mathematical structure does not explain the success of the mathematical representation.

The condition that descriptions have to be true is a problem for the extensional abstraction account, because there are many cases of successful science in which descriptions are false (or at least, are believed to be false). Famously, there are historical cases, like scientific representations that involve reference to ether or phlogiston. But also in modern science, people sometimes give a description of the unobservable part of their target system that they believe to be false, and achieve a successful representation anyway. For example, in ultracold atom physics, the Lorentz model is used. It suggests that the electron in an atom is connected to the core with a tiny spring, which makes the atom act like a classical harmonic oscillator that can be driven by a light field. Of course, no one believes that an electron is connected to an atomic core with a spring. So the description is clearly false. Yet the model is very successfully used, for example in experiments with light traps for atoms. The extensional abstraction account cannot explain cases of that kind. But without involving descriptions, as the extensional abstraction account does, the gap in the mapping account persits. So proponents of the mapping account need to find a different way of closing the gap for the problematic cases.

Direct Inference in the Material Theory of Induction

William Peden
Durham University

John D. Norton's Material Theory of Induction (MTI) is one of the most intriguing recent additions to the philosophy of induction. Norton's account has many merits, but his theory has also attracted considerable criticisms. A particular point of controversy is whether the MTI commits us to inductive scepticism: many critics have argued that it succumbs to the Problem of Induction, as it implies inductive scepticism when combined with common epistemological claims. I defend the MTI against this criticism: provided that the theory can be combined with a suitable theory of direct inference, then material inductivists nothing to fear from the Problem of Induction.

According to Norton, inductions are not justified by formal relations between evidential reports and hypotheses. Instead, their justification requires the combination of the evidence with relevant background knowledge of uniformities in nature, provides these uniformities give us reasons to expect our samples to be (at least approximately) representative of our target populations. However, unlike similar traditional theories of induction, there is no general principle of the uniformity of nature in the MTI. Instead, there is a vast multiplicity of background information about local uniformities that can justify belief in the representativeness of some samples and thus justify particular local inductions.

For example, we can expect that all or almost all of a newly discovered species of insect will reproduce by the same mechanisms as a sample, because we know that species of animal usually only have one basic reproduction mechanism each. (We know that there is not perfect uniformity here: aphids and some other species can reproduce both sexually and asexually.) In contrast, the fact that a sample of a species of bird share a common colour of their plumage does not justify the hypothesis that the plumage all members of that species is that colour, because our background knowledge includes the information that this characteristic tends to vary within species of birds.

To date, critics of Norton's theory (Thomas Kelly, Samir Okasha, John Worrall etc.) have not objected to inferences from local uniformities to the representativeness of samples. Instead, they appeal to the Humean Problem of Induction and argue that the MTI is especially vulnerable to it. They offer a regress problem: presumably, we can only know local uniformities via antecedent inductions. However, according to the MTI, these antecedent inductions will also require justification via local uniformities. The regress is vicious: at each step, there is an essential element in the process of justification that is unfinished. If the regress terminates at a point where there are no available local uniformities to justify the initial induction, then the MTI commits us to saying that induction is a house built on sand. If the regress is infinite, then it is always vicious, and the MTI commits us to saying that induction are eternally tardy borrowers, offering promissory notes that are always ultimately unfulfilled.

My response to this criticism is that there is an answer to the Problem of Induction for which the MTI is very felicitous. Philosophers such as Donald Williams (The Ground of Induction, 1947) have noted that we know some local facts about populations via the mathematical principles of combinatorics and the fact that we have n-fold samples of them. For example, given a well-defined random variable, a very large proportion of the means of the 3,000-fold subsets of any finite population will match the population mean within a small margin of error. Therefore, Williams argued, if we know the mean for such large subsets, then we can infer the approximate population mean with (up to) a high degree of probability, and thereby confirm various hypotheses about it. This is the 'Combinatorial Justification of Induction' (CJI).

The key step in this argument is the inference from (1) the uniformity facts about a target population's large subsets to (2) the claim that the particular observed sample is representative of the target population. This reasoning, often called "direct inference", is already crucial in the MTI: I used it earlier in the insect reproduction example. It is Furthermore, whereas in some theories of induction (e.g. Bayesianism) we must justify a suitable prior probability distribution to allow for such reasoning, in the MTI there are no priors beyond those we know via the relevant local background information. (For instance, we can accept scientific theories that imply physical probabilities for an event.) Finally, since the CJI uses the local facts about a population, it is consistent with the MTI, yet these facts do not require antecedent inductions.

However, Norton's own discussions of direct inference are far too permissive. They do not account for the crucial role of defeaters in direct inference. For instance, even if you know that 99.9% of the balls in an opaque box are red, you can still have good reasons to expect that the next ball to be drawn will not be red, if you know that the ball will be drawn from the top of the box and that 99.9% of the balls at the top are not red. Similarly, the literature on the CJI is full of ways that inferences from large subsets can be unreasonable – biased sampling procedures, gerrymandered reference classes ('grue'), information about narrower reference classes, etc. These examples not only threaten the use of the CJI, but the MTI in general.

If a suitable theory of direct inference can be combined with the MTI, then these problems would be addressed, and thus the CJI would be available for answering the Problem of Induction within the MTI. The main issue would be just connecting exiguous inductions with mature science. However, as Alan Hájek has argued, developing a viable theory of direct inference is an extremely slippery task; there are severe problems with the standard approaches. I close by arguing for some intuitive principles about direct inference and desiderata for Nortonian theories of direct inference and note that Henry Kyburg's theory of "Evidential Probability" meets these. The integration of Evidential Probability and Norton's theory of induction is a natural next step.

Direct Perception and Computation

Manolo Martínez
Universitat de Barcelona

Introduction

In this paper I examine the notion of ecological information, developed as part of ecological psychology, the picture of the mind spearheaded by Gibson (2014). Gibson's main interest was to substitute what he saw as the excessively rationalistic mainstream in cognitive science with a view of cognition where the interactions of agents and their environment is first in the order of explanation.

The two main theoretical innovations in ecological psychology are, first, the idea of affordances, or aspects of the world that are of direct relevance to, and actionable by, a certain agent; and, second, the idea of ecological information. This is supposed to offer an alternative construct to Shannon's understanding of information (Chemero); one, in particular, able to ground the direct perception of affordances. That is, the fact that sensory states carry ecological information about affordances is offereed as an explanation of how behavior attuned to the presence of affordances is possible without agents needing to perform the complicated, inference-like computation that undergirds this behavioral sensitivity in the representationalist approach.

In this paper I argue, first, that ecological information is just Shannon information—one that, depending on the ecological theorist that develops it, will meet further conditions, not unknown of in representationalist quarters (Dretske). And, second, that the fact that information meets these stringent conditions doesn't mean that perception based on it will be direct (in the sense of not mediated by computation over information-bearing vehicles): it will not be in the very widespread case of crossmodal perception.

Ecological Information is Shannon Information

The central informational relation postulated in ecological psychology is the one that holds between an energy array and an affordance. Energy arrays record the structure of the energy passing through a certain, small-ish spatio-temporal region—such that an agent can tap into it if it is situated in that spatio-temporal region. So, for example, light in many points of a room is structured in a way that carries information about objects situated elsewhere in the room. Affordances are, roughly and in ways that differ slightly from account to account, properties that can be directly acted upon by the agent—say, eatable, or scalable.

The energy array carries ecological information about an affordance if the affordance and the array are connected in such a way that certain configurations in the latter nomologically necessitate the presence of the former (Turvey et al. 1981). This nomological necessitation used to be seen as necessary direct perception by early ecological psychologists: it makes misrepresentation of the relevant feature impossible. (but see Chemero 2011 on direct perception as 'causal coupling'.) On the other hand, the main notion of informational connectedness in Shannon's information theory (Shannon & Weaver 1998; Cover & Thomas 2006) is the mutual information that holds between two random variables.

Nothing in the Shannon formalism, or the ecological information literature, prevents energy arrays on the one hand, and affordances, on the other, to be modeled as random variables. In fact this is tacitly done by ecological psychologists and commentators: (Golonka 2015, p. 238), (Chemero 2011). Furthermore, if one of these random variables carries ecological information about the other, then the mutual information between them is nonzero.

The upshot is that we can equate the Turvey-Mace-Shaw brand of information with Dretske's early 1981 proposal, where there is no information without certainty; and more contemporary developments in so-called "semantic information" (Scarantino 2015; Stegmann 2015; Green 2018) are equivalent to more contemporary developments in ecological information.

Ecological Information Without Direct Perception

Furthermore, there is no guarantee that energy arrays carrying ecological information about an affordance will enable the direct perception of said affordance. Not even if information is carried in the strictest, Turvey-Mace.Shaw sense. I will present a simple model in which this is the case. The model exploits the fact that two different sensory surfaces of an agent are sensitive to two different aspects of the wnergy array (say, structure in ambient light and structure in the pattern of sound waves), and the presence or absence of the affordance is necessitated by information contained synergistically in both such aspects, but not in any in particular. In such a situation, typical of so-called cross-modal perception (Bertelson & De Gelder 2004; Nanay 2014; Vroomen, Bertelson & de Gelder 2001), under very minimal assumptions about brain architecture, exploitation of ecological information won't be possible without computation, if at all.

References

Bertelson, P & De Gelder, B 2004, 'The psychology of multimodal perception', Crossmodal space and crossmodal attention, pp. 141–177.

Chemero, A 2011, Radical embodied cognitive science, MIT press.

Cover, TM & Thomas, JA 2006, Elements of Information Theory, New York: Wiley.

Gibson, JJ 2014, The ecological approach to visual perception: Classic edition, Psychology Press.

Golonka, S 2015, 'Laws and conventions in language-related behaviors', Ecological Psychology, vol. 27, no. 3, pp. 236–250.

Green, M 2018, 'Organic Meaning', in A Capone (ed.), Further Advances in Pragmatics and Philosophy, Springer.

Nanay, B 2014, 'Empirical problems with anti-representationalism', Does perception have content, pp. 39–50.

Scarantino, A 2015, 'Information as a probabilistic difference maker', Australasian Journal of Philosophy, vol. 93, no. 3, pp. 419–443.

Shannon, CE & Weaver, W 1998, The Mathematical Theory of Communication, University of Illinois press.

Stegmann, U 2015, 'Prospects for Probabilistic Theories of Natural Information', Erkenntnis, vol. 80, no. 4, pp. 869–893.

Turvey, MT, Shaw, RE, Reed, ES & Mace, WM 1981, 'Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981)', Cognition, vol. 9, no. 3, pp. 237–304.

Vroomen, J, Bertelson, P & de Gelder, B 2001, 'Auditory-visual spatial interactions: Automatic versus intentional components', Out of mind, pp. 140–150.

Disease, Sex, Senescence and Pregnancy. Who's Normal?

Jonathan Grose
University of Southampton

This paper defends Christopher Boorse's biostatistical account of health and disease (1977, 2014) against influential objections from Rachel Cooper (2002) and Elselijn Kingma (2007) regarding choice of reference classes. My focus is restricted to the reference class problem and sets to one side other objections to the biostatistical account. I argue that evolutionary theory yields wide, non-arbitrary classes for identifying pathology. These classes align closely with Boorse's, although I briefly examine potential problems regarding pregnancy and race.

Boorse defines normal functioning as the level of function statistically typical in a reference class of organisms. "The fundamental idea is that a pathological condition is a state of statistically species subnormal biological part-function (Boorse, 1997, 4), relative to sex and age." … "1. the reference class is a natural class of organisms of uniform functional design; specifically, an age group of a sex of a species. 2. A normal function of a part or process within members of the reference class is a statistically typical contribution by it to their individual survival [or] reproduction". (Boorse 2014: 684)

Kingma (2007) argues that there is no good way in which to make sense of what counts as a “natural” class of organisms. She concludes that lack of objectively “natural” reference classes implies that their choice in medicine is arbitrary or could even be circular. Cooper’s objection (2002) is that reference classes might, potentially, collapse to small populations or even single individuals and, hence, distinguishing dysfunction from accidental effects may not be possible.

My response to Cooper and Kingma expands upon and explores in detail Boorse’s recent comment that, “as for choice of reference class, the one that I suggested medicine uses— an age group of a sex of a species or subspecies—could hardly be a more biologically natural choice.” (Boorse 2014: 695) This brief suggestion requires clarification and further detail because part of Kingma’s critique rests on questioning how properly to define biological naturalness.

I argue that medicine can look to biological theory to deliver reference classes. Physiological theory would be one place to look but doing so risks narrowing reference classes to pathological conditions recognised by physiology, and this would make the biostatistical account vulnerable to Kingma’s objection. Instead, evolutionary biological theory includes at its core a number of reference classes which closely but not precisely match Boorse’s criteria. Hence my analysis provides qualified support for his biostatistical account. Use of evolutionary theory has the advantage that it is a high level, cross-taxa theory with no prior non-naturalistic commitments with regard to pathology. Core evolutionary reference classes are, firstly, the species category. Secondly, sexual selection theory is a core branch of theory across taxa (Servedio et al 2017), including humans and this relies on categorising organisms by sex. Thirdly, life history theory models categorise organisms according to mature versus immature organisms, acknowledging different behavioural and morphological strategies across a lifetime including, in humans, the existence of senescence (Mace 2000). Here there is a point of difference with Boorse because, while life history theory models are typically stratified according to age, such stratification is a proxy for the genuine biological classifications, which are the species’ various life stages. Hence an evolutionary approach to human classes would categorise homo sapiens as infants, pre-pubescent juveniles, adults and senescent adults. Thus, Kingma’s worry that reference class choice cannot be non-arbitrary is answered. This argument also eliminates most of the reference class narrowing features raised by Cooper. At this point I briefly consider implications of the evolutionary approach for the normality versus pathology of “diseases” associated with “ageing”.

Finally, I briefly consider one or two problematic issues regarding this evolutionary approach to medical reference classes. First, pregnancy is an interesting difficult case for biostatistical disease ascription, given the significant but (sometimes) temporary physiological changes that occur. Evolutionary theory, compared to physiology or developmental human biology, does not recognise pregnancy as a reference class (inasmuch as it does not feature in standard life history models). Under an evolutionary approach, various changes during pregnancy, which would normally be considered “normal for pregnancy” will potentially be regarded as pathological relative to the mature adult female population (Formelli et al 2016). Secondly, race is a class which may be problematic. Cooper raises race as an additional narrowing to Boorse’s criteria and worries that it will lead to overly narrow reference classes. Given the consensus that race is a suspect evolutionary classification in humans (Hochman 2016), Cooper’s concern appears unfounded. However, if we switch our attention from “race” to sub populations more generally, there is good evidence that some quite small populations have been subject to strong, non-typical evolutionary pressures and have developed non-typical physiological adaptations as a result (Ilardo et al 2018, Ilardo & Nielsen 2018). Hence, on a case by case basis, there may be small, evolutionarily significant human populations for which statistical normality does differ significantly from other homo sapiens.

I conclude that application of evolutionary theory to the reference class objection broadly vindicate this aspect of Boorse's biostatistical account of pathology, although such application also draws out some interesting potential problems and complexities regarding statistical normality.

Boorse, 1977. Health as a theoretical concept. Phil Sci,44(4),pp.542-573.

1997. A rebuttal on health. In What is disease?(pp.1-134)

2014. A second rebuttal on health. J Med and Phi,39(6),pp.683-724.

Cooper, 2002. Disease. Studs Hist Phil Sci C:33(2),pp.263-282.

Formelli et al,2016. "Transient osteoporosis and pathological fractures in pregnancy and puerperium: A case report and review of literature." Italian J. Gynaecology, Obstetrics 28,2:45-47.

Hochman, Adam. 2016 "Race: Deflate or pop?." Studs Hist Phil Sci:C:57:60-68.

Ilardo et al,2018 . Physiological and Genetic Adaptations to Diving in Sea Nomads. Cell, 173(3),pp.569-580.

Ilardo et al,2018. Human adaptation to extreme environmental conditions. Current ops. genetics & development,53,pp.77-82.

Kingma, 2007. What is it to be healthy?. Analysis, 67(294),pp.128-133.

Mace, 2000. Evolutionary ecology of human life history. Animal behaviour,59(1),pp.1-10.

Parikh et al 2018. Peripartum cardiomyopathy and preeclampsia: overlapping diseases of pregnancy. Current hypertension reports,20(8),p.69.

Servedio et al,2017. The role of sexual selection in local adaptation and speciation. Annual Rev Ecol, Evo, Syst,48,pp.85-1

Drawing the Semantics–pragmatics Distinction in Animal Communication

Mihnea Capraru
Nazarbayev University

In 1978 Richard Dawkins and John Krebs have argued that animal signaling is to be interpreted not as communicating information, but simply as influencing or manipulating the receivers' behavior [endnote 1]. This view has been influential and seems to be the motivation for eusocial insect experts who posit a separate meaning for every separate behavioral influence. E. O. Wilson and Bernd Hölldobler maintain, for instance, that the alarm pheromones of ants are polysemous because they engender different reactions in their receivers (2009). When an ant close to the nest receives the alarm signal, it fights to defend the nest, while an ant far from the nest is more likely to flee and return home. Similarly, older ants are more likely to fight while younger ones are more likely to flee. Wilson and Hölldobler conclude that the pheromone carries two meanings, 'fight!' and 'flee!', which are conveyed to different receivers in different contexts.

This conclusion, however, seems to violate Occam's Razor by multiplying meanings beyond necessity. Instead of positing a separate meaning for every observed behavior, it would be preferable if we could explain the behavior as the resultant of two factors: 1) a constant meaning and 2) a context-dependent behavioral repertoire that would need to be posited anyway, on independent grounds, even if the danger signal did not exist. This strategy is inspired by Paul Grice's approach to conversational implicature (1975), with the difference that Grice averts to human reasoning abilities, whereas in the case of ants we must appeal to instinctive behavior. In David Kaplan's terms (1989), the alarm pheromone means danger and encodes the semantic content that there is danger at the specific place where the pheromone is released [endnote 2]. The pheromone is best seen as transmitting information about the presence of danger at a place, and not as eliciting a specific behavioral response. As for the variable responses that ensue, we can explain them as the result of the following factors:

1. The perception of danger. In our case this perception is produced by the alarm pheromone, but in general it can be produced in any number of ways (e. g., by being under direct attack or by sensing the chemical signature of a rival colony) [endnote 3].

2. Territorial behavior. Territorial animals have the propensity to fight when close to home but to flee when far away. This is so regardless whether they are social or not, or whether they receive alarm signals or not.

3. The social structure of ant colonies. As is well known, older ants perform jobs far from the colony center, while younger ants work closer to the queen. Defense outside the nest is a relatively remote job, so it is performed by older ants. When there is a soldier caste, defense is performed primarily by the soldiers.

The chemical signal, thus, is best understood as having a constant, informational meaning, as well as a moderately context-sensitive, still informational semantic content; meanwhile, the signal's variable behavioral influence is explained by the ants' generic behavioral dispositions in the presence of danger. This means that ant signaling supports a distinction analogous to the semantics/pragmatics distinction we draw for human languages, both on the ‘near side’ (the determination of semantic content based on constant meaning and variable context), and on the ‘far side’ (the context-dependent influence of semantic content on the receivers' behavior).

To be sure, when we say that ant signals have informational meaning, we are not saying that the ants possess any form of substantial mental understanding of these signals. The reason why danger signals transmit information about danger is not because the ants interpret them as such, but simply because the ants react to the signals in ways which are likely to increase their fitness in the presence of danger [endnote 4]. (I adopt this minimalist understanding of semantic information from Ruth Millikan (1984).) The fitness-enhancing behavior at issue varies with the location, age, and caste of the receiver; however, the fitness-enhancing information is the same. This is why the most economical theory identifies the signals' content not with the elicited behavior, which results ‘for free’ from the interplay between information and the preexisting behavioral dispositions, but with the information itself.

This economical approach benefits us not only through increased explanatory power, but also by reducing the empirical commitments of the theory, and thus by increasing its epistemic probability. If the alarm pheromone had different meanings, then we would expect each of these meanings to be underlain by a separate evolutionary explanation: e. g., we would need to posit one adaptation for responding to the alarm pheromone far from the nest, and another adaptation for responding close to home. But if the pheromone has only one meaning, then we only need to posit one adaptation, i. e., one that connects sensing the pheromone to perceiving danger.

=== Endnotes [1] Note, however, that Krebs and Dawkins moderated their view in 1984. The influence-only view received recent support from Rendall, Owren, and Ryan (2009). The collection edited by Ulrich Stegmann in 2013 pursues the influence–information debate at length.

[2] This means that the place of emission carries semantic information about itself; in Ruth Millikan's terms, the place is a reflexive sign (2004).

[3] This approach is inspired by J. J. Gibson's views of perception laid out in 1968 and 2015[1986].

[4] This is not to say that ants do not possess any internal representations at all; Charles Gallistel has argued convincingly to the contrary (1990; 2010). But such internal representations appear to be limited, and quite plausibly they are not involved in the forms of communication under discussion.

Evidence in Cancer Epidemiology at IARC

Michael Wilde
University of Kent

A main aim of the Monographs programme of the International Agency for Research on Cancer is to evaluate the strength of the available evidence concerning whether a particular exposure is carcinogenic to humans. In some cases, the evidence from epidemiological studies alone is taken to be sufficient to establish carcinogenicity in humans (Bouvard et al 2015). This practice has received some criticism from philosophers of science. In particular, Bert Leuridan and Erik Weber (2011) have objected that there is never sufficient epidemiological evidence for carcinogenicity because this evidence comes only from observational studies, the design of which does not sufficiently rule out the possibility of alternatives to the causal hypothesis, for example, confounding, bias, or chance. Against this, I argue that in the evaluation of the consumption of processed meat as a cause of colorectal cancer, the epidemiological evidence alone was in fact sufficient to establish carcinogenicity. The argument appeals to the viewpoints for causal inference in epidemiology provided by Austin Bradford Hill (1965). In particular, causation could be inferred because the consistency of the epidemiological data together with a clear dose-response curve helped to rule out confounding, bias, and chance.

I then consider two ways of modelling such causal inferences in cancer epidemiology. On the one hand, a probabilistic model in terms of inference to the best explanation seems to give a more faithful description of the fallibility of causal inference in epidemiology (Lipton 2004). However, I argue that this comes at the cost of understanding the causal hypothesis as properly established. Although the causal hypothesis provides the better explanation of the epidemiological data compared to confounding, bias, and chance, the causal hypothesis is not established because it does not provide the only explanation of the data. On the other hand, a non-probabilistic model in terms of inference to the only explanation does not share this cost (Bird 2005, 2007, 2010, 2011). However, I argue that this model gives a less faithful description of the fallibility of causal inference in epidemiology. Intuitively, the causal hypothesis is not the only hypothesis consistent with the epidemiological data, given the possibility of confounding, bias, or chance. I argue in favour of a model that combines the probabilistic and non-probabilistic models (Williamson 2000). It is this model that provides the correct way to understand the causal inference behind the conclusion that the consumption of processed meat is a cause of colorectal cancer. I intend this model to provide a qualified defence of the practice of the Monographs programme of the International Agency for Research on Cancer.

Explanatory Pragmatism as a Philosophy for the Science of Explainable Artificial Intelligence

Rune Nyrup
University of Cambridge

A common objection to deploying AI systems in ethically sensitive domains is that it can be difficult to adequately explain their decision-making to humans. Many current forms of AI, especially those based on advanced machine learning techniques, are often accused of being “opaque”, “black boxes”, “uninterpretable” or “incomprehensible”. In response, a new sub-field is currently emerging within AI research, aiming to create methods for making ‘interpretable’ or ‘explainable’ AI, sometimes abbreviated XAI.

Early work in this field tended to rely on researchers’ intuitive sense of whether a given model or system was more intelligible. Recently, however, a number of researchers have grown dissatisfied with this approach and started calling for more ‘rigorous’ or ‘scientific’ approaches to XAI.

Two main such approaches are currently being pursued: (1) Empirical approaches, which seeks to devise experimental methods for measuring whether a system is explainable, e.g. by (a) measuring how representative users evaluate the adequacy of explanations; or (b) testing their performance on some domain-relevant task. (2) Theoretical approaches which seek to design AI systems based on some existing account of explanation from psychology or philosophy.

Both approaches represent plausible steps forward, but also face limitations in their current forms. Regarding (1a), there is evidence that people sometimes overestimate how much understanding they get from an explanation; (1b) gets around this problem by focusing on behavioural measures but faces the question of which tasks are most relevant to determine understanding. Regarding (2), given the fact of explanatory pluralism, i.e. that there are many different explanatory models, there is unlikely to be a single account of explanation which can form the basis for XAI. This suggest a more contextual approach, but the field is currently lacking a principled method for choosing which model of explanation to implement in a given application of AI.

I propose a package of philosophical views, which I call Explanatory Pragmatism, as a promising guiding framework for the field of XAI research. Explanatory Pragmatism combines: A communicative view of explanation, i.e. the view that explanations are in the first place speech-acts aiming to generate understanding; a manipulationist view of understanding, i.e. degrees of understanding consist in the ability to successfully perform certain tasks; and a contextualist semantics for understanding, i.e. ascriptions of understanding are true if the subject possesses certain degrees of understanding deemed contextually important.

This provides a promising framework on two grounds: First, it fits naturally with the move towards a contextual and practice-based conceptions of explanation already underway in the field of XAI research. Second, it suggests a constructive approach to overcoming the limitations highlighted previously. Rather than trying to device general theories or measures for AI explanation, the field should aim to formulate plausible mid-level theories of what constitutes adequate explanations in specific contexts. Explanatory Pragmatism suggests that this should be done by focusing on: (a) what kinds of tasks should be deemed important in a given domain for the audience of explanations to be able to perform; (b) what information do they need in order to successfully perform these tasks; and (c) what kinds of explanations will best supply that information.

Fields, Loops, and the Strong Cp Problem

John Dougherty
Munich Center for Mathematical Philosophy

I argue that the holonomy approach to gauge theories does not solve the Strong CP problem. The Strong CP problem is one of the most prominent theoretical defects of the standard model of particle physics, and is recently a topic of much experimental interest. It arises because the observed CP symmetry of the strong force apparently requires fine-tuning of the standard model Lagrangian, which comes about because of the vacuum structure of quantum chromodynamics (QCD). The Strong CP problem is also important in the philosophical literature, where it has been used to criticize popular philosophical accounts of symmetry and to argue in favor of some interpretations of gauge theories like QCD. In particular, Healey (2007, 2010) has claimed, following Fort & Gambini (2000), that interpreting gauge theories in terms of properties attaching to loops in spacetime---rather than the more common field-theoretic interpretation---dissolves the Strong CP problem by simplifying the vacuum structure of QCD. I show that this is incorrect, and that the holonomy interpretation has no special resources for answering the Strong CP problem.

The first part of this talk reviews the origins of the Strong CP problem, with a particular emphasis on the role played by the distinction between gauge and physical symmetries. In order to solve another problem with the standard model, the axial U(1) problem, 't Hooft argued that some gauge transformations are physical symmetries---i.e., they relate distinct physical possibilities ('t Hooft, 1976). This contradicts the common philosophical view that gauge symmetry is "surplus structure", as Belot has argued (2018). These distinct but gauge-related configurations are also distinct classical vacua of the theory, and so the quantum vacuum in QCD is a superposition of these configurations, parametrized by an angle θ appearing in the QCD Lagrangian. If θ is nonzero then the strong force violates CP symmetry, which has not been observed. So we face a fine-tuning problem: why is θ exactly zero?

The second part of the talk considers 't Hooft's argument that some gauge-related field configurations are distinct physical states of affairs. I argue that this is a generic feature of gauge theories. Following recent work that characterizes gauge theories in category-theoretic terms (Dougherty 2017, Nguyen et al. 2018), I show how gauge structure can lead to distinct but gauge-related configurations in any gauge theory, if one restricts attention to appropriate sectors of the theory. In fact, this phenomenon is a consequence of the view that gauge-related configurations in the full theory represent the same physical state of affairs. The category-theoretic characterization of gauge theories makes no distinction between field-theoretic theories and other theories, and so 't Hooft's argument makes no use of the difference, either.

The final part of the talk shows that the Strong CP problem arises for holonomy formulations of QCD just as much as for field-theoretic versions. In the holonomy formulation defended by Healey (2007, 2010) and Fort & Gambini (2000), gauge-related configurations always represent the same physical state of affairs, both in the full theory and in any particular sector of the theory. As such, there is only one classical vacuum in this formulation, the vacuum of QCD is not a superposition, and the parameter θ never arises. But this is just because their formulation does not solve the axial U(1) problem. Solving this problem requires more attention to the gauge structure of holonomy formulations, an overlooked issue since the development of these formulations by Wu & Yang (1975). The vacuum structure of QCD depends on the gauge structure of the theory, not on its specific formulation, and so appealing to holonomy formulations gives no new tools for resolving the Strong CP problem.

References

Belot, G. Fifty million Elvis fans can't be wrong. Noûs, 52:946–981, 2018.

Dougherty, J. Sameness and separability in gauge theories. Philosophy of Science, 84(5), 2017.

Fort, H. and Gambini, R. U(1) puzzle and the strong CP problem from a holonomy perspective. International Journal of Theoretical Physics, 39:341–349, 2000.

Healey, R. Gauging What's Real. Oxford University Press, 2007.

Healey, R. Gauge symmetry and the theta-vacuum. In M. Suárez, M. Dorato, and M. Rédei, editors, EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association, pages 105–116. Springer, 2010.

Nguyen, J., Teh, N. J., and Wells, L. Why surplus structure is not superfluous. The British Journal for the Philosophy of Science, page axy026, 2018.

't Hooft, G. Computation of the quantum effects due to a four-dimensional pseudoparticle. Physical Review D, 14:3432–3450, 1976.

Wu, T. T. and Yang, C. N. Concept of nonintegrable phase factors and global formulation of gauge fields. Physical Review D, 12:3845–3857, 1975.

Gauge and Boundary: A Complicated Relationship

Henrique Gomes
University of Cambridge

Forces such as electromagnetism and gravity reach across the Universe - they are the long-ranged forces. And yet, in many applications of physics, we only have access to finite domains of the world. For instance, in computations of entanglement entropy, e.g. for black holes or cosmic horizons, we put boundaries in the world, separating the known from the unknown. Similarly, we would like to assign to a given bounded region all the total charges - sources of the given force-contained therein.

But, in this talk, I will argue that we do not know precisely how to do this. I will argue we do not understand gauge theory as well as we think we do, when boundaries are present. Such misunderstandings have been rekindled in the recent debate surrounding the meaning of certain charges at asymptotic infinity - the so-called 'soft-charges' (see [Str18] and references therein).

Indeed, local gauge theories are in a complicated relationship with boundaries. It is agreed by all that we should aim to construct variables that have a one to one relationship to the theory's physical content within bounded regions-regional observables. But puzzles arise if we try to combine definitions of strictly physical variables in different parts of the world.

This is most clearly gleaned by employing the simplest tool for obtaining unique physical representatives-gauge-fixings-and finding its shortcomings. Whereas fixing the gauge can often shave off unwanted redundancies, the coupling of different bounded regions requires the use of gauge-variant elements. Therefore, the coupling of regional observables is inimical to gauge-fixing, as usually understood. This resistance to gauge-fixing has led some to declare the coupling of subsystems to be the raison d'être of gauge [Rov14].

Indeed, while gauge-fixing is entirely unproblematic for a single region without boundary, for finite bounded regions it introduces arbitrary restrictions on the gauge degrees of freedom themselves. Such arbitrary boundary choices enter the calculation of charges through Noether's second theorem, barring the assignment of physical charges to local gauge symmetries. The confusion brewn by gauge at boundaries is well-known, and must be contended with both conceptually and technically.

It may seem natural to replace the arbitrary boundary choice with new degrees of freedom, for using such a device we resolve some of these confusions while leaving no naive gauge-dependence on the computation of Noether charges [DF16]. This resolution has recently become popular, but, concretely, such boundary degrees of freedom are rather arbitrary - they have no relation to the original field-content of the field theory. How should we conceive of them?

Here I will explicate the problems mentioned above and illustrate a difierent possible resolution. The resolution was introduced in a recent series of papers [GR17,GR18,GHR18]. It requires the notion of a connection-form in the field-space of gauge theories. Using this tool, a modified version of symplectic geometry - here called 'horizontal' - is possible. Independently of boundary conditions, this formalism bestows to each region a physically salient, relational notion of charge: the horizontal Noether charge. It is relational in the sense that it only uses the different fields already at play and relationships between them; no new "edge-mode" degrees of freedom are required.

The guiding requirement for the construction of the relational connection-form is simply a harmonious melding of regional and global observables. I show that the ensuing notions of regional relationalism are different from other attempts at resolving the problem posed by gauge symmetries for bounded regions. The distinguishing criterion is what I consider to be the 'acid test' of local gauge theories in bounded regions: does the theory license only those regional charges which depend solely on the original field content? In a satisfactory theory, the answer should be "yes". Lastly, I will introduce explicit examples of relational connection-forms, and show that the ensuing horizontal symplectic geometry passes this 'acid test'.

References

[DF16] William Donnelly and Laurent Freidel. Local subsystems in gauge theory and gravity. JHEP, 09:102, 2016.

[GHR18] Henrique Gomes, Florian Hopfmuller, and Aldo Riello. A unified geometric framework for boundary charges and dressings: non-Abelian theory and matter. NPB, 2019.

[GR17] Henrique Gomes and Aldo Riello. The observer's ghost: notes on a field space connection. JHEP, 05:017, 2017.

[GR18] Henrique Gomes and Aldo Riello. Unified geometric framework for boundary charges and particle dressings. Phys. Rev. D, 98:025013, Jul 2018.

[Rov14] Carlo Rovelli. Why Gauge? Found. Phys., 44(1):91{104, 2014.

[Str18] Andrew Strominger. Lectures on the infrared structure of gravity and gauge theory. Princeton University Press, 2018.

Gibbs' Solution of Gibbs' Paradox

James Wills
London School of Economics

Gibbs' paradox, in its broadest form, is a puzzle in the foundations of thermal physics concerning the difference between the entropy change on mixing two distinguishable samples of gas and the entropy change on mixing two indistinguishable samples of gas. This essay is about Gibbs' paradox in thermodynamics.

There is also a paradox in statistical mechanics which concerns justifying a factor of N! (where N is the number of particles in the gas) in the expression for the entropy in order to make it extensive (proportional to N). This paradox has attracted and continues to attract more attention in the philosophy and physics literature than the thermodynamic version. The consensus is that the thermodynamic paradox has been satisfactorily solved.

This paper pushes against this consensus through a combination of philosophical analysis of the mathematical and physical foundations of thermodynamics and historical interpretation and reconstruction of Gibbs' 1875-1878 argument, thereby analysing the paradox in thermodynamics in more detail than has so far been done. The first aim of the paper is to point out that there are three distinct versions of the paradox in thermodynamics in the literature. This is important because all discussions and solutions so far are targeted at only one of the versions and thus no treatment to date can claim to provide the full picture. Furthermore, I show that they all follow from two premises. The second aim is to argue that the paradoxes disappear when we derive from first principles the equation used to calculate entropy changes. The third aim is to show that this analysis reconstructs and clarifies Gibbs' reasoning concerning gas mixing.

Gibbs considers the ''increase in entropy which takes place when two different gases are mixed by diffusion, at a constant temperature and pressure.'' Assuming equal volumes of the (ideal) gases are mixed by removing a partition, the process described by Gibbs is two separate processes occurring simultaneously, one for each sample of gas, called Joule expansions. Gibbs calculates the entropy of mixing to be 2nRln2, where n is the number of moles of each sample of gas and R is the molar gas constant, whereas the entropy of mixing when the gases are indistinguishable is zero. There are three puzzles associated with this result:

GP1: The entropy of mixing is independent of the kinds of gas.

GP2: The entropy of mixing changes discontinuously as the gases go from distinguishable to indistinguishable.

GP3: There is a non-zero increase in entropy on mixing indistinguishable gases and no increase in entropy on mixing indistinguishable gases.

The reason GP1 and GP2 are paradoxes is that they contradict intuitions. GP1 contradicts an intuition we may have that the entropy of mixing should depend on the type of gas. GP2 contradicts an intuition we may have that the entropy of mixing should vary continuously as the gases become more similar. GP3 is a straightforward logical contradiction.

In order to look for resolutions of the paradoxes and see where they come from, I provide a reconstruction of the reasoning in the literature which leads to the paradoxes. No source gives the full line of reasoning I provide because each source is concerned only with one of the paradoxes. I show that the arguments for the paradoxes are based on two premises:

P1: The entropy increase for one gas undergoing a Joule expansion is nRln2.

P2: There is no increase in entropy on mixing indistinguishable gases.

Stating the reasoning explicitly is very useful for two reasons. Firstly, it helps us see what it takes to block the paradoxes. For GP1 and GP2, we either have to explain why our intuition is wrong or block the derivation of GP1 or GP2 by rejecting one or more of the premises from which they follow. Secondly, the derivation makes it possible to match each attempted solution in the literature to the rejection of an intuition or premise. This will help us see clearly how each solution works and allows us to assess them against a common logical background. This adds clarity to the debate and helps us see what is at stake.

A curious feature of these puzzles is that Gibbs himself was only puzzled by GP1, not GP2 or GP3, whereas physicists and philosophers writing about the paradox since seem to have been puzzled by GP2 and GP3 and not GP1. My explanation for this is that Gibbs understood the gas mixing scenario he described perfectly well, although his argument and explanations are not all that clear and sometimes difficult to follow.

I reconstruct and clarify Gibbs' reasoning by arguing that P1 requires the proviso that there be a reversible process between the initial and final states. This proviso then carries through the argument, blocking GP3 and demonstrating that GP1, GP2 and P2 are all true. I argue for this proviso by showing that it is required by the theory of thermodynamics; it comes out explicitly when deriving from first principles the equation used to calculate the entropy change.

Further, I argue that this proviso leads to a clear and precise definition of distinguishability in thermodynamics. This result is important in three respects. Firstly, this definition is not stipulated or 'ad hoc' in any sense; it is shown to arise from the physics and formalism of thermodynamics. Secondly, the confusion in the literature on the Gibbs paradox can plausibly be traced to the lack of such a definition. Thirdly, it shows that there is a very particular sense in which distinguishability is meant in thermodynamics, indicating that distinguishability in other contexts and theories (such as statistical mechanics) could plausibly have very different meanings and definitions. This is a warning against conflating definitions when discussing distinguishability. Fourthly, I show that Gibbs can and should be interpreted as giving an identical definition in his original analysis of gas mixing.

This paper reconstructs Gibbs' result and reasoning and so the resolution of the paradox in thermodynamics presented here is grounded in the authority of Gibbs and the physics and formalism of thermodynamics.

Going It Alone (epistemically)

Katherine Furman
University College Cork

Public distrust in science is on the rise. With it, so too is the temptation to 'go it alone' epistemically. That is, to disregard mainstream science and engage in independent evidence gathering in an effort to become your own expert on the topics that matter to you. There are plenty of examples of going it alone with varying degrees of success.

On the negative end of the spectrum, former South African president Thabo Mbeki's AIDS policies in the late 1990s and early 2000s provide a cautionary tale against going it alone. Mbeki distrusted the mainstream scientific view, suspecting that the scientists were racist. He went on to engage in large scale independent evidence gathering; first online and later consulting fringe scientists. He came to the conclusion that HIV does not cause AIDS, and that anti-retrovirals (ARVs) – the treatment that prevents the virus from replicating – is toxic. He thus prevented the distribution of ARVs via the public health system and best estimates indicate that this resulted in 171, 000 avoidable new infections and 343,000 deaths over the 1999–2002 period.

On the other end of the spectrum, Steven Epstein's (1996) 'Impure Science' details how gay activist groups in the United States in the 1980s resisted the mainstream scientific view on AIDS, ultimately becoming the experts themselves and playing a fundamental role in developing the testing and treatment protocols for early AIDS treatment regimens. This is case of a group going it alone very successfully.

Given the diverse range of possible outcomes associated with going it alone epistemically; with severe harm on the Mbeki end of the spectrum and heroism on the early AIDS activist end, how might you decide when it is a good idea for you to try to forge your own epistemic path? This talk will provide some guidance on the factors you should consider when deciding whether to go it alone epistemically.

I suggest that there are at least four factors to consider when deciding whether to go it alone. These are: 1) what is your pre-existing knowledge base like; 2) what is your available time frame; 3) what are the stakes of getting the answer wrong; and 4) is the issue of resistance one of facts or values? More needs to be said on each.

The first thing you might consider when deciding whether or not to go it alone is your pre-existing knowledge base. Goldman (2001) notes that there is no one unified group of laypeople – rather, on any particular issue, individuals will have a wide range of pre-existing expertise, from absolute novices to almost-experts. If the epistemic area you are considering is close to your pre-existing skill set – perhaps you are a biologist considering going it alone on a medical matter – then it would be more permissible for you to do so than for the complete novice. The Dunning-Kruger effect – the cognitive glitch by which we over-estimate our own abilities – might interfere with our ability to do this self-assessment well, but it would still be a good idea to reflect on your own skills before deciding to embark on independent evidence gathering.

A second factor to consider is your time frame. Can you acquire the relevant skills in the time available, or will the problem no longer be relevant? Empirically, most cases in which you might want to go it alone are time sensitive (typically health-related cases), and it may not be possible to develop the required skills quickly enough for them to be relevant. Further, like considerations about your pre-existing skills set, you might not be well-placed to make your own assessments about how long the task will take; perhaps we systematically under-estimate the time required to develop new skills.

Third, you should consider what the stakes are of getting the wrong answer. Going it alone will be a epistemically riskier than following the mainstream scientific view, given the training and checks that mainstream science involves. If the foreseeable harm to others is very high (like it was in the Mbeki case) this will speak in favour of deferring to the experts. However, if the expected outcome of following the mainstream view is already very bad, then it will be more permissible to go it alone. We already accept that it is permissible to suspend standard decision procedures when the expected outcome is dire. Consider the case of the recent West African Ebola outbreak, during which there were only 3 doses of the very experimental treatment, ZMAPP. ZMAPP had not even gone through animals trials at the time of the outbreak, but given how deadly Ebola is and the dearth of available treatments, it was decided that they should be distributed anyway. When there is nothing left to lose, it is permissible to go it alone.

Finally, you should consider whether the point of contention is one of facts or values. Much of the resistance between early AIDS activists and mainstream AIDS scientists in the 1980s were disputes over values. For instance, activists insisted on much shorter timeframes for pharmaceutical trials. The activists recognised that this would lower safety standards, but this was preferable to no treatments at all. Resistance on values is likely to be more permissible than resistance on facts.

Overall, I argue that there is no uniform advice that can be given about when it is appropriate to go it alone epistemically. Rather, the agent contemplating this course of action should consider a range of factors; including their own skills, the type of problem at hand, the nature of the dispute, and the stakes of getting the answer wrong. Further, doing first-person assessments of these issues is difficult, given our cognitive biases and should be approached with caution.

Heaviside's Operational Calculus and the Application of Unrigorous Mathematics

Colin McCullough-Benner
University of Leeds

It is common in science, especially in physics, to apply mathematics that does not meet the standards of rigor of pure mathematics. In this paper, I argue that such applications present the standard account of the applicability of mathematics, the mapping account, with serious problems. Using Heaviside's applications of his operational calculus as a case study, I argue that even the most plausible versions of the mapping account provide at best a misleading picture of central features of the practice of applying unrigorous mathematics, particularly the inferential restrictions characteristic of such applications.

According to the mapping account, mathematical scientific representations represent their target systems as bearing a structural similarity to a structure picked out by the mathematics, with this similarity cashed out in terms of a structure-preserving mapping between the structure of the target system and the relevant mathematical structure. Scientists' mathematically mediated inferences are then justified by the existence of such mappings.

I argue that such accounts give at best a misleading picture of what goes on in applications of unrigorous mathematics, obscuring central features of the practice of applying such mathematics. In particular, applications of unrigorous mathematics typically involve what Davey (2003) calls an “inferentially restrictive methodology”: limits are placed on when and how mathematically problematic concepts may be used, so that such concepts are quarantined to contexts in which they behave in the desired way. For example, the Dirac delta function is not a mathematically well-defined function, but it can be profitably applied as long as it only appears as a factor within an integrand. Heaviside’s application of his operational calculus, I argue, is a less tidy example of the same phenomenon.

The central idea behind Heaviside's operational calculus (like other operational calculi) is that, by treating differentiation and integration as operators, we can reduce difficult problems involving differential equations to simpler algebraic problems—in practice, essentially by moving from a representation of a physical system in terms of differential equations to an algebraic representation taken to be in some sense equivalent to it. A typical application of the operational calculus goes as follows: first, formulate differential equations characterizing the target system; second, replace each time derivative d/dt with the operator p and solve, treating p (and similar operators) as ordinary algebraic quantities; third, “algebrize” the solution by turning it into a function of t, typically by expanding the solution as a power series to put it in a form that allowed each occurrence of p to be eliminated (interpreting p^n as the nth time derivative and 1/p^n as n definite integrals from 0 to t).

These techniques failed to live up to the standards of mathematical rigor—in the hands of Heaviside, at any rate—for several reasons. For instance, Heaviside treated his p and 1/p operators as if they were inverse operators, though this is not generally the case if they are interpreted as above. He frequently manipulated his resistance operators as if they were commutative, though this is again not generally the case. And he worked with divergent series with reckless abandon. Heaviside never attempted to articulate in any kind of generality the conditions in which these techniques were valid. In fact, he thought doing so was counterproductive (see, e.g., Electromagnetic Theory, vol. 2, section 282).

In this sense, we can understand Heaviside as adopting an inferentially restrictive methodology. Rather than define the mathematical concepts he used in a fully rigorous way, so that the conditions under which his techniques apply were clear from the outset, he determined whether he could fruitfully apply particular techniques to particular cases by checking whether they yielded the right results, made good physical sense, and so on. Heaviside restricted the use of his techniques and the concepts they involved to contexts in which they worked—i.e., where these conditions were met. So, unlike in the cases described by Davey (2003), the restrictions are messy and observed only once one has seen that a mathematical technique or concept leads to a problematic result in a particular case.

We can represent this in terms of the mapping account in a minimal sense: we can find a structure and mapping that give the right accuracy conditions for each representation using the operational calculus. This can’t be the simple matter of choosing a structure picked out by a later, rigorized version of the operational calculus like Bromwich’s reconstruction in terms of integral transformations. Heaviside never did anything like any of the (numerous) strategies proposed to put his operational calculus on a rigorous footing. In fact, Heaviside was critical of these strategies, writing to Bromwich, for example, “I never could stomach your complex integral method” (quoted in Nahin, 2002, p. 230). Rather, an appropriate structure is any extension of a structure for real or complex analysis (depending on the application) that contains structure interpreting Heaviside’s differential and integral operators in which the inferences Heaviside makes use of in applying his operational calculus come out as truth-preserving. However, this leads to two problems.

First, we don’t really have a grip on such a structure independently of the inferential moves Heaviside took to be licensed by his operational calculus. As a result, it is misleading to claim that those inferences are justified (in any sense) by the positing of an appropriate mapping between such a structure and a target structure. Such a structure just is one that supports the relevant inferences. This problem, I argue, generalizes beyond the Heaviside case to any application of mathematics requiring an inferentially restrictive methodology.

Second, mapping accounts lack the resources to accommodate the messiness of Heaviside’s inferential restrictions. In particular, Heaviside frequently appeals to the physical interpretation of the mathematics in solving particular problems, using such considerations to help determine where particular techniques and concepts may legitimately be used, thereby mitigating the risks of using unrigorous mathematics. Such reasoning does not fit neatly in existing versions of the mapping account, which focus on inferences from the mathematics to the world rather than inferences from the world to the mathematics.

Indeterminism and the C Theory

Matt Farr
University of Cambridge

Summary.

Can a theory be indeterministic without implying a preferred direction of time? Intuitively it can, since the two concepts appear to be quite independent of one another. Despite this, a number of authors have made arguments to the effect that indeterminism entails the directionality of time. This paper critiques such types of argument and argues that: (1) such arguments establish that if there is a linear one-way law of temporal evolution, it cannot run in both directions; (2) that a time symmetrised understanding of transitional probabilities is perfectly coherent that fits with a temporally adirectional metaphysics. Given (1) and (2), the existence of indeterministic laws is insufficient to establish a time-directed metaphysics.

The one-way-applicability argument.

A number of authors have made structurally analogous arguments to the effect that if there are probabilistic laws of nature, they can only run in one direction in time. This kind of argument has been raised with respect to both thermodynamics/statistical mechanics and quantum measurement.

Watanabe (1965) puts forward a theorem which holds that “if we can use transitional probabilities as a statistical law in one direction [of time] then we cannot apply them in the opposite direction” (p. 168). Sober (1993) constructs a similar proof, taking it to establish that systems “cannot have both a forward-directed translationally invariant probabilistic law and a backward-directed translationally-invariant law” and that "science seems to favor the former” (p. 171). Call this feature of probabilistic laws their one-way applicability (OWA).

Penrose (1989) takes the OWA of quantum mechanics to make the case that quantum mechanics is predictive and not retrodictive (“[i]t is only for calculating the probabilities of future states on the basis of past states that this procedure works”) and consequently “cannot be time-symmetric” (Penrose, 1989, p. 359). Arntzenius (1995, 1997) argues that such cases entail that there is “an objective direction of time” (Arntzenius, 1995, p. 68) according to collapse interpretations of quantum mechanics. An analogous argument has also been made in statistical mechanics: for any typical non-equilibrium system, it evolves in such a way that its forwards evolution is towards more probable states and its backwards evolution is towards more improbable states. Maudlin (2007) argues that this is best explained by hypothesising that later states are ‘produced’ out of earlier states: “[t]his sort of explanation requires that there be a fact about which states produce which […] earlier states produce later ones […] [a]bsent such a direction, there is no account of which evolutions from states should be expected to be atypical and typical in which directions” (Maudlin 2007, p. 134). What we see is the OWA of probabilistic/statistical laws, in both quantum and statistical mechanics, being interpreted as supporting or requiring the hypothesis that time is directed, conversely implying that a temporally adirectional metaphysics cannot accomodate indeterministic theories.

Temporal Adirectionality and the C theory.

In order to assess the significance of such a line of reasoning, I distinguish between a time-directed metaphysics—the ‘B theory’—and a temporally-adirectional metaphysics—the ‘C theory’—, and outline and defend the commitments of the C theory. These two theories disagree as follows: the B theory holds that moments of time are ordered by an asymmetric temporal relation (‘earlier than’), with later states exhibiting an asymmetrical metaphysical dependence upon earlier states (e.g. earlier states produce later states and not vice versa); the C theory holds that moments of time are ordered by symmetric temporal relation (‘temporal betweenness’), meaning that no two worlds can differ solely over the directionality of time (which events are earlier than with other). The C theory’s adirectionality is instructive as to how to think of states of affairs standardly represented as directed in time, such as motions, velocities, etc. On the C theory, time reversing some description of a process results in an equivalent description of the same process (in line with Reichenbach’s [1956] passive interpretation of time reversal), meaning that our standard application of time-directed terms to systems (such as having some particular velocity) is dependent upon the convention of using our past-to-future direction to fix the description, rather than picking out some intrinsic time-directionality of the system itself.

The key question then is this: does OWA support the B theory over the C theory? I argue that OWA is insufficient to support the B theory over the C theory. What is established by the argument is that, if we are to assume the B theory—that is, if statistical laws are understood as governing evolutions from some time to another—it is clear that, OWA demonstrates that statistical laws favour one direction over the other. However, this argument specifically assumes the B theory and is insufficient to lend independent support for the central claims of the B theory over the C theory. On the contrary, I show not only that the C theorist can perfectly well accomodate OWA, but also that understanding statistical laws in terms of temporal betweenness relations offers a clear way of demonstrating that a theory can be both indeterministic and time symmetric. An instructive case in point is Aharonov, Bergman & Lebowitz's (1964) demonstration of how probabilities can be time symmetrised in such a way to remove the apparent time asymmetry of quantum measurement.

References.

Aharonov, Y., Bergmann, P. G., & Lebowitz, J. L. (1964). Time symmetry in the quantum process of measurement. Physical Review, 134(6B), B1410

Arntzenius, F. (1995). Indeterminism and the direction of time. Topoi, 14(1), 67-81

Arntzenius, F. (1997). Mirrors and the Direction of Time. Philosophy of Science, 64, S213-S222

Penrose, R. (1989). The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics. Oxford: Oxford University Press Press.

Reichenbach, H. (1956). The direction of time. Univ of California Press.

Sober, E. (1993). Temporally oriented laws. Synthese, 94(2), 171-189.

Watanabe, S. (1965). Conditional probability in physics. Progress of Theoretical Physics Supplement, 65, 135-160.

Inequivalent Representations and the Coalesced Structures Approach: Non-Radically Unpristine

Caspar Jacobs
University of Oxford

Ruetsche (2011) argues that the occurrence of unitarily inequivalent representations in quantum theories with an infinite number of degrees of freedom (QM∞) forces us to reject the ideal of pristine interpretation. Ruetsche puts the non-pristine Coalesced Structures Approach forward as an alternative. In this paper, I critically assess one of Ruetsche's cases in favour of the Coalesced Structures Approach: the broken symmetries of the ferromagnet. I agree with Ruetsche that a departure from the pristine ideal is indeed necessary to account for ferromagnetic phase structure. However, I disagree on the consequences of this departure. I argue that the version of the Coalesced Structures Approach necessary to account for the ferromagnet is non-radically unpristine: it leaves the core tenets of the ideal of pristine interpretation untouched.

The ideal of pristine interpretation requires interpretations are 'intrinsic' to a theory: they must follow from the theory itself and a handful of logical, metaphysical and methodological principles. By contrast, 'geographical' considerations, such as contingent facts about the world or context-dependent considerations about the use a theory is put to, are not allowed to play a role in pristine theory interpretation.

In this paper, I distinguish between two ways of understanding Ruetsche's alternative Coalesced Structures Approach (CSA):

Modest CSA: The interpretation of a theory may be influenced by contingent facts about the way the world is.

Radical CSA: The interpretation of a theory may be influenced by the context in which the theory is used.

I then argue that we only need the Modest CSA in order to account for the broken symmetries of the ferromagnet. This departure from the pristine ideal is modest in the following senses: firstly, it is continuous with scientific practice before the advent of QM∞; secondly, the it is compatible with scientific realism; and thirdly, it does not pose a threat to the distinction between laws and initial conditions. The Radical CSA, by contrast, would both threaten realism and radically blur the laws/initial conditions distinction.

The account of the ferromagnetic on the Modest CSA essentially follows Ruetsche (2006). I briefly repeat her arguments against 'Hilbert Space Conservatism' and 'Algebraic Imperialism', and take as a starting point Ruetsche's assertion that no pristine interpretation can account for ferromagnetic behaviour. We can explain this behaviour on the Modest CSA, as follows. Start with the set of algebraic states as our 'primordial' possibilities for the ferromagnet, and the set of algebraic observables as initially physically meaningful quantities. One of these states, ωa, is the state the particular ferromagnet we are considering is actually in. With the Gelfand-Naimark-Segal (GNS) representation of ωa, we obtain a particular Hilbert space containing observables parochial to that space. Among these observables is the polarisation observable m, which characterises phase behaviour. We can now 'coalesce' this observable into our theory, such that it becomes physically meaningful, and hence we can express whether the magnet is in its paramagnetic or ferromagnetic state. Furthermore, with this account of the ferromagnet we can understand how its rotational symmetry could have been broken in any direction: for each possible direction of polarisation, the algebra contains a corresponding state ωi such that, if the ferromagnet is in ωi, it is polarised in that direction. Thus, with our adulterated interpretation we can give a conspicuous account of spontaneous symmetry breaking.

Firstly, this mode of reasoning is continuous with scientific practice pre-QM∞. In classical mechanics, for example, the particular phase space we use depends on the number of degrees of freedom of the target system. We can imagine a 'Phase Space Conservative' who privileges a single phase space and declares that it exhausts the set of physical possibilities. This implies, implausibly, that a system necessarily has a certain number of degrees of freedom. In reality, scientists are opportunistic: they choose a phase space based on the actual, contingent features of the system. Similarly, the Past Hypothesis is an example of the Modest CSA in statistical mechanics: on the basis of the contingent fact that entropy increases with time, we narrow down the set of allowed states to those with a sufficiently low initial entropy. In both cases, then, scientists do not choose the full set of possible states of their theories a priori; rather, as the Modest CSA advocates, they pick the most suitable phase space based on contingent factors about their target system.

Secondly, Ruetsche (2011) argues that Pristinism sustains the No Miracles Argument (NMA) for realism. On the Radical CSA, there is no single interpretation of QM∞ to which all of its theoretical virtues belong. Instead, we need different interpretations in different circumstances. This poses a dilemma: either the interpretations are individually successful enough to merit the status of approximate truth; or they are not. In the first case, the NMA generates contradictory conclusions: it implies that each interpretation of QM∞ is correct, even though they may disagree on what is physically possible. On the second option, the NMA cannot be applied at all, which would severely limit scientific realism.

If, on the other hand, we accept the Modest CSA, these problems do not arise. For the Modest CSA still enables us to give a single interpretation of QM∞ to which all its explanatory benefits accrue. On the Modest CSA, interpretations are at most indexed to possible worlds; but within a world, each theory is given a unique interpretation. Finally, this also allows us to maintain a distinction between laws and initial conditions: all we need to concede is that laws are also dependent on contingent facts about the world (as is the case on Lewis' Best Systems Account).

In arguing against Ruetsche's claims regarding the implications of unitarily inequivalent representations in QM∞, I have limited myself to the issue of spontaneously broken symmetries, as exemplified by the ferromagnet. Ruetsche offers further cases to support her defence of unpristinism, such as her W*-argument and her arguments regarding Hadamard States. Unfortunately, I cannot address these arguments here; whether the Radical CSA is after all needed to account for these cases remains to be seen.

Is Physically Significant the Analogy Between Shannon's Information and Mechanical Statistical Entropies?

Javier Anta

In this talk, I will assess the conceptual relation between physical (especially its "mechanical statistical" version) entropy and Shannon's information. Although there exist stronger arguments for regarding the latter concept as not having physical content, as we will see later on, the formal analogy between these two quantities might be epistemologically relevant for scientific purposes.

The thermodynamic concept of "entropy" was coined by Clausius in 1865 in order to refer to the ratio between a minimal amount of energy dissipation and a certain temperature. That concept was translated by Boltzmann into mechanical terms with probabilistic tools (and microphysical assumptions), which properly refer to the probability for a physical system of having a particular microscopic configuration given certain macrostatistical measure. On the other hand, Shannon (et al. 1949), also called his information-theoretic quantity of uncertainty generated in the transmission of a message (sequence of symbols) as entropy "H". Based on the formal similarity between these two last quantities, Jaynes (1956) (who firstly considered them as "conceptually identical") developed a reduction of statistical mechanics into Shannon's information-theory terms. But, in which sense are the two entropies identical?

In this talk, I am going to analyses the main arguments within the scientific literature defending and attacking this "Jaynesian identification" (Wüthrich 2017). One of the main strategies is to justify the conceptual identity between Shannon's entropy "H" and mechanical statistical (particularly Gibb's) entropies "SG" from their mathematical or formal equivalence (1), since both have a logarithmic form and a probabilistic nature; this is what Steiner called "Pythagorean analogy" (2002):

H = - p In p - Formal Analogy -S = - p In p(1)

A second argument favoring the Jaynesian identification is the fact that both quantities (as we as their respective areas of inquiry) are deeply rooted in probability theory. This idea was mainly defended by neo-Jaynesians like Be-Naim (2008), who supported the eliminativistic strategy of substituting the unclear concept of "entropy" by that of "missing information", since the latter could be interpreted as an objective quantity. Other authors, particularly during the 90s, had gone even further in semantically identifying Shannon's information (or even Chaitin-Kolmogorov's complexity) with Clausius's thermodynamic entropy (Zurek 1990b, Peres 1990) following Wheeler's "it from bit" information-centrist thesis.

On the other side of the debate, there exist strong reason to regard the Jaynesian identification as unjustified. Firstly, Callender (1999) famously defended the conceptual disconnectedness not just between the quantities in (1), but moreover between two statistical-mechanical entropies: one dynamic-dependent and the other grounded on probability-functions (Boltzmann's and Gibb's, respectively). Although both Shannon's and statistical-mechanical entropy refers to "uncertainties", some authors (e.g. Denbigh 1982; Dougherty and Callender 2017) have argued that the probabilistic content of the former is intrinsically subjective since it reflects the epistemic state of an observer, namely the uncertainty generated in the transmission of a message; while the statistical-mechanical probabilities should be characterized as objective because they reflect the physical 'uncertainty' in determining the microstatistical state of a system. A thought experiment (firstly proposed by Wicken 1987) might help in illuminating this conceptual difference:

Let assume that a Laplace's demon gives us the microstatistical data determining the configuration of a particular gas. In this case, the value of Shannon's information become cero (there is no uncertainty about the gas) after this demonic action; while the value of the statistical mechanical entropy of the gas remain exactly the same, since its actual configuration does not change.

A second robust argument against the Jaynesian identification is that, the concept of statistical mechanical entropy is ultimately based on the microstate-macrostate (namely, the idea that a measurable quantity could be implemented in different microphysical configurations) distinction, while Shannon's entropy is not. I would argue, in the line of Wicken (1987, p.180), that messages (as possible chain of symbols) cannot even have an "entropic-like" isomorphic structure, since (i) symbols cannot be taken as microstates nor (ii) sequences could be viewed as macrostates: mere sequential combinations of symbols are not isomorphic to the statistical (or phase space) combination of positional and momentum-based data.

Grounded on the above conceptual assessment, I will propose an alternative way of evaluating whether the Shannon-Gibbs entropy analogy is physical relevant or not.

The core idea is to ground this empirically meaningful/meaningless conceptual connection onto its epistemic success; namely, (i) its explanatory power within a particular physical domain, or (ii) its predictive capacity in an experimental area. For this purpose, I will evaluate the role of the Jaynesian identification in the development of the concept of "black hole entropy" (codified via the area theorem or Bekenstein-Hawkin formula) within black hole thermodynamics. Particularly, Bekenstein (1973) originally defined black hole entropy as the measure of the inaccessibility of information about its internal configuration, explicitly relying on Shannon's information. Wüthrich (2017) argued that, although there is no direct experimental evidence, we have good theoretical reasons regarding the empirical validity of Bekenstein-Hawkin formula without its "Shannonian flavor", since it would immediately imply Platonism-Pythagoreanism. Against Wüthrich, I will defend (a) that we can reject the Platonist-Pythagorean conclusion derived from the epistemic power of Shannon's information in the empirical domain (Pincock 2012); and (b) the existence of positive experimental data confirming (although in an indirect fashion) Hawking radiation (see Dardashti et al. 2017), therefore showing the epistemic usefulness of relating the two quantities mentioned in our title.

I will conclude by claiming that making analogical inferences between Shannon's information and thermal entropies might be epidemically useful for certain scientific tasks (as we have just defended for the case of black hole thermodynamics), even when the former concept is regarded as having no physical meaning at all: or in other words, the analogy would be physically significant as long as it robustly contributes to our knowledge of the empirical world.

Jeffrey Conditionalisation: Proceed with Caution

Borut Trpin
University of Salzburg / University of Ljubljana

How should an agent update her degrees of belief when she is not fully certain of her evidence? A common prescription in Bayesian epistemology is that she needs to update by Jeffrey Conditionalization (JC), a generalisation of standard Bayesian conditionalisation for cases like this (see Jeffrey 1983, 164-83, for his explication). But why should an agent update by JC and not by some other rule? A common response is based on a proof that any agent who does not update by JC is vulnerable to a so-called dynamic Dutch book. In other words, a bookie who knows just as much as the agent can offer the agent a series of bets that the agent evaluates as fair but that lead to a guaranteed loss (Armendt 1980). The converse was also proven: any agent who updates by JC is invulnerable to dynamic Dutch books (Skyrms 1987).

The argument is convincing. A rational agent must avoid sure losses. However, as the problems described below show, invulnerability to Dutch books does not provide a be-all and end-all justification of JC. We will show that there exist many situations where JC gradually prescribes the agent to assign an arbitrarily high probability to a false hypothesis after observing specific sequences of uncertain but nonetheless non-misleading evidence. When we say that the evidence is non-misleading, we mean that when an agent becomes more certain of some evidence E than its negation ¬E, E is actually the case. Hence, while JC offers a pragmatic advantage (invulnerability to Dutch books), we believe that this advantage is offset by the epistemic disadvantage -- a rational agent ought, after all, not assign high probability to a false hypothesis (given that the evidence is not misleading). Although an agent cannot know whether some evidence is misleading or not in the described sense, a well-performing update rule should at least not lead to problematic outcomes in the latter cases. The problem is even more worrying because it is (at least in some outlined cases) robust with respect to the agent's prior probabilities. In other words, even if an agent who updates by JC is initially highly confident of the true hypothesis, there exist such sequences of non-misleading uncertain observations that she will eventually become highly confident of a false hypothesis.

Consider the following scenario for an illustration of how JC prescribes the agent to become highly confident of a false hypothesis: Freya is a Bayesian microbiologist. She updates her beliefs by Bayesian conditionalisation or by JC if she is not fully certain of her evidence and the rigidity condition is satisfied (The rigidity condition is satisfied when $Pr^*(H|E_k) = Pr(H|E_k)$ for all k (Jeffrey 1983, 174).) She has identified some bacteria in a sample and correctly believes it may only be of the A or B strain but not both. She knows that both strains have similar biochemical characteristics, except for characteristic E, which is 75% likely to be present in a given inspected part of strain A, and is present in all parts of samples containing strain B. It does not matter what her prior probability distribution is like. However, for the ease of calculations, suppose that her prior probabilities are 0.5 for both mutually exclusive and jointly exhaustive hypotheses. Further, suppose that her sample actually contains strain B, so that characteristic E is present in all inspected parts of the sample. Finally, suppose Freya inspects various parts of the sample 40 times and is constantly 70% certain that she observed characteristic E in each inspected part (e.g., because her instrument only affords her ineffable learning experiences).

It is easy (if a bit lengthy) to verify that after 40 such observations Freya becomes approximately 0.99 certain that her sample contains strain A (the one where characteristic E is 0.75 likely), and merely 0.01 certain that it contains strain B which she is actually inspecting. Considering that Freya's evidence was always such that she was reasonably certain that E was present in all inspected parts of her sample (she was constantly 0.7 certain about the presence of E), it is problematic that she assigned a very high probability to strain A and a very low probability to strain B hypothesis. What went wrong in this case was that the hypothesis with the (objective) likelihood of E closest to her (subjective) certainty of observing E was favoured. But this is not what we want from an updating rule -- we are not interested in confirming subjective certainties of evidence (at least when we are not fully certain). After all, Freya's observations perfectly fit strain B hypothesis as she was always more certain that E is present in the sample than that it is not.

The scenario is admittedly oversimplified to serve as an actual example from scientific practice. However, it affords a precise analysis of what principles of JC lead to the problem. We also discuss a number of variations of this problem (e.g., the cases where an agent operates with more hypotheses and different likelihoods) and show when and why the formal properties of JC lead an agent to assign very high probability to a false hypothesis despite non-misleading sequences of observations.

References

Armendt, B. (1980). Is there a Dutch book argument for probability kinematics? Philosophy of Science 47(4), 583-8.

Jeffrey, R. (1983). The Logic of Decision. Chicago and London: University of Chicago Press.

Skyrms, B. (1987). Dynamic coherence and probability kinematics. Philosophy of Science 54(1), 1-20.

Justifying the Norms of Inductive Inference

Olav Vassend
Nanyang Technological University

Bayesianism has become the most common formal framework used by philosophers of science to study scientific methodology, and it is also an influential framework for statistical inference. But it rests on an assumption that is often violated in scientific practice, namely that one of the hypotheses under consideration is true. Suppose none of the hypotheses under consideration is true, so that the goal is instead to find the hypothesis that is – in some sense – best. Depending on what is meant by ``best,'' the likelihood may not be an appropriate measure of evidential favoring. For example, suppose the goal is to identify the hypothesis whose expected maximal prediction error on future data is as low as possible. Then, as Vassend (Forthcoming) shows, the likelihood is not an appropriate measure of evidential favoring because the hypothesis that has the best likelihood score on the evidence will in general not be the hypothesis that has the lowest expected maximal prediction error on future data. In this context, a more reasonable measure of evidential favoring is one according to which the evidence favors H1 over H2 if and only if H1's maximal prediction error on the evidence is lower than H2's maximal prediction error on the evidence. The fact that Bayesianism is tied to using the likelihood as a measure of evidential favoring is therefore a limitation of the framework.

The purpose of my presentation is to study inductive inference in a very general setting where finding the truth is not necessarily the goal and where the measure of evidential favoring is not necessarily the likelihood. I use an accuracy argument to argue for probabilism and I develop a new kind of argument to argue for two general updating rules, both of which are reasonable in different contexts. One of the updating rules – which I call “inferential” updating – is characterized by the fact that it never assigns a posterior probability of 0 to a hypothesis that has a prior probability of 0. In other words, inferential updating obeys a kind of “regularity” principle. Inferential updating has standard Bayesian updating, Bissiri et al.’s (2016) general Bayesian updating, and Vassend’s (Forthcoming) quasi-Bayesian updating as special cases. The other updating rule – which I call “predictive updating” – is novel and is characterized by the fact that it violates regularity, but that it does so in a maximally conservative way.

My argumentative strategy is to divide inductive updating into two steps: in the first step, the prior plausibility of a hypothesis is combined with the hypothesis's score on the evidence according to some measure of evidential favoring in order to produce a posterior score. In the second step, the posterior scores are normalized so that they are probabilistic. The requirement that the combination step and normalization step commute in certain reasonable ways, together with a few other plausible assumptions, entail that the combination step and normalization step must both be either multiplicative or additive. That is, if the evidential score (according to some evidential measure, such as the likelihood) of some hypothesis is e and the prior plausibility of the hypothesis is h, then the posterior plausibility of the hypothesis must be proportional to either h*e or h+e. Adding the requirement that the updating rule obey the previously mentioned regularity principle is sufficient to entail inferential updating; on the other hand, the requirement that the updating rule violate the regularity principle in a maximally conservative way entails predictive updating.

After sketching the characterizations of inferential and predictive updating, I go on to discuss the relationship that inferential and predictive updating bear to each other as well as their relationship to previously proposed updating rules. In particular, I show why standard Bayesian updating and Bissiri et al.’s (2016) general Bayesian updating are a special case of inferential updating. Finally, I discuss two applications of the resulting normative framework. First, I show how we can give an analysis of the evidential import of unification that radically generalizes Myrvold’s (2017) Bayesian analysis. Second, I show how the Bayesian argument for Ockham’s razor (MacKay, 2003) may be generalized, and how standard objections to that argument may be addressed.

References

Bissiri, Pier G., Christopher Holmes, and Stephen G. Walker (2016). “A General Framework for Updating Belief Distributions.” Journal of the Royal Statistical Society. Series B (Methodological) 78(5): 1103–1130.

MacKay, David (2003). Information Theory, Inference, and Learning Algorithms. Cambridge University Press.

Myrvold, Wayne C. (2017) "On the Evidential Import of Unification," Philosophy of Science 84 (1): 92-114.

Vassend, Olav B. (Forthcoming). “A Verisimilitude Framework for Inductive Inference, with an Application to Phylogenetics.” The British Journal for the Philosophy of Science.

Levels and a New Role for Mathematics in Empirical Sciences

Atoosa Kasirzadeh
University of Toronto

Most philosophers tend to divide the roles of mathematics in scientific explanations between representational and genuinely autonomous, purely mathematical roles. Accordingly, they maintain that either mathematics is simply representational of the empirical phenomena in science or it has a non-representational, genuinely autonomous explanatory role. The representational role of mathematics is uncontroversial. Mathematics as an integral part of scientific explanations plays a significant role in idealized representations of the empirical world. In contemporary literature, this representational role is often analyzed in two ways: either by appealing to the mapping account of Pincock (2007) which suggests that there is some kind of structural isomorphism between mathematics and the empirical world; or by the extended mapping account of Bueno and Colyvan (2011), which emphasize pragmatic and context-dependent features, in addition to the structural isomorphism, in the process of applying mathematics to the empirical world.

In contrast to these representational views, some philosophers have promoted a genuinely explanatory role for mathematics in scientific explanations. For instance, Lange (2012, 2017) argues that mathematics can enter into scientific explanations by constraining the ways things can be in the empirical world through enforcement of necessarily true constraints that are modally stronger than ordinary laws of nature.

In this paper I identify a third, distinct role for mathematics, a role that may actually be pervasive in empirical sciences. I call this the "bridging" role, according to which mathematics acts as a translation scheme in our explanatory reasoning from lower-level to higher-level phenomena. First, I support this proposal by analyzing a scientific explanation of color pattern formation by mathematical biologists. Second, I suggest that the bridging role of mathematics is actually pervasive in other explanations in condensed matter physics that use limiting idealizations such as Renormalization Group explanations of critical phenomena (Batterman, 2009).

The paper is organized as follows. In Section 2, I argue that the bridging role of mathematics is distinct from both the mapping role and the genuinely autonomous role. In Section 3, I use a case study from mathematical biology to examine my proposal for a new role for mathematics, the bridging role, in detail. I discuss how mathematical biologists have appealed to different explanatory roles of mathematics to approach the problem of pattern formation on animal skins. The explanation that I discuss is of particular interest because there are two known mechanisms, one at the macro-level (Section 3.1) and the other at the micro-level (Section 3.2), that use mathematics to explain the formation of skin patterns. Here is the explanatory gap to be considered: how can we explain the cellular automaton pattern on the macro-level by the Turing models on the micro-level? The gap consists of a set of missing mathematical explananda that figures as a bridge to the macro-level explanandum.

I show how the scientists argue that the generation of the macroscopic cellular automaton pattern can be explained as the result of the combination of two components. (1) Turing models which highlight a macro-level mechanistic explanation by representing the interactions and activities among micro-scale biological skin cells; (2) mathematical facts of geometry that bridge the gap in our explanatory reasoning from the Turing microscopic model of the interactions between biological cells to the macroscopic cellular automaton pattern (Section 3.3). In particular, I discuss that the bridging role of mathematics contributes to confirming the scientific intuitions about the relation between the mechanistic explanations of the two levels. In Section 4, I argue that the bridging role of mathematics can offer an analysis for other instances such as explanations that appeal to infinite limits. In Section 5 I discuss how the two roles of mathematics, the mapping and the bridging, complement each other in multi-level explanations of empirical phenomena. Section 6 contains concluding remarks.

References

Batterman, R. W. (2009). On the explanatory role of mathematics in empirical science. The British Journal for the Philosophy of Science 61(1), 1–25.

Bueno, O. and M. Colyvan (2011). An inferential conception of the application of mathematics. Nous 45(2), 345–374.

Lange, M. (2012). What makes a scientific explanation distinctively mathematical? The British Journal for the Philosophy of Science 64(3), 485–511.

Lange, M. (2017). Because without cause: Non-causal explanations in science and mathematics. Oxford University Press.

Pincock, C. (2007). A role for mathematics in the physical sciences. Nous 41(2), 253–275.

Multiple Discoveries, Multiple Errors, and the Inevitability of Science

Luca Tambolo
University of Trieste

This paper brings together two issues that, so far, have been discussed separately: the phenomenon of multiple discovery, which sociologists and historians of science studied for decades while philosophers have, with few exceptions, ignored, and the inevitability/contingency of science controversy, to which a growing number of philosophers and historians of science have contributed since the 1990s. More specifically, it investigates the question of whether a scientific realist ought to embrace the following thesis connecting multiple discovery with the inevitability of science:

Multiple Discovery-to-Inevitability Thesis (MDIT)

If multiple discovery—whereby different scientists independently and sometimes simultaneously arrive at (more or less) the same results—is a robust, indeed, pervasive feature of scientific inquiry (as claimed, e.g., by Merton 1973), then there is reason to believe that, had scientist S not arrived at results R at time t, someone else would have: the results of successful scientific research are, in this sense, inevitable.

It would be strange for a realist not to be instinctively sympathetic with MDIT. Realism is, among other things, the view that there is a mind-independent world which science aims at, and is to a large degree successful in, describing. The realist can then offer a straightforward explanation of the fact that multiple researchers investigating a certain fragment of the world happen to independently hit upon (more or less) the same results or findings. According to such explanation, multiple discovery—epitomized by such famous cases as Newton and Leibnitz inventing the calculus independently of each other, Wallace and Darwin both putting forward the theory of evolution by natural selection, etc.—is a sign of the results in question being, in fact, a genuine, if imperfect, description of the relevant aspects of the fragment of the world under investigation. From the vantage point of the realist, it is then tempting to embrace MDIT, which views the phenomenon of multiple discovery as providing evidence for the claim that science is inevitable in the sense specified by Hacking: the results of successful science, “if correct, must be ‘essentially’ and ‘implicitly’ contained in the end-run science” (2000, S60). We shall nevertheless suggest that, especially in light of the phenomenon of multiple errors (systematically investigated for the first time by Seeman 2018), the realist should not overrate the epistemic significance that they at-tach to multiple discoveries. We shall proceed as follows.

As a preliminary step, we shall need to discuss the standing of the phenomenon of multiple discovery, since the interest of MDIT hinges on the phenomenon being a pervasive feature of the scientific enterprise. Trading on the fact that, by deploying a strict criterion of identity of discovery, many alleged instances of the phenomenon can be readily dismissed, various authors have challenged the claim that multiple discovery is a pervasive feature of science. We shall argue that, when an appropriately loose notion of multiple discovery as a matter of degree is adopted, it turns out to be an indispensable feature of scientific research—one that, due to the reward system characterizing science, is notoriously an enduring concern for scientific communities (see, e.g., Merton 1973 and, for recent examples, The PLOS Biology Staff Editors 2018).

We shall then move to briefly consider the mechanisms that produce multiple discoveries, identified by the few philosophers who systematically investigated the phenomenon in the communal nature of scientific knowledge. More specifically, it has been claimed that, given a community of researchers sharing the same background knowledge, having access to comparable equipment, and devoting an appropriate amount of resources to the investigation of a certain research problem, numerous cases of multiple discoveries are only to be expected (Kuhn 1959/1977; Lamb and Easton 1984; Scerri 2016).

Yet, it is pretty easy to imagine also cases of multiple errors, that is, scenarios in which several competent researchers working within a certain scientific field independently come up with the same wrong solution to the problem that they are investigating and the community embraces it, or they fail to recognize a certain solution to the problem as the correct one. Multiple errors may involve every aspect of scientific research, ranging from wrong interpretations of experimental data or computational results to misjudgments concerning the importance of other researcher’s achievements and the positing of spurious theoretical entities, etc. Indeed, as Seeman (2018) forcefully argues, multiple errors in science seem to be as common as multiple discoveries—if not more common, given that there are more ways to be wrong about some-thing than there are to be right. Sometimes multiple errors are quickly detected and lead to progress (as in the case of the multiple mischaracterization of the structure of ferrocene, which led to the identification of its real structure, thereby paving the way for the birth of a new area of organic chemistry), sometimes they persist for long stretches of time. Seeman suggests that one may go so far as to take the concept of multiple error to encompass all the theories that the relevant scientific communities accepted for a certain period, but later abandoned.

We shall argue that, since there seem to be no significant epistemic differences between the situations in which multiple errors occur and those in which multiple discoveries are made, the realist ought to be cautious when assessing alleged cases of multiple discovery. The sound intuition underlying the realist’s sympathy for MDIT is that if a certain result R is “out there,” waiting to be discovered so to speak, then it is likely that multiple researchers will, in the course of successful inquiry, independently hit upon it; the fact that several competent individuals arrive at the same result enhances one’s confidence in the result. The results of successful inquiry, if correct, are then inevitable in the sense that they will be preserved within the ideal final stage of science. Nevertheless, whether multiply discovered results are in fact correct—whether they are discoveries, and not errors—can only be determined after the fact. Therefore, the realist ought to remain neutral with respect to MDIT.

Natural Kinds as Real Patterns

Ana-Maria Crețu
The University of Edinburgh

The dialectic between realism and nominalism about natural kinds gives rise to a problem regarding the ontological commitment to natural kinds, which a third, recently popularized naturalist position cannot solve. Here, the 'real patterns' strategy is twice modified to solve the problem of natural kinds. First, it is shown how the extant views on real patterns can be modified to deliver a substantial ontological commitment to autonomous real patterns. Second, the real patterns strategy is further modified to incorporate natural kinds specific ontology and methods. Both modifications are novel merits of the paper.

The Commitment Problem for Natural Kinds

Recent discussions have focussed on ways in which classifications are guided and constrained by norms or interests (Slater (2017), Bursten (2016), and Cooper (2014)). However, this refocussing has exacerbated an old problem for realism about natural kinds, call it the commitment problem for natural kinds. The problem resides in an alleged incompatibility between a substantial ontological commitment, to objective or mind-independent natural kinds, and the observation that in scientific practice classifications are indexed to interests or scales, they are practice-relative. The realist is faced with a dilemma: admit that classifications are indexed to interests and scales and thus preserve a sense of alignment with scientific practice, or cling to a substantial ontological commitment to preserve objectivity. Choosing the first horn of the dilemma leads the realist dangerously close to a nominalist position. The nominalist embraces a permissive practice-relativity and argues against mind-independent natural kinds. In contrast, the second horn of the dilemma imposes such a high standard on what counts as a natural kind that few natural kinds if any can be identified in contemporary science.

The commitment problem is equally worrying for naturalists such as Khalidi (2016), Slater (2015), and Massimi (2014), who, whilst defending practice-relativity forgo mind-independence, without adopting nominalism. Sacrificing mind-independence can however leave the naturalist unable to maintain a clear-cut distinction between natural kinds and non-natural kinds. Moreover, as noted by Hacking (2007), echoing Goodman (1988/1978), "many questions posed in the context of natural kinds – induction for example – arise equally for other kinds of things such as machinery or musical works" (p. 204). So, whilst naturalised accounts are practice-relative, such accounts can at best preserve a paper-thin distinction between natural kinds and non-natural kinds. Hence, despite efforts from naturalists the commitment problem remains unresolved.

The Real Patterns Strategy

Confronted with an equally thorny ontological problem, Dennett (1991) introduced the real patterns strategy as a means to deal with ontological problems. Dennett claimed that real patterns are 'real' although they need to be recognized from a perspective. Dennett fails to explain what a perspective is and whether patterns can be independent of a perspective. The real patterns strategy has been enriched in different ways by Ross (1995), Ladyman and Ross (2007), and Wallace (2011). However, none of these further developments of the strategy secured a substantial ontological commitment to mind-independent entities. Ross' view explains what underpins the phenomena of microeconomics but describes real patterns in terms of ‘usefulness’ thus inviting the same criticism as Dennett's view. Ladyman and Ross provide a way of making the ontology of science compatible and continuous with recent developments in a variety of scientific practices (e.g. physics, astrophysics, geology among others). However, they make real patterns dependent on a “physically possible perspective” which is at the same time too restrictive and too liberal. Wallace's application of the real patterns strategy scores high on both practice-relativity and scientific objectivity, but his view makes patterns dangerously dependent on (successful) theories. None of these views clearly explains how one can think of patterns as independent or autonomous entities irrespective of perspectives.

Natural Kinds as Real Patterns

If the real patterns strategy can indeed be used to tackle ontological problems, in particular the commitment problem for natural kinds, real patterns need to be disentangled from perspectives. It will be argued that one can adapt Laudan’s (1977) distinction between ‘research traditions’ and ‘theories’ to fix the real patterns strategy. In particular, it will be shown that the real patterns advocates ought to have distinguished between the stage of identifying a pattern from within a research tradition and the stage of endowing a pattern with a theoretical identity within a perspective or ‘theory’ in Laudan’s sense. On the modified account real patterns are ‘authenticated’ within a ‘research tradition’ and this is what makes them real, mind-independent entities. Real patterns are further studied from within a perspective, which preserves the ontological commitment of the research tradition. The perspective seeks to understand the nature of the real patterns and in that it can be successful or fail. However, when a perspective is abandoned, one need not also abandon the ontological commitments warranted by the research tradition. By failing to distinguish between the authentication of real patterns and the various perspectives that can be had on their nature, the advocates of the real patterns’ strategy have invited an unnecessary doubt with regard to the ontological commitment that the real patterns strategy can deliver.

Having shown how one can fix the real patterns strategy in general terms, the commitment problem for natural kinds is further tackled using the modified approach. The only attempt to apply the real patterns strategy to natural kinds was undertaken by Ladyman and Ross. Whilst Ladyman and Ross are right to suggest that natural kinds are real patterns, their view cannot in fact support this claim. Since natural kinds are inherently tied to the ideas of likeness and difference which require for their application the existence of objects, contra Ladyman and Ross, at least some things must stay. In particular, it will be argued that their view must be modified to include a distinction between real patterns (qua relations) and objects (qua relata) to be applicable to natural kinds. With the real patterns strategy twice modified, natural kinds as real patterns provides a solution to the commitment problem.

Negotiating History: Contingency, Canonicity, and Case Studies

Joseph D. Martin and Agnes Bolinska
University of Cambridge

Recent work on the use of historical case studies as evidence for philosophical claims has advanced several objections to this practice. Our two-fold goal is first to systemize these objections, showing how an appropriate typology can light the path toward a resolution, and second, to show how some of these objections can be recast as advantages for the historically sophisticated philosopher, specifically by describing how attention to contingency in the historical process can ground responsible canonicity practices.

Systematizing objections to the use of historical case studies for philosophical ends shows that they fall largely into two categories: methodological objections and metaphysical objections. The former, we argue, fail to be distinctive—they do not identify special challenges from other forms of philosophical reasoning are immune. Case studies demand responsible handling, but this is unsurprising. History is messy and philosophy is difficult. But the need for care is hardly the mark of a hopeless endeavor. Rather, attention to the ways in which history is messy and in which philosophy is difficult can be resources for developing better historiographical and philosophical practices.

Metaphysical objections do, however, raise special problems for the use of historical case studies. We show that attention to what makes for a canonical case can address these problems. A case study is canonical with respect to a particular philosophical aim when the philosophically salient features of the historical system provide a reasonably complete causal account of the results of the scientific process under investigation. We show how to establish canonicity by evaluating relevant contingencies using two prominent examples from the history of science: Eddington's confirmation of Einstein's theory of general relativity using his data from the 1919 eclipse and Watson and Crick's determination of the structure of DNA. These examples suggest that the analogy between philosophical inquiry and the natural sciences, although imperfect, has important elements that make it worth retaining. This is not to say that we should think of philosophy as modeled on scientific practice, but rather that both succeed by virtue of something more general: their reliance on shared principles of sound reasoning.

Taking seriously the practices necessary to establish the canonicity of case studies makes clear that some examples of the historical process of science are more representative of its general ethos than others. With historiographical sense, we can pick these examples out. Doing so requires attention to the contingencies of history. Rather than undermining the use of historical cases, philosophical attention to contingency aids the development of case studies as resources by making explicit otherwise tacit assumptions about which features of them are most salient and why.

These considerations help us address the question of the value of history of science for the philosophy of science. It is possible, even easy, to use the rich resources that history provides irresponsibly to make a predetermined point. But that is not a genuine case of history of science informing philosophy of science—in part because it proceeds in the absence of historiographical sense. By outlining the practices that render particular cases canonical for certain philosophical aims, we have offered a route by which such sense can be integrated into standard philosophical practices.

On the Individuation of Choice Options

Roberto Fumagalli
King's College London; London School of Economics; University of Pennsylvania

Several purported violations of decision theory's axiomatic requirements can be accommodated by modifying how agents' choice options are individuated and formally represented. In recent years, prominent authors have criticized these modifications for trivializing decision theory, undermining the theory's falsifiability, imposing cognitively overdemanding requirements on real-world agents and hampering the internal coherence of decision theory's mathematical formalism. In this paper, I draw on the best available empirical and theoretical works in contemporary decision theory to address these prominent criticisms. In doing so, I articulate and assess several different criteria for individuating and formally representing agents' choice options.

Extended Abstract

Standard decision theory builds on specific axiomatic requirements on agents' preferences, together with the representation theorems derivable from these requirements (e.g. von Neumann and Morgenstern, 1944, Savage, 1954). Such representation theorems demonstrate that if an agent's preferences satisfy specific axiomatic requirements, then this agent's choices can be represented as if the agent maximizes expected utility (e.g. Bradley, 2017, ch.2, Okasha, 2016). Over the last few decades, several purported violations of decision theory's axiomatic requirements have been documented across experimental settings (e.g. Anand, 1993, Machina, 2008, Starmer, 2000). Many of these purported violations can be accommodated by modifying how agents' choice options are individuated and formally represented (e.g. Broome, 1991, ch.5, Dietrich and List, 2016). In recent years, prominent authors have criticized these modifications for trivializing decision theory (e.g. Hausman, 2000, Steele, 2010), undermining the theory's falsifiability (e.g. Bhattacharyya et al., 2011, Hampton, 1994), imposing cognitively overdemanding requirements on real-world agents (e.g. Bales et al., 2014, Gilboa et al., 2009) and hampering the internal coherence of decision theory's mathematical formalism (e.g. Alexander, 2012, Sugden, 1991).

In this paper, I draw on the best available empirical and theoretical works in contemporary decision theory to address these prominent criticisms. In doing so, I articulate and assess several different criteria for individuating and formally representing agents' choice options. The contents are organized as follows. In Section 2, I briefly outline decision theory's axiomatic requirements and examine these requirements' observed violations. In Section 3, I explicate how such violations can be accommodated by modifying how agents' choice options are individuated and formally represented (re-individuation strategy). In Sections 4-7, I identify and address four major objections put forward against such re-individuation strategy, namely: the trivialization objection (e.g. Hausman, 2000, Steele, 2010); the falsifiability objection (e.g. Bhattacharyya et al., 2011, Hampton, 1994); the objection from cognitive overdemandingness (e.g. Bales et al., 2014, Gilboa et al., 2009); and the objection from theoretical incoherence (e.g. Alexander, 2012, Sugden, 1991).

Over the last few decades, decision theorists have made substantial advances in specifying how distinct versions of decision theory are to be applied when one lacks precise probabilities (e.g. Bradley and Stefánsson, 2017, Levi, 1974, Joyce, 1999) and well-defined utility functions (e.g. Buchak, 2013, Gilboa, 2009, Hare, 2010). However, comparatively little progress has been made in the provision of plausible and informative criteria for individuating and formally representing agents' choice options (e.g. Broome, 1993, Hedden, 2012). My evaluation aims to fill this major lacuna in the decision theoretic literature and thereby contribute to the development of a descriptively and normatively adequate decision theory.

REFERENCES

Alexander, J. 2012. Why the angels cannot choose Australasian Journal of Philosophy, 90, 619-640.

Anand, P. 1993. The philosophy of intransitive preference. The Economic Journal, 103, 337-346.

Bales, A., Cohen, D. and Handfield, T. 2014. Decision theory for agents with incomplete preferences. Australasian Journal of Philosophy, 92, 453-470.

Bhattacharyya, A., Pattanaik, K. and Xu, Y. 2011. Choice, internal consistency and rationality. Economics and Philosophy, 27, 123-149.

Bradley, R. 2017. Decision Theory with a Human Face. Cambridge University Press.

Bradley, R. and Stefánsson, O. 2017. Counterfactual Desirability. British Journal for the Philosophy of Science, 68, 485-533.

Broome, J. 1991. Weighing Goods. Basil Blackwell.

Broome, J. 1993. Can a Humean be moderate? In Frey, R. and Morris, C. Ed. Value, Welfare and Morality, 51-73. Cambridge University Press.

Buchak, L. 2013. Risk and Rationality. Oxford University Press.

Dietrich, F. and List, C. 2016. Reason-based choice and context dependence: an explanatory framework. Economics and Philosophy, 32, 175-229.

Gilboa, I. 2009. Theory of Decision under Uncertainty. Cambridge University Press.

Gilboa, I., Postlewaite, A. and Schmeidler, D. 2009. Is it always rational to satisfy Savage's axioms? Economics and Philosophy, 25, 285–296.

Hampton, J. 1994. The failure of expected utility theory as a theory of reason. Economics and Philosophy, 10, 195-242.

Hare, C. 2010. Take the sugar. Analysis, 70, 237-247.

Hausman, D. 2000. Revealed preference, belief, and game theory. Economics and Philosophy, 16, 99-115.

Hedden, B. 2012. Options and the subjective ought. Philosophical Studies, 158, 343-360.

Joyce, J. 1999. The Foundations of Causal Decision Theory. Cambridge University Press.

Levi, I. 1974. On indeterminate probabilities. Journal of Philosophy, 71, 391-418.

Machina, M. 2008. Non-expected Utility Theory. In The New Palgrave Dictionary of Economics. 2nd Ed., Durlauf, S. and Blume, L., 74-84. New York: Palgrave Macmillan.

Okasha, S. 2016. On the interpretation of decision theory. Economics and Philosophy, 32, 1-25.

Savage, L. 1954. The Foundations of Statistics. John Wiley and Sons.

Starmer, C. 2000. Developments in nonexpected utility theory: the hunt for a descriptive theory of choice under risk. Journal of Economic Literature 3, 38, 332-382.

Steele, K. 2010. What are the minimal requirements of rational choice? Arguments from the sequential-decision setting. Theory and Decision, 68, 463-487.

Sugden, R. 1991. Rational choice: a survey of contributions from economics and philosophy. The Economic Journal, 101, 751-785.

Von Neumann, J. and Morgenstern, O. 1944. Theory of Games and Economic Behavior. Princeton University Press.

Outline of a Kairetic Account of Explanation in Mathematics

Michele Lubrano
Università di Torino

I would like to present an account of explanation in mathematics along the lines of Strevens (2011), namely an account based on the notion of difference-maker. I'm going to illustrate what such an account consists in and why it deserves attention and further research effort.

Explanation in mathematics is one of the most interesting aspects of mathematical practice. Professional mathematicians not only want their theorems to be correctly proven, they also want them satisfactorily explained. Only in relatively recent times, philosophers have started to pay attention to the issues of what a mathematical explanation consists in and how it works (see Steiner 1978).

There are two classical views of explanation within mathematics: a local model and a holistic model (see Mancosu 2018). The local model, first presented by Steiner (1978), is one in which a proof of a theorem T is explanatory when T is deduced from the essence, or nature, of the mathematical objects involved. The holistic model (see Kitcher 1989) says that a proof of a theorem T is explanatory if it shows that the behaviour of the structures or entities mentioned in T can be subsumed under a general unifying pattern, from which the behaviour of different structures or entities can be deduced. These two models have both virtues and limits, which I'm not going to examine in my presentation. What can be said is that there is a general consensus on the fact that the two models work well in some cases and are unsatisfactory in others (see Mancosu 2018).

Before giving up every hope of a unified account of explanation in mathematics it's worth trying to formulate an alternative account along different lines. The point I would like to start from is the simple observation that one of the tasks that an explanatory proof of a theorem T must accomplish is to clearly indicate which are the properties T depends on. The notion of dependence might be the key for a deep understanding of explanation in mathematics. In order to illustrate this notion, I choose, as a guiding example, the case of reverse mathematics.

Reverse mathematics is an important research program initiated by Friedman (1975), whose aim is do the reverse path of the most common mathematical research: instead of going from axioms to new theorems, it goes from already known theorems to their axioms. More precisely, the kind of questions that it aims at answering is: which is the weakest group of axioms that we need in order to prove theorem T of ordinary mathematics? For a surprisingly high number of theorems, this question has a perfectly defined answer. Such an answer is often one of the several subsystems of Full Second Order Peano Arithmetic, in symbols, Z2. Several subsystems of Z2 have been extensively investigated and have been ordered on the basis of their demonstrative power: the closer they are to demonstrating all the theorem demonstrated by Z2 the higher they are in the hierarchy.

I'll point my attention to two important subsystems: Arithmetic Comprehension Axiom (also known as ACA0) and Recursive Comprehension Axioms (also known as RCA0). The differences between the two are entirely due to a difference in the strength of the Comprehension Axioms Schema. Now, this difference in demonstrative power can be made more precise by listing some theorems that can be proven in a subsystem, but not in a weaker one. For example, Bolzano-Weierstrass Theorem can be proven in ACA0, but not in RCA0. It can be shown that RCA0 is the most powerful subsystem of Z0 in which Bolzano Weierstrass statement is false and ACA0 is the weakest in which it is true. Since the only difference between the two lies in the strength of Comprehension Axiom Schema, the most natural way to describe the situation is to say that the strengthening of Comprehensions (i.e. its upgrading from Recursive to Arithmetical) is what makes the difference between Bolzano-Weierstrass statement being true and its being false. In the context of a hierarchy of subsystems of Z2, Arithmetic Comprehension is what such a statement crucially depends on.

The relation of crucial dependence that is in play here and how it is related with the notion of difference-maker can be illustrated by means of this definition:

Crucial Dependence: the truth of a statement T crucially depends on axiom A if and only if, given a hierarchy of systems of increasing strength (S1, …, Sn),

Si instead of Si-1 proves T instead of non-T,

and Si = Si-1 + A}.

This definition is a good generalisation of the phenomenon described in the example of RCA0 and ACA0. Indeed, ACA0 instead of RCA0 proves Bolzano-Weierstrass Theorem instead of its negation. Moreover, by adding Arithmetic Comprehension to RCA0, we get ACA0. Therefore, Arithmetic Comprehension is what makes the difference between proving Bolzano-Weierstrass statement or proving its negation, in the context of subsystems of Z2. Bolzano-Weierstrass statement crucially depends on Arithmetic Comprehension, in such a context.

Now, a proof of a theorem T is explanatory if and only if T is deduced, among other things, from a make-difference axiom in the context of a suitable hierarchy of formal systems.

In addition to a presentation of this account of explanation in mathematics, my talk will include a brief comparative analysis of two proofs: one explanatory according to my account, the other not explanatory.

Bibliography

- Friedman, Harvey 1975. 'Some systems of second-order arithmetic and their use'. Proceedings of the International Congress of Mathematicians, Vancouver 1974, Vol. 1: 235-242.

- Kitcher, Philip 1989. 'Explanatory Unification and the Causal Structure of the World, in P. Kitcher and W. Salmon (eds.), Scientific Explanation, (Minnesota Studies in the Philosophy of Science, Volume XIII), Minneapolis: University of Minnesota Press, 410–505.

- Mancosu, Paolo 2018. 'Explanation in Mathematics'. The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.).

- Steiner, Mark 1978. 'Mathematical Explanation'. Philosophical Studies, 34 (2): 135-151.

- Strevens, Michael 2011. Dept: An Account of Scientific Explanation. Harvard University Press.

Perspectival Realism About Mechanistic Functions

Joe Dewhurst
Ludwig Maximilian University of Munich

The attribution of a function to a (putative) mechanism plays an important role in mechanistic explanation, both for characterising the phenomenon that is to be explained and for developing an explanation of that phenomenon. Advocates of mechanistic explanation must therefore offer some account of functional attribution, and they typically do so either in teleological terms, endorsing the idea that each mechanism has a distinct and determinable proper function, or in non-teleological terms, where any causal process can potentially qualify as functional. In this paper I will explore an alternative, perspectival realist approach to functional attribution, where the function performed by a mechanism depends on the explanatory context, but is also constrained by objective features of the world, such as the physical structure of the mechanism and its environmental context. I will first describe the role played by functional attribution in mechanistic explanation, before considering some of the strengths and weaknesses of existing accounts of mechanistic functions, and then finally proposing my own novel account and explaining its benefits.

All accounts of mechanistic explanation emphasise that there are no mechanisms simpliciter, but rather every mechanism must be identified as a mechanism for the production of some phenomenon (see e.g. Machamer, Darden, & Craver 2000: 3; Glennan 2002: S344; Bechtel & Abrahamsen 2005: 423; Illari & Williamson 2012: 120). This is commonly understood in terms of a mechanism having the function of producing some phenomenon (cf. Garson 2013), and in some cases the phenomenon itself is partly characterised in terms of a function that we have prior reason to attribute to a mechanism. An additional contribution that functional attribution makes to mechanistic explanation is to guide the investigation of a mechanism once a phenomenon has been identified. This is particularly important for functional decomposition, where the aim is to identify the component parts necessary for carrying out some function, and which can only proceed once a function has been attributed to a mechanism.

Any proponent of mechanistic explanation must therefore provide some account of how we determine the function of a mechanism, and there are many options available. Etiological accounts determine the function of a mechanism in terms of its evolutionary history (see e.g. Millikan 1989, Neander 1991), selectionist accounts appeal more generally to a mechanism's causal history (see e.g. Garson 2017), while goal-directed accounts look at the current contribution a mechanism makes to the aims of an organism or system (see e.g. Maley & Piccinini 2017). Each of these accounts is usually understood to be teleological, insofar as we can say what the 'proper' function of any given mechanism is. There are also non-teleological alternatives to functional attribution, such as the causal role account, where every causal contribution made by a system qualifies as functional (Cummins 1975; cf. Craver 2001), or perspectival accounts (Craver 2013; cf. Hardcastle 1999), according to which the function attributed to a mechanism depends on our explanatory interests. The advantage of these latter approaches is that they avoid any outstanding epistemological or metaphysical concerns about natural teleology, and can accomodate functional descriptions of systems that we would not normally describe as teleological, such as the astrophysical mechanisms discussed by Illari & Williamson (2012). The disadvantage is that they can appear overly liberal, making it seem too easy to attribute any function to any system.

My aim in this paper is to build on existing perspectival approaches by exploring the ways in which our perspectival attribution of a function to a mechanism might be constrained by objective features of that mechanism, such as its physical structure and environmental context. The account proposed here therefore constitutes a kind of perspectival realism about mechanistic functions (cf. Giere 2006, Massimi 2012), where the function that we (perspectively) attribute to a (real) mechanistic structure depends on certain objective features of that structure, but is at the same time sensitive to our explanatory aims and objectives (cf. Kästner 2018 for a related approach). I will present the account, illustrated with examples drawn from biology and cognitive science, and defend it against some initial concerns to do with both its realist and perspectivalist credentials. The account will therefore constitute a middle path between these two extremes. In concluding I will suggest some ways in which this account might prove beneficial for mechanistic explanation more generally, including its capacity to provide a foundation for a motivated and non-trivial scientific pluralism.

Bechtel, W. & Abrahamsen, A. 2005. “Explanation: A Mechanistic Alternative.” Studies in History and Philosophy of the Biological and Biomedical Sciences, 36: 421-41.

Craver, C.F. 2001. “Role Functions, Mechanisms and Hierarchy.” Philosophy of Science, 68: 31-55.

Craver, C.F. 2013. “Functions and Mechanisms: A Perspectivalist Account.” In Huneman (ed.), Functions: Selection and Mechanisms. Dordrecht: Springer.

Cummins, R. 1975. “Functional Analysis.” The Journal of Philosophy, 72/20: 741-65.

Garson, J. 2013. “The Functional Sense of Mechanism.” Philosophy of Science, 80: 317-33.

Garson, J. 2017. “A Generalized Selected Effects Theory of Function.” Philosophy of Science, 84/3: 523-43.

Giere, R. 2006. Scientific Perspectivism. Chicago, IL: CUP.

Glennan, S. 2002. “Rethinking Mechanistic Explanation.” Philosophy of Science, 69: S342-53.

Hardcastle, V. 1999. “Understanding Functions: A Pragmatic Approach.” In Hardcastle (ed.), When Biology Meets Philosophy. Cambridge, MA: MIT Press.

Illari, P.M. & Williamson, J. 2012. “What is a Mechanism?: Thinking about Mechanisms Across the Sciences.” European Journal for Philosophy of Science, 2: 119-35.

Kästner, L. 2018. “Integrating mechanistic explanations through epistemic perspectives.” Studies in the History and Philosophy of Science, 68: 68-79.

Machamer, P.K., Darden, L. & Craver, C.F. 2000. “Thinking about Mechanisms.” Philosophy of Science, 67: 1-25.

Maley, C.J. & Piccinini, G. 2017, “A Unified Mechanistic Account of Teleological Functions for Psychology and Neuroscience.” In Kaplan (ed.), Explanation and Integration in Mind and Brain Science. Oxford: OUP.

Millikan, R.G. 1989. “In Defense of Proper Functions.” Philosophy of Science, 56/2: 288-302

Neander, K. 1991. “The Teleological Notion of ‘Function’.” Australasian Journal of Philosophy, 69: 454-68.

Massimi, M. 2012. "Scientific perspectivism and its foes." Philosophica 84: 25-52.

Predictable Behaviour and Intentional Action: Disentangling the Two

Catherine Greene
London School of Economics

The social sciences often aim to predict human behaviour. In doing this, it is taken for granted that social scientists deal with intentional behaviour, rather than reflexive, or biological phenomena. Using studies of middle bias, which is an example of predictable behaviour, this paper argues that predictable behaviour may not be intentional, in the way that intentional behaviour is ordinarily understood. Drawing on O'Shaughnessy's description of sub intentional acts, it argues that there is a category of behaviour that is neither intentional, nor non-intentional; but derivatively intentional. Derivatively intentional behaviour is predictable, largely because of the characteristics that distinguish it from intentional behaviour.

Attali and Bar Hillel (2003) describe the phenomenon of middle bias, which is the tendency for people to pick middle options from linearly arranged choices; this includes the placement of correct answers when setting multiple choice questions, picking a number between 1 and 10, and deciding which toilet stall to go into. This behaviour cannot be described as straightforwardly intentional. This is because agents showing middle bias are unaware of their tendency and have no psychological, or conscious attitudes, whether beliefs or desires, or plans, regarding this tendency. Middle bias therefore does not fit well within the desire-belief account of intentional action (Davidson 1963 & 1978, Garcia 1990, Davis 1984), the planning account (Bratman 1979 & 1999), or non-causal accounts (Frankfurt 1978, Grunbaum 2007 & 2010, Castaneda 1982 & 1992). This paper draws on two insights to characterise middle bias. Castaneda argues that some actions are only intentional by virtue of the action of which they are a part. He says that intentional action may bring with it other, unintended, actions which nevertheless form part of the larger, intended, action. O'Shaughnessy (1980) describes sub-intentional acts which, he says, are acts which are intentional under no description. O'Shaughnessy's examples include the movements of a person's tongue, or fingers, while they are performing other actions such as talking, or listening to music. Neither Castaneda nor O'Shaughnessy characterise middle bias satisfactorily, but their analysis suggests how this is to be done.

Middle bias is an example of derivatively intentional behaviour. Derivatively intentional behaviour is defined by the following necessary and jointly sufficient conditions:

1. Derivatively intentional behaviour is a subset of intentional behaviour.

2. Derivatively intentional behaviour occurs whenever the intentional behaviour (of which it is a subset) is undertaken.

3. An actor is, initially, unaware of this derivatively intentional behaviour and has no beliefs, desires, plans, or other deliberations relating to the derivatively intentional behaviour.

4. If an actor becomes aware of the derivatively intentional behaviour they realise that it is something they are doing. It is not something that is merely happening to them.

5. If an actor becomes aware of the derivatively intentional behaviour, it can be brought under the actor's control, if they choose to do so.

Derivatively intentional behaviour is often predictable. This paper argues that this is because people have no intentions about the aspects of behaviour that are predictable- their intentions are directed towards the larger behaviour, of which they are a part. Nevertheless, it would be incorrect to describe derivatively intentional behaviour as non-intentional because it is tied in with intentional behaviour. Furthermore, when people become aware of derivatively intentional behaviour they may try to alter it, thereby making it less predictable. In the example above, people intend to set a multiple-choice exam, or pick a number in answer to a request to do so, but they do not intend their behaviour to conform to a pattern. However, once they are made aware of this pattern, they may consciously change their behaviour. Other examples of derivatively intentional behaviour include the sentence structure of novels, patterns in internet surfing (Halvey et al 2006), patterns in individuals locations over time (Song et al 2010). More tentatively, the paper proposes that the concept 'demand' from economics meets many of the criteria for derivative intentionality so it is no accident that the relationship between demand and supply is one of the best confirmed regularities in the social sciences. The analysis of derivative intentionality also suggests why predictions, or regularities, concerning derivatively intentional behaviour may not persist. If people become aware of the predictable nature of their behaviour, they are able to change it.

Bibliography

Attali, Y.; Bar-Hillel, M. (2003) "Guess Where: The Position of Correct Answers in Multiple-Choice Test Items as a Psychometric Variable", Journal of Educational Measurement Summer 40. 2. 109-128

Bratman, M (1979) "Simple intention" Philosophical Studies. 36.3. 245-259

Bratman, M. (1999) Intention, Plans and Practical Reason. CSLI Publications

Castaneda, H-N. (1982) "Conditional intentions, intentional action and Aristotelian practical syllogisms" Erkenntnis 18.2. 239-260

Castaneda, H-N. (1992) "Indexical reference and bodily causal diagrams in intentional action" Studia Logica 51.3. 439-462

Davidson, D. (1963) "Actions, Reasons, and Causes" The Journal of Philosophy. 60.23. 685-700

Davidson, D (1978/ 2002) "Intending" in D. Davidson 'Essays on Actions and Events' Oxford: Clarendon Press. Second Edition 83-102

Davis, W. (1984) "A causal theory of intending" American Philosophical Quarterly. 21.1. 43-54

Frankfurt, H.G. (1978) "The problem of action" American Philosophical Quarterly 15. 2. 157-162

Garcia, J. L. A. (1990) "The intentional and the intended" Erkenntnis 33.2. 191-209

Grunbaum, T. (2007) "The body in action" Phenomenology and the Cognitive Sciences 7.2. 243-261

Grunbaum, T. (2010) "Action and Agency" in 'Handbook of Phenomenology and Cognitive Science' edited by S. Gallagher and D. Schmickinh. Springer Science. 337-354

Halvey, M.; Keane, M. T.; Smyth, B. (2006) "Mobile web surfing is the same as web surfing" Communications of the ACM 49.3. 76-82

O'Shaughnessy (1980) The Will, a Dual Aspect Theory Volume 2. Cambridge University Press

Song, C.; Qu, Z.; Blumm, N.; Barabási A-L. (2010) "Limits of Predictability in Human Mobility" Science, New Series 327.5968. 1018-1021

Prediction Markets and Extrapolation

Robert Northcott
Birkbeck, University of London

Extrapolation enables us to predict outcomes in a new domain by using knowledge drawn from a different domain. The problem of the extrapolator's circle is, roughly speaking, that in order to extrapolate a predictive model or causal relation to a new domain it is necessary to know that it applies in this new domain, but that in order to establish the latter it is necessary in turn to examine the new domain – thus negating the main benefit of extrapolation, which is precisely that we can avoid having to examine the new domain (Steel 2008). Extrapolation is hugely important in social science: will a policy intervention in one region have the same effect in a different one, or the same effect 10 years later? A closely related issue is the external validity of findings from experiments: will an economic behavior observed in the laboratory be repeated in the field, or the result of a field trial in one country still hold in another? These methodological issues have been the focus of much recent philosophy of social science (e.g. Cartwright and Hardie 2012).

How might the extrapolator's circle be overcome? Steel's own suggestion is that a mixture of causal process tracing and background knowledge may enable identification of relatively downstream causal nodes in the new domain that screen off any disanalogies between domains at upstream nodes. But, as several have pointed out, this requires us to identify which causal nodes can play this role, which in turn requires knowledge of the new domain, and so the circle problem returns. A commonly proposed alternative solution is knowledge of mechanisms, presuming these to operate stably across domains. The main problem with this is that, in practice, mechanisms are often fragile, i.e. behave erratically. Their behavior in new domains may be unknown, as may their behavior in interaction with other mechanisms – and often we don't know all the relevant mechanisms that might be present. Howick et al (2013) emphasize these difficulties in the context of medicine. Worries about the robustness of mechanisms are even stronger in social science; thus, so are worries about extrapolation.

We argue here for a previously unappreciated solution to the extrapolator's circle, namely predicting outcomes in a new domain by using prediction markets (when they are available, which is mainly in social science cases). Prediction markets are markets for placing bets on future or otherwise unknown events. The price signals arising in such markets, if interpreted as probability assignments, constitute implicit predictions. Prediction markets are attractive because they have a track record of (relative) predictive success (e.g. AUTHOR 2016). They work well when there are at least some informed traders on the market – indeed, going by the current empirical literature, this seems close to a sufficient condition. If so, then the only thing you need to know as a market maker is that, somewhere in the pool of traders you attract, there will be some who are informed. Given that, a prediction market may then be used to predict outcomes in a new domain. Crucially, what you don't need to know is any particular theory about the new domain. Of course, individual traders on the market might make any number of theoretical assumptions, and (lucky guesses aside) those assumptions will usefully inform the market's output only in so far as they lead to good predictions. But the market maker need presuppose almost no theory whatsoever.

In effect, prediction markets are mechanisms that do extrapolate across domains easily because they require unusually minimal assumptions. In particular, they require only that there exist some informed traders, plus that there is sufficient market liquidity, available data, legal infrastructure, and so forth. There is no need to identify key nodes in causal sequences. There is also no need to assume what is often implausible in social science, namely that causal relations will be robust across domains.

Instead, the relevant causal relation is, so to speak, a higher-level one, namely that prediction markets cause accurate predictions. The mechanisms underpinning that causal relation are likely to be stable across domains: the higher-level fact that informed traders and background infrastructure are sufficient for prediction markets to predict successfully, would seem to hold independently of the particular topic that the predictions are about.

It is one thing to predict actual events. But extrapolation also often concerns conditional 'predictions', e.g. about the result of possible or counterfactual interventions. Hitherto, there has been no evidence that guidance about such conditional predictions can be given by prediction markets because by definition in conditional cases no actual event ever occurs that settles market participants' bets, at least not within the timeframe of interest. But recent experimental research (co-conducted by one of us) now suggests that so-called self-resolving prediction markets, i.e. markets for non-actual events, operate just as reliably as markets for actual events (AUTHOR forthcoming). We report on that ongoing research here, including the intricate comparisons it requires between the behavior of actual and self-resolving markets, and the insight it provides into what factors might threaten the efficient operation of self-resolving markets in particular. When self-resolving markets do work, prediction markets as a whole in effect achieve all of the goals of extrapolation, namely successful prediction of both actual and non-actual events in a new domain. In which case, where applicable, they solve the extrapolator's circle.

Predictive Infelicities and the Neo-Humean Conception of Laws

Chris Dorst
Washington University in Saint Louis

Recent Humean theories of laws have increasingly emphasized the predictive function of the laws in scientific practice. [See Hicks's (2018) Epistemic Role Account, Dorst's (2018) Best Predictive System Account, and Jaag and Loew's (forthcoming) Cognitive Usefulness Account. The germ of these theories traces back to Hall (ms).] This paper presents several related problems for these theories and then considers how a supporter could try to defend them.

One of the chief motivations behind the neo-Humean view of laws is the observation that putative laws of nature found in scientific practice exhibit a number of features that make them predictively useful:

1) Highly informative dynamical implications

2) Wide applicability

3) Spatiotemporal locality

4) Spatial, temporal, and rotational symmetries

These features allow us to use the laws to forecast the behaviors of a wide variety of systems using information that we are usually in a position to ascertain empirically. For example, the laws' extensive dynamical implications provide us with useful information about the future temporal evolution of physical systems in our vicinity, and their various symmetries allow us to calculate these temporal evolutions without having to locate and orient ourselves in spacetime.

Neo-Humean views are developments of orthodox Humeanism. According to orthodox Humeanism, the laws of nature are the regularities that figure into the simplest and strongest systematization of the totality of the particular matters of fact. [See Lewis (1973, 1986).] By contrast, according to the neo-Humean views articulated by Hicks, Dorst, and Jaag and Loew, the laws of nature are the regularities that figure into the systematization of the particular matters of fact that is maximally predictively useful. These neo-Humean theories thus replace the standards of simplicity and strength with standards that are designed to generate predictively useful principles—standards such as informative dynamics, wide applicability, spatiotemporal locality, etc.

The benefits of this shift in standards from orthodox to neo-Humeanism are manifold, but perhaps the most significant payoff is that it generates a compelling selectionist explanation of the laws' manifest predictive utility. The rough picture is that while most of the patterns in the particular matters of fact would be utterly useless for predictive purposes, a small subset of them would possess the right combination of features to make them predictively useful, and on the neo-Humean view, the laws of nature just are those regularities in the particular matters of fact that are maximally predictively useful.

One of the main obstacles confronting the neo-Humean view is that even though actual putative laws of nature are predictively useful on the whole, there are nevertheless various predictive "infelicities" in our best physical theories that stand in need of explanation. For example, the indeterminism and nonlocality of quantum mechanics sometimes render it impossible to use that theory to generate informative and accurate predictions about systems of interest. Relatedly, it is often difficult to obtain a complete state description of the system of interest which can be plugged into the dynamical equations of theories like quantum mechanics and general relativity. And lastly, the requirement that laws of nature be exceptionless is difficult to square with the aim of predictive utility, for it seems plausible that the most predictively useful system could occasionally generate incorrect predictions, as long as those errors were (a) relatively minor and (b) compensated for by the system's predictive usefulness in other respects.

This paper reviews the details of these predictive infelicities and then considers various strategies that the neo-Humean might use to account for them. The strategies considered fall into three categories.

First, in some cases the neo-Humean can argue that the alleged predictive infelicity is actually not an infelicity at all, but is rather a requirement on the predictive utility of the theory in the first place. This strategy can be used to address the suggestion that there is a tension between the laws' exceptionlessness and their predictive utility. Roughly, the thought is that in evaluating the predictive utility of a system, creatures like us can never know for sure that it won't lead us radically wrong in the future. We thus need to be responsive to evidence about the reliability of that system, and if it occasionally licenses false predictions, this is evidence against that system's reliability. Hence the aim of predictive utility leads us to prefer systems with exceptionless regularities.

Second, the neo-Humean can explain predictive infelicities by appealing to tensions between different standards that are meant to facilitate prediction. As one example, the two standards of informative dynamics and spatiotemporal locality may come into a rather straightforward conflict. For notice that a highly informative dynamics is easier to achieve the more variables we allow to figure into it. But conversely, the standard of spatiotemporal locality effectively acts as a restriction on the class of variables that we may appeal to in our dynamics: only spatiotemporally local variables are allowed. Thus it is not too surprising that we end up with theories, like quantum mechanics, that cannot satisfy both standards perfectly.

Third, the neo-Humean can point to the inherent instability of the very concept of predictive utility. Predictively useful principles both (1) function to increase our epistemic grasp, and (2) are influenced by our epistemic grasp. The issue is that anything that satisfies both (1) and (2) is going to be systematically unstable. More specifically, since their function is fundamentally ampliative, predictive principles may eventually allow us to ascertain the very sorts of facts of which we were previously ignorant, and ignorance of which informed our very standards of predictive utility. In other words, as they increase our epistemic grasp, they may simultaneously undermine their own optimality for that task. The more we know, the fewer constraints there are on what variables our set of predictive principles may appeal to. One consequence of this is that many of the standards designed to facilitate prediction are in fact defeasible, and their importance is likely to diminish as our epistemic grasp increases. The neo-Humean can appeal to this fact to explain the appearance of predictive infelicities in our best theories.

Probing Novelty at the Large Hadron Collider: Heuristic Appraisal of Disruptive Experimentation

Sophie Ritson
Alpen-Adria Universität

In this paper, I will explore 'novelty' through a recent historical episode from high-energy experimental physics, to offer an understanding of novelty as disruption. I call this the '750 GeV episode', an episode where two Large Hadron Collider (LHC) experiments, CMS and ATLAS, each independently observed indications of a new resonance in the same mass region. Several physicists indicated, both at the time and then in subsequent reflections, that if the statistical significance of the result increased with more data, to the point where a discovery claim could be made, then this would be more novel than the Nobel Prize winning discovery of the Higgs Boson. Whilst the observed result ultimately turned out to be a statistical fluctuation, these expressions of greater novelty motivated a deeper investigation of the episode as a case study, including interviews with those who conducted the search and analysis in CMS and ATLAS to probe their reflections on the 750 GeV episode and novelty.

Philosophical treatments of novelty, in discussions of scientific realism, emergence, and scientific progress, have very often focused on the theoretical generation of novelty (see, for example, (Butterfield, 2011; Lakatos, Worrall, & Currie, 1978; Leplin, 1997). In these accounts, experimentation plays no role in the generation of novelty; instead experimentation is relegated to confirmation. The LHC at CERN is one of the largest and most complex experiments ever built, consisting of a 27 km ring in which protons are accelerated and made to collide in bunches of proton proton collisions in four detectors. Each of these detectors was independently built and is run a by large experimental collaboration: the ALICE, ATLAS, CMS, and LHCb experiments. ATLAS and CMS are multi-purpose detectors, originally designed to understand the origin of electroweak symmetry breaking, to search for physics beyond the standard model, and to perform precision measurements of processes within and beyond the standard model (ATLAS, 2003; CMS, 2002). These epistemic goals make the ATLAS and CMS experiments the ideal place to explore novelty in the context of the practices of scientific experimentation. Not only do the epistemic goals of ATLAS and CMS suggest that the LHC experiments aim to be novelty-producing machines, and in their diversity, they also hint at diverse understandings of ‘newness’. This presents an opportunity to examine, from the perspective of high-energy experimental physics, novelty: a concept often used but rarely interrogated.

In this paper, I locate and disambiguate different expressions of novelty found in interviews with experimental physicists. In order to do this, I will consider novelty as a relational concept, i.e. I locate novelty in the interviews where two or more things are connected, are in some way differentiated, and in which the difference is positively valued. This conceptually driven analysis allows for the exploration of the diversity of novelty. Instead of attempting an exhaustive taxonomy of novelty, some of the differences between understandings of novelty, that are significant within the context of the case study, will be explored. I will outline and explore some differing expressions of ontological novelty found in relation to different expressions of the standard model (differing ontologies). Also located are different kinds of novel contributions to the high-energy physics ontology: properties vs entities identified through differing expressions of epistemic practices. Across each of these differing expressions of novelty are differing positive appraisals (in that there are different attributions of value).

I show that the kinds of novelty framed as most valuable are those that violate expectations and are difficult to incorporate into the existing structures of knowledge. In such instances, disruption to the existing ontology or ways of knowing are valued. This positive appraisal of disruption, and contradiction over confirmation, is explored in the recent context of high-energy physics, where several physicists have claimed that there is a lack of promising directions for the future, or even that the field is in a ‘crisis’. I show that the role of disruption explains the differences between the differing notions of novelty. Furthermore, I show that the positive appraisal of disruption is based on forward looking assessments of future fertility, or heuristic appraisal (Nickles, 1989, 2006). Within the context of concerns of a lack of available promising future directions, disruption becomes a generator of alternative futures.

References

ATLAS. (2003). ATLAS high-level trigger, data-acquisition and controls: Technical Design Report: ATLAS-TDR-016.

Butterfield, J. (2011). Less is Different: Emergence and Reduction Reconciled. Foundations of Physics, 41(6), 1065-1135. doi: 10.1007/s10701-010-9516-1

CMS. (2002). Data Acquisition and High-Level Trigger Technical Design Report (Vol. Two).

Lakatos, I., Worrall, J., & Currie, G. (1978). The methodology of scientific research programmes. Cambridge; New York;: Cambridge University Press.

Leplin, J. (1997). A Novel Defense of Scientific Realism. New York: Oxford University Press.

Nickles, T. (1989). Heuristic Appraisal: A proposal. Social Epistemology, 3(3), 175-188. doi: 10.1080/02691728908578530

Nickles, T. (2006). Heuristic Appraisal: Context of discovery or justification? In J. Schickore & F. Steinle (Eds.), Revisiting Discovery and Justification: Historical and philosophical perspectives on the context distinction (Vol. 14, pp. 159-182). Dordrecht: Springer Netherlands.

Proper Functions: Etiology Without Typehood

Geoff Keeling and Niall Paterson
University of Bristol (Keeling), University of Helsinky (Paterson)

Biologists distinguish the functions which a trait is supposed to serve from the functions it has only accidentally. For example, whilst the heart is supposed to pump the blood, but only accidentally makes a thumping sound. Those functions a trait is supposed to serve are called its proper functions. Under what conditions is some function a trait's proper function?

There are two dominant classes of answer to this question. There is agreement between both that the notion of a proper function, at least in a biological context, is closely connected to natural selection. Where they disagree is in whether that connection concerns the trait's history. According to backwards-looking accounts (Millikan 1986; Neander 1991a, 1991b; Godfrey-Smith 1994), the truth conditions for proper functional ascriptions concern facts about an organism's ancestry. Thus on a fairly standard view the proper function of a trait are the effects of previous tokens of the same type, which conferred a selection advantage to the organisms that served as the bearers of those tokens. On this view, proper functions are primarily properties of trait types, and only derivatively of trait tokens. In contrast, proponents of forwards-looking accounts (Pargetter & Bigelow 1987; Nanay 2010) take the truth conditions to concern the trait's modal or dispositional properties. Roughly, they are those dispositional or modal properties that would confer a selection advantage under suitable conditions. On this view, proper functions are primarily properties of trait tokens, not types.

This paper has two central aims. The first is to argue that all existing forwards- and backwards-looking accounts are inadequate. Drawing on recent work by Bence Nanay (2010), we argue that backwards-looking accounts are inescapably viciously circular. We then argue that forwards-looking accounts fail to capture the explanatory power of functional ascriptions in the biological sciences. In particular, we argue that at most forwards looking accounts show that proper functions can be explanatory, not that they are. Drawing on the debate concerning the proper function of the giraffe's long neck, we argue that only the latter is in keeping with actual biological practice.

The second is to outline and defend a novel backwards-looking account of proper functions that takes proper functions to be primarily attributable to trait tokens, and only derivatively to trait types. We call this the token etiological view. Standardly, etiological accounts that apply primarily to tokens have been thought impossible, as selection only acts on trait types. We argue, however, that whilst proper functions do have an important connection to selection, that connection can be adequately understood at the token level in terms of the dual notions of inclusive fitness and comparative similarity alone. Roughly, it is argued that a trait's function is amongst its proper functions just in case there is a previous ancestor of the organism that bears the trait token, such that the ancestor's most similar trait served that function, and by doing so contributed to that organism's inclusive fitness. Since we make no appeal to types, the charge of circularity is avoided. Moreover, we argue that this conception properly accounts for the explanatory role of proper functional ascriptions in the biological sciences. More precisely, where the conditions above are satisfied, an organism's standing in the relevant relation to ancestral traits raises the probability that the organism possesses that trait.

Quantifying Causal Specificity Comes Up Short

Ulrich Stegmann

University of Aberdeen

In recent work, Paul Griffiths and his collaborators have provided an information-theoretic measure of causal specificity (e.g. Griffiths et al. 2015). The quantification is intended to precisify causal specificity, which until then was mostly analyzed in qualitative terms (e.g. Waters 2007, Woodward 2010). Causal specificity refers to the phenomenon that some causes appear to be more specific in their effects than others. For instance, a light dimmer has a high degree of causal specificity because the many positions to which it can be set control a wide range of lighting levels. The dimmer's causal influence is fine-grained. The quantification of causal specificity is regarded as a major advance. It has been advertised, and welcomed, as yielding a more rigorous account of biological specificity, enabling principled comparisons between the specificities of different biological causes, and accounting for several scientific practices, such as the use of informational language and the emphasis on genes over other causes (e.g. Weber 2017). The attention is now shifting to subsidiary issues, e.g. detailed comparisons between types of biological causes and determining the appropriate kind of variation (e.g. actual variation in a given population vs total potential variation).

Underpinning much of this work is the assumption that "biological specificity is simply causal specificity in biological systems" (Griffiths 2016), or at least that it is mostly so. This assumption is misguided, however. Here I argue that (1) causal specificity captures only one kind of biological specificity (out of several distinct phenomena) and that (2) the kind it does capture is not even the source of the scientific practices that proponents claim to explain. Therefore, quantifying causal specificity comes up short if, as professed, the goal is to understand biological specificity.

I will support this conclusion by highlighting two types of causal relations. In one, a given cause is responsible for only one effect, or very few. This relation was at stake in Beadle's (1945) notion of "gene specificity" as well as Pauling's (1956) comparison between the specificity of genes and enzymes. For scientists like Beadle and Pauling, the specificity of a cause (like a gene or an enzyme) increased to the extent it had *fewer* effects. This kind of biological specificity is therefore not the causal specificity that has been quantified, which increases with more possible effects. I suggest, instead, that this form of biological specificity exemplifies Woodward's (2010) "one cause-one effect notion" of specificity. This notion has been overlooked in the excitement about fine-grained specificity, its popular twin.

The second type of causal relation resides at the level of several cause-effect pairs, rather than at the level of only one such pair. In linear macromolecules like proteins, each subunit can be considered as an effect variable that can take different values. As Stegmann (2014) argued, the cause variables can form a separate entity or simply be identical with the chain of effect variables (since a given effect may be the cause for a subsequent effect). This is a difference with respect to how several cause-effect pairs hang together, and it is independent of the degree of fine-grained specificity within individual cause-effect pairs. Here, I will argue that the biological specificity of molecular templates is due to how the cause-effect pairs interrelate and is, therefore, not captured by causal specificity.

In sum, quantifying causal specificity is of limited value when accounting for biological specificity. However, the philosophical literature already provides additional conceptual resources for a broader and more nuanced understanding of biological specificity. Future work in this area should focus on integrating these different aspects.

References

Beadle, George W. (1945). "Biochemical Genetics". Chemical Reviews, 37: 15-96

Griffiths, Paul. (2016). “Proximate and Ultimate Information in Biology.” In The Philosophy of Philip Kitcher, edited by Mark Couch and Jessica Pfeifer, 74–91. New York and Oxford: Oxford University Press

Griffiths, Paul, Arnaud Pocheville, Brett Calcott, Karola Stotz, Hyunju Kim, and Robert Knight. (2015). “Measuring causal specificity”. Philosophy of Science, 82: 529-555

Pauling, Linus. (1956). "The Future of Enzyme Research". In Enzymes: Units of Biological Structure and Function, edited by Oliver H Gaebler. New York, Academic Press

Stegmann, Ulrich. (2014). “Causal control and genetic causation”. Nous, 48: 450-464

Waters, C. Kenneth. (2007). "Causes that make a difference". Journal of Philosophy, 104, 551-579

Weber, Marcel. (2017). “Which kind of causal specificity matters biologically?” Philosophy of Science, 84: 574-585

Woodward, James. (2010). "Causation in biology: stability, specificity, and the choice of levels of explanation". Biology and Philosophy, 25, 287-318

Quantisation as a Method of Discovery

Niels Linnemann
University of Geneva

Discussions of philosophers of physics on quantizing gravity have so far largely focused on the question in how far general relativity needs to be (or rather should be) quantised: Huggett and Callender (2001) as well as Wüthrich (2005) for instance consider the prospects of a mere semi-classical theory in which the gravitational sector itself stays classical (albeit perhaps subject to slight corrections). In this talk I will take the need for a quantisation of gravity as given. So, rather than asking whether quantisation is necessary, I would like to raise the consequent question in line, namely of the prospects of such a procedure: given that quantisation is a highly ambiguous mapping from a classical to a quantum theory, why think that quantisation can be a sensible rationale for the theory change from general relativity (GR) to quantum gravity (QG) at all? In the first half of my talk, I will work out that we in fact face a genuine challenge here, and argue that we can only address it through the imposition of well-chosen principles. Quantisation is an ideal context to demonstrate the usefulness of the idea of principles for theory construction.

Now, under the question for the prospects of a specific quantisation project such as that of gravity, lies an arguably even deeper and more general one: how to think of quantisation per se? What kind of procedure is quantisation? What should we make of its apparent success? I will address these issues in the second part of my talk. For this, I will flesh out the often heard claim that quantisation is a form of recipe for translating between two theoretical frameworks, namely a classical and a quantum one. After comparing quantisation to other forms of recipes of theory change to be found in the evolution of physics, I will argue that theory changes in physics fall into two categories, namely (1) theory changes realizable through prescriptions of contents from the old theory's framework into the successor theory's framework, and (2) those which do not admit such a prescription (or, for that matter, at most only highly artificial ones). I will then point out how this distinction between theory changes has consequences for quantum gravity research.

I end this abstract with a short overview on the concrete conceptual insights to be expected from dealing with quantisation both for quantum gravity and more generally - to emphasize that there is indeed more involved in quantisation than just technical issues:

I. A theory change via quantisation suffers from ambiguities at several stages, namely at the level of (1) choosing the appropriate operator algebra, of (2) giving it an unitary representation, and of (3) ordering the quantum operators. Usefulness of quantisation as a method of discovery can however be enhanced through principles of different kinds (geometric, topological, physical) some of which even allow for getting rid of these ambiguities completely. Quantisation approaches thereby also provide an ideal background for concretely displaying how principles play a role in theory change (if not clear already).

II. A second issue is in how far quantisation qua being a prescription (or any other prescription for this matter) is special to physics: I suggest that it is not but that at the same time this aspect about quantisation cannot be taken for granted either. One should rather distinguish between theory changes, those involving and those not involving notable prescription schemes for carrying over content from the old to the new theory. Quantisation as well as the (minimal) coupling prescription for translating matter content from special relativity to general relativity are examples for the first kind of theory change, the theory change from thermodynamics to statistical mechanics an example for the second kind.

III. A third issue is (building on the second issue) how to `conceptualise' quantisation (prescriptions) in the end. I claim that quantisation (and prescriptions more generally) plays a threefold role as a method of generation, (weak) generative justification and as an intertheory relation.

IV. A fourth issue is in how far the term `theory change' is still adequate for denoting prescriptional changes. From the perspective of a quantizer, it might seem as if the actual theory change (or at least, the bigger chunk of it) has already occurred with the set-up of a general quantum framework including the notion of a quantisation map (which we are by now acquainted with from non-relativistic as well as relativistic quantum (field) theory). What is left, is to apply this quantisation scheme (or, more precisely, some technical adaptation of it) but not more. To speak in the language of Kuhn, the quantisation of gravity might still be nothing but puzzle-solving (Cf. Feintzeig (2017)).

Quinean Realism and a New Defence of Antirealism

Arthur Harris
University of Cambridge

Quine's broadly Humean argument in ‘Posits and Reality’ (1966 in The Ways of Paradox) aims to show the how our ontological commitments relate to the evidence. Quine argues that, in some sense, the unobservable posits of scientific theory are on an epistemic par with everyday objects: both are underdetermined by the only evidence available, sensory stimulation. So if we have good reason to accept that everyday objects like tables and chairs are real, we should also accept electrons as real. This powerful argument has received insufficient attention in debates concerning scientific realism. First I distinguish two forms which the argument may take, both constituting serious attacks on the tenability of constructive empiricism. On the one hand, the Quinean argument may be construed in terms of what Wright has called ‘Group II’ propositions. In this case Quine’s argument can be made to push the antirealist towards implausible sceptical denial of particular everyday objects like tables and chairs. On the other hand, Quine's argument may be construed in terms of 'Group III' propositions, in which case it pushes the antirealist towards incoherent denial of the external world. Van Fraassen has not responded to this argument directly, but I argue that what comments he has made on Quine are unsatisfactory.

I propose a response to Quine on behalf of the antirealist, although it involves dropping van Fraassen's insistence that philosophy of language is irrelevant to the question of scientific realism. For it is partly in Wittgenstein's On Certainty that compelling responses to Quine may be found. Hinge epistemology dismantles the Quinean argument's 'Group III' variant. The other part of my suggested response is supplied by epistemological disjunctivism, which disarms the Quinean argument's 'Group II' variant. One advantage to this new defence of scientific antirealism is that it offers a less question-begging response to the radical sceptic than Quine's naturalised epistemology.

Relations in the Metaphysics of Science

Stavros Ioannidis, Elina Pechlivanidi and Stathis Psillos
University of Athens

Relations in the Metaphysics of Science

Although the ontic status of relations has recently drawn some attention within metaphysics (cf. Heil 2012, Simons 2010), it has not drawn too much attention within metaphysics of science. This paper will examine the status of relations in two prominent views in metaphysics science, dispositionalism and structuralism, and it will sketch an alternative to both. The key question is how exactly do relations enter the world: are relations dependent on the properties of the relata, or are they independent ontic features of reality? We show that dispositionalism and structuralism lead to opposite but equally controversial accounts of the metaphysics of relations. We argue that an alternative account is possible, which addresses the shortcomings of both of these views when it comes to their metaphysics of relations, and best combines metaphysical rigour with a naturalistic stance.

Currently, there are two main views about the metaphysics of relations. The first view is that there exist both internal and external relations. Internal relations depend on their relata in the following sense: if we fix the properties of the relata, we thereby fix the relations they enter into. But although such a view is plausible regarding, for example, comparative relations ('a is bigger than b'), causal and spatiotemporal relations seem to provide counterexamples. Hence, many philosophers think that such relations are external to their relata, i.e. they are not fixed by the relata and their properties. Issues of conceptual economy, however, have led several metaphysicians to think that all relations are internal, and hence that prima facie external relations, such as causal and spatiotemporal ones, should also be understood as supervenient on or reducible to the properties of the relata. But these views require further ontological commitments. For example, a typical move in this debate is to adopt a theory of causation in terms of the manifestation of powers, which is supposed to make causal relations internal (Heil 2012, 148).

How does this debate play out in the metaphysics of science? Two currently prominent views in the metaphysics of science, dispositionalism and structuralism, seem to take opposing views on the status of relations. Dispositionalists hold that properties are powerful and their causal role is essential to them; by building on the intrinsic powers of things, they attempt to offer a unified ontology of the fundamental structure of the world. Laws on this view are said to supervene on powers (Bird 2007) or to be eliminated altogether (Mumford 2004). In light of this it appears that all relations that particulars enter into in virtue of their powers are fixed by these powers; hence all relations are internal to the relata. We identify three key problems for the dispositional account. First, it is not clear how spatiotemporal relations can be viewed as internal; indeed, less extreme dispositionalists, e.g. Ellis (2001), allow for spatiotemporal relations to be external to the relata. This means that, at best, we 've got a mixed view which allows internal as well as external relations. Second, not all relations in science seem to be the result of powers, e.g. relations embodied in symmetries and conservation laws. Third, dispositionalists do not seem to have the resources to account for the quantitative relations among powers, as they are specified in laws of nature (e.g. in Coulomb's law).

Ontic structuralism, on the other hand, does away with 'metaphysically robust or 'thick' objects' (French 2014, 64) and posits just relations and structures. Structuralists 'begin with the symmetries and laws, from which the relevant properties effectively 'drop out'' (French 2018). So, they consider relations to be fundamental, arguing that such a metaphysical picture is more consonant with our best scientific theories. The resulting picture is in stark contrast to dispositionalism; by taking relations as primary, ontic structuralism aims to recover the properties from them. Thus, fundamental relations are external to the properties, since properties depend for their existence on the existence of relations. We identify three key problems for the structuralist account. First, it is not at all clear how structuralists can recover properties from relations. Second, laws for structuralists are taken to relate determinable quantities; but then it is not clear how the specific quantities that there are in the world are related to each other in such a way that they fall under the law that relates determinable quantities. Third, what do relations relate for structuralists?

Dispositionalism and structuralism lead thus to two opposing views concerning the metaphysics of relations. In both cases the resulting picture is characterised by conceptual economy: for dispositionalists, fundamental relations are internal; for structuralists, they are external. But this shared monistic attitude is grounded in different motivations. Dispositionalists start from a preferred metaphysics and attempt to use it to account for the relational features uncovered by science; they adopt a metaphysics-first top-down approach. Structuralists start from science, and attempt to read off metaphysics from our best scientific theories; they adopt a science-first bottom-up view. Both approaches have problems when it comes to the metaphysics of relations. But there is an alternative: both properties and relations may by needed to offer a metaphysical picture consonant with the content of laws as revealed by science. Although this account is less parsimonious concerning the metaphysics of relations, it best combines two important methodological virtues in metaphysics of science: a naturalistic stance with a rigorous attitude towards fundamental metaphysics.

References

Bird, A. (2007) Nature's Metaphysics: Laws and Properties, Oxford: Oxford University Press.

Ellis, B. (2001) Scientific Essentialism, Cambridge: Cambridge University Press.

French, S. (2014) The Structure of the World: Metaphysics and Representation, Oxford: Oxford University Press.

--- (2018) 'Defending Eliminative Structuralism and a Whole Lot More (or Less)', Studies in History and Philosophy of Science, https://doi.org/10.1016/j.shpsa.2018.12.007.

Heil, J. (2012) The Universe As We Find It, Oxford: Clarendon Press.

Mumford, S. (2004) Laws in Nature, London: Routledge.

Simons, P. (2010) 'Relations and Truthmaking', Proceedings of the Aristotelian Society Supplementary Volume 84: 199-213.

Revisiting Abstraction and Idealization in Molecular Biology

Martin Zach
Charles University

The philosophical debates over the last several decades have made abundantly clear that much of scientific practice relies on models that, in some sense, are simplified versions of their target systems. Philosophers of science often distinguish between abstraction and idealization, both of which contribute to the various ways in which scientists simplify.

In this paper I argue for two things. First, the popular and widely used notions of abstraction and idealization face numerous issues with regard to their characterization. I provide a survey of the literature (e.g., Godfrey-Smith [2009]; Frigg and Hartmann [2012]; Levy and Bechtel [2013] etc.) to show that, few notable exceptions notwithstanding, there currently exists a "standard view" according to which, roughly speaking, an idealization concerns a distortion of a feature in a model, and as such it introduces a falsehood into the model, whereas an abstraction concerns an omission of an (irrelevant) feature. Second, the lack of conceptual clarity with respect to these notions poisons various other debates including the one on mechanistic explanation.

Regarding the first problem, I argue that the standard view fails to provide an adequate characterization of either of the notions. Number of issues beset the individual concepts as well as the distinction between them. Here, I list only a couple. For instance, not every distortion counts as an idealization, nor does just any falsehood (see Levy [2018] for a related idea). Abstraction, understood as a procedure by which one subtracts individual features from the target system, is akin to rational reconstruction rather than a description of the model-building process. This is because scientists often do not know what features the target system has, and the point of modeling is to find out precisely that (see Portides [2018]). It also proves rather difficult to apply the standard view to actual scientific cases. Indeed, authors often disagree about whether a particular assumption counts as a case of abstraction or idealization. Hence, in practice the distinction is often blurred. Furthermore, it is often the case that an important distinction between the nature of simplifying assumptions and the types of functions these assumptions serve is conflated, as seen in many instances (e.g., in Wimsatt [2007]; Rohwer and Rice [2013]). All this shows that the standard view is based on a confusion, and more importantly, it introduces additional confusion into other debates in which the notions of abstraction and idealization play key roles (see below).

The second problem concerns the implications this conceptual confusion of the standard view has for various other debates. Depending on the context, a rough characterization of these notions may very well be good enough. However, the standard view is often discussed in context in which the precise characterization of the notions of abstraction and idealization is key. Here, I argue that some of the recent attempts to challenge the framework of the new mechanistic account of explanation is wrong-headed precisely because it relies on the standard view. In a recent paper, Love and Nathan ([2015])

argue that “accounts of mechanistic explanation face a problem in accommodating the deliberate misrepresentation [i.e. idealization] of causal relations among components and activities that play a difference-making role in producing the explanandum” (p. 770). They discuss the case of modeling gene transcription and argue that scientists commonly introduce three misrepresentations into their models (i.e. treating molecular complexes as if they were single molecules, disregarding the fuzzy nature of the process in which various molecules constantly bind and detach, and omitting the role of concentrations). Love and Nathan take these assumptions to be idealizations, using the standard view characterization. However, there are several issues with their approach. In accord with the standard view they define abstraction as “the intentional omission of detail” (p. 763), yet they claim that when “known difference makers are intentionally omitted from the representation” (p. 767) we are to understand it as an act of idealization. Thus, they seem to conflate abstraction with idealization. Using their own example, I further argue that abstraction cannot be distinguished from idealization by characterizing the latter as an introduction of a distortion, since abstraction-as-omission-of-detail can, and often does, distort features as well.

Importantly, the approach and analysis of Love and Nathan has explicitly been embraced by several authors in the debate on mechanisms(e.g., van Eck and Mennes [2016]; Rice [2017]; Halina [2018]). This introduces a dangerous precedent, one that could spark a long-lasting debate without realizing that it builds on a wrong footing.

Frigg, R. and Hartmann, S. [2012]: ‘Models in Science’, in E. N. Zalta (ed.), The Stanford Encyclopedia of Philosophy (Summer 2018 Edition),.

Godfrey-Smith, P. [2009]: ‘Abstractions, Idealizations, and Evolutionary Biology’, in A. Barberousse, M. Morange and T. Pradeu (eds), Mapping the Future of Biology, Dordrecht: Springer Netherlands, pp. 47–56.

Halina, M. [2018]: ‘Mechanistic Explanation and Its Limits’, in S. Glennan and P. Illari (eds), The Routledge Handbook of Mechanisms and Mechanical Philosophy, New York: Routledge, pp. 213–24.

Levy, A. and Bechtel, W. [2013]: ‘Abstraction and the Organization of Mechanisms’, Philosophy of Science, 80, pp. 241–61.

Levy, A. [2018]: ‘Idealization and Abstraction: Refining the Distinction’, Synthese, pp. 1–18.

Love, A. C. and Nathan, M. J. [2015]: ‘The Idealization of Causation in Mechanistic Explanation’, Philosophy of Science, 82, pp. 761–74.

Portides, D. [2018]: ‘Idealization and Abstraction in Scientific Modeling’, Synthese, pp. 1–23.

Rice, C. [2017]: ‘Models Don’t Decompose That Way: A Holistic View of Idealized Models’, The British Journal for the Philosophy of Science,.

Rohwer, Y. and Rice, C. [2013]: ‘Hypothetical Pattern Idealization and Explanatory Models’, Philosophy of Science, 80, pp. 334–55.

van Eck, D. and Mennes, J. [2016]: ‘Design Explanation and Idealization’, Erkenntnis, 81, pp. 1051–71.

Wimsatt, W. C. [2007]: ‘Re-Engineering Philosophy Fo Limited Beings’, Cambridge (Mass.): Harvard University Press.

Salience and the Sure-Thing Principle

Chloé de Canson
London School of Economics

This paper is about the Sure-Thing Principle, which originates with Savage. Savage introduces this principle with the help of the following case: `A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant. So, to clarify the matter to himself, he asks whether he would buy if he knew that the Democratic candidate were going to win, and decides that he would. Similarly, he considers whether he would buy if he knew that the Republican candidate were going to win, and again finds that he would. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains' (p. 21). The intuition behind this case can be captured by the following principle. Where O={w1, w2, ...} is the set of all the states an agent considers possible, or sample space:

Unrestricted Sure-Thing Principle (STPu). Let P={P1, P2, ...} form a partition of O. If, for every proposition Pi in P, an agent would perform an action T if she were to learn Pi, then the agent ought to perform T.

Two decision-theoretic paradoxes have been presented as counterexamples to the (STPu): the Two-Envelope Paradox, and Allais' Paradox. My aim in this paper is to isolate the feature of the (STPu) which generates these paradoxes, and to formulate a restricted version of the Sure-Thing Principle, which both captures the intuition behind Savage's businessman case, and avoids the paradoxical conclusions of Allais and the Two Envelopes. Because of time limitations, I will not be able to talk in detail about both paradoxes: I therefore concentrate on the Paradox of the Two Envelopes. My analysis carries over straightforwardly to Allais' Paradox: I will explain briefly how if time permits.

The upshot of my analysis of the Sure-Thing Principle is surprising: the Sure-Thing Principle is relative to the way in which possibilities are described, in a way I make precise. This has important implications for decision theory: it implies that there is an additional layer of description-relativism in the individuation of decision-theoretic situations, beyond standard intensionality. More precisely, it implies that fixing a sample space of possibilities, and a credence function and a utility function on this space, is not sufficient to yield a recommendation for action. One must also specify what I call a salient feature of the situation, which in turn allows us to pick out a privileged partition of the sample space. Only once this salient feature (and therefore partition) has been specified, can we get a recommendation.

I proceed in the following way.

(1) I begin by showing how the paradox of the two envelopes is a counterexample to the (STPu). In a nutshell, the (STPu) applied to the decision situation of the paradox of the two envelopes yields two inconsistent recommendations for action. Two instances of the same reasoning, applied to the same case, yield contradictory recommendations.

(2) Then, I argue that a refinement of the (STPu) proposed by Dietrich and List is unsatisfactory. Although this refinement does avoid the paradoxical conclusion in the case of the two envelopes, it yields very unintuitive results in two cases I describe.

(3) Then, I study the two cases I levelled against Dietrich and List in more detail. I ask what a Sure-Thing Principle needs to look like, if it is going to adjudicate these cases correctly. The answer, it turns out, appeals to a notion I call "salience". Salience is the notion involved in Goodman's discussions of what he calls the "new riddle of induction". In a nutshell, the idea is this. When one describes an

object, there are several ways that one can partition logical space, and therefore several true descriptions that one can give of that object. Suppose for instance that one wants to describe the appearance of an emerald. If one partitions logical space as green-observed/green-unobserved and blue-observed/blue unobserved, then the emerald is (truly) described as green. If by contrast one partitions logical space as green-observed/blue-unobserved and blue-observed/green unobserved, then the emerald is (truly) described as grue. The emerald is thus both green and grue. However, for the purposes of reasoning (in Goodman's case, about unobserved emeralds), one of these two descriptions must be retained. Our intuitions tell us that "green" is the predicate to be retained---in my terminology, the emerald's greenness is its salient feature.

(4) The introduction of salience in the previous section leads to a refinement of the (STPu), which I show solves the paradox of the two envelopes. The newly-formulated Sure-Thing Principle, the (STPsa), is restricted to salient partitions of O---where the (STPu) was unrestricted in its applications to partitions. I show that it yields no paradox and all the intuitively correct judgements in all the cases mentioned in the paper.

(5) Finally, I examine the costs and benefits of adopting my position. My main conclusion is that, even though the description-relativism of the Sure-Thing Principle for which I advocate is a radical departure from orthodox decision theory, the costs of rejecting this restricted version are too heavy to bear. Rejecting my proposal implies that decision theory must be silent on a wide range of cases, including those of the paradox of the two envelopes and the Allais paradox.

Socially Extended Scientific Understanding

Harry Lewendon-Evans
Durham University

Over the past decade, a significant body of work has been developed within epistemology and philosophy of science that examines the nature of scientific understanding, and its relation to the vehicles of scientific inquiry, such as models, explanations, and thought experiments. These analyses of scientific understanding have largely followed traditional accounts of knowledge by primarily focusing on the necessary and/or sufficient conditions required for an individual to possess scientific understanding, whether through the possession of the right kind cognitive abilities, representational states, explanatory information or connections between beliefs. However, this individualism is decidedly at odds with the social dimensions of scientific enquiry. In the era of 'Big Science', scientific research is typically spread out over multiple research groups and teams, distributed on a massive scale, and fundamentally collaborative in nature. In this context, no individual researcher has the evidence required to justify particular claims to understanding; instead numerous researchers with specific types and levels of expertise are required to achieve understanding. What does this mean for the concept of scientific understanding? As yet, insufficient attention has been paid to the collective and collaborative nature of scientific research and its implications for our understanding of scientific understanding.

This paper offers a new perspective on scientific understanding by arguing that scientific understanding is fundamentally social in nature. To support this claim, the paper draws on the conceptual framework of extended theories of cognition, and in particular distributed theories of cognition, to provide a theoretical model that captures the social dimensions of scientific understanding. While some scholars have already proposed the application of extended cognition theory to accounts of scientific understanding (e.g. Toon 2015), this work has remained largely limited to the analysis of the use of tools and material artefacts. In contrast, I argue that an adequate account of scientific understanding also needs to accommodate the way in which scientific understanding is socially extended. In order to demonstrate this, I build upon Andrea Woody's (2015) recent work on the function of explanatory discourse in scientific practice and argue that such an account serves to illustrate the way in which a socially extended conception of scientific understanding can accommodate the widely held view that understanding and explanation are closely connected.

This paper challenges a widely held assumption in the current literature on scientific understanding, namely that scientific understanding is adequately analysed at the level of the individual subject. By proposing a socially extended conception of scientific understanding, this paper seeks to shed new light on the social nature of understanding, its relation to scientific explanation, and the collective, collaborative and distributed nature of scientific research.

Spacetime Emergence and Functional Realization

Baptiste Le Bihan
University of Geneva

Some approaches in contemporary physics entail that space (or spacetime) emerges from a structure in which interesting features of space and time are missing. This is the case in many approaches to quantum gravity (for instance in string theory and loop quantum gravity), but also in a particular approach to quantum mechanics (configuration space realism). A general issue is then to understand the nature of this relation of emergence and to determine what sort of ontological picture follows from spacetime emergence. Lam and Wüthrich (forthcoming) have recently suggested---in the context of quantum gravity---to identify this relation of emergence or, as I prefer to call it, of constitution with a relation of functional realization---thereby drawing inspiration from the philosophy of mind where functional realization is a popular way to analyse the relation obtaining between physical states and mental states.

Spacetime is regarded as potentially emerging from, or being constituted by, a non-spatio-temporal structure in various approaches to QG, to various degrees. At first glance, it may mean several things depending on whether space, time or spacetime, comes under attack. As it has been argued by Le Bihan and Linnemann (forthcoming), if one defines the existence of a minimal spacetime as the existence of a local split between two structures---"space" or "quasi-space" on the one hand, and "time" or "quasi-time" on the other hand---then we find such a distinction implemented in most approaches to quantum gravity, either with Lorentz symmetries or through another diachronic principle. However, and as suggested by the expression "quasi-space" and "quasi-time", this is not to say that no interesting features usually ascribed to space and time are missing in quantum gravity. (For a general review, cf. Huggett and Wüthrich (2013).)

In string theory for instance, the 4D spacetime emerges, prima facie, from a 10D structure. According to a naive understanding there is no problem of spacetime constitution in this context since the additional dimensions are compactified: it is simply that we fail to notice them when we zoom out. But there are five different dual 10D string theories, and models of these theories are empirically equivalent. Some of those models are related by T-duality and possess different compactification radius. As a result, the network of dual theories has to be related to a more fundamental M-theory, which still has to be found, and to GR. Therefore, the classic story about the compactification of unobserved dimensions does not explain away the emergence of GR spacetime. And importantly for our purpose, as demonstrated by Huggett (2017) the target spaces on which the strings live cannot be identified with GR spacetime, leading to a problem of spacetime emergence.

In this paper, I examine how exactly functional realization may help us understand situations of spacetime emergence in quantum gravity. More precisely, I examine which problems can be solved by functionalism. Indeed, Lam and Wüthrich (forthcoming) focus on a particular epistemological issue--namely, the problem of empirical (in)coherence (introduced by Huggett and Wühtrich (2013)): How are we going to justify a theory which threatens its own evidence which, arguably, is located in space and time? But they also take their functionalist strategy to have a broader application since, according to them, functionalism might, in principle, do more than just solve the problem of empirical coherence by also providing an answer to the more general philosophical issue of accounting for the ``metaphysical gap'' obtaining between a spatio-temporal theory (General Relativity, GR for short) and a non-spatio-temporal theory (of quantum gravity)---a problem that amounts, in the context of Quantum Gravity, to asking, first, whether GR spacetime does (not fundamentally) exist or, on the contrary, does not exist and, second, about how we may connect a more fundamental non-spatio-temporal structure with a less fundamental spatio-temporal structure.

During the talk, I will defend the three following claims:

1*) functionalism may be regarded as a view orthogonal to the metaphysical problem of spacetime when subscribing to a particular sort of functionalism,

2*) functionalism may be regarded as a particular solution to the metaphysical problem of spacetime when subscribing to another particular sort of functionalism and,

2**) functionalism helps to solve some---but not all---problems of spacetime emergence.

Bibliography

Crowther, K. (2016). Eective Spacetime. Springer.

Huggett, N. (2017). Target space6= space. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 59, 81-88.

Huggett, N. and C. Wüthrich (2013). Emergent spacetime and empirical

(in) coherence. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 44 (3), 276-285.

Knox, E. (2013). Eective spacetime geometry. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 44 (3), 346-356.

Knox, E. (2017). Physical relativity from a functionalist perspective. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.

Lam, V. and C. Wüthrich (forthcoming). Spacetime is as spacetime does. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.

Le Bihan, B. and N. Linnemann (forthcoming). Have we lost spacetime on the way? narrowing the gap between general relativity and quantum gravity. Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.

Maudlin, T. W. (2007). Completeness, supervenience and ontology. Journal of Physics A: Mathematical and Theoretical 40 (12), 3151.

Monton, B. (2002). Wave function ontology. Synthese 130 (2), 265-277.

Monton, B. (2006). Quantum mechanics and 3 n-dimensional space. Philosophy of science 73 (5), 778-789.

Ney, A. and D. Z. Albert (2013). The wave function: Essays on the metaphysics of quantum mechanics. Oxford University Press.

Wüthrich, C. (2017). Raiders of the lost spacetime. In Towards a theory of spacetime theories, pp. 297-335. Springer.

Yates, D. (forthcoming). Thinking about spacetime. In C. Wüthrich, B. Le Bihan, and N. Huggett (Eds.), Philosophy Beyond Spacetime. Oxford University Press.

Stability and the Looping Effects of Human Kinds

Riana Betzler
University of Cambridge

How to distinguish natural kinds—and whether there exists a substantive difference between natural and social or human kinds—is a longstanding problem in the philosophy of science. Recently, there has been a trend of thinking of natural kinds as grounded in scientific practice; these kinds of accounts place emphasis on the epistemic value of attributing natural kindhood to entities. They capture what "natural kind" is supposed to be doing—that is, enabling reliable inference. In this paper, I focus on one such account—Matthew Slater's (2015) Stability Property Cluster (SPC) account of natural kindness.

This account has many virtues. It foregrounds—rightly, I think—the issue of stability as central to natural kind ascriptions. It avoids widely-acknowledged problems with traditional essentialist accounts, thereby framing itself as ripe for use within the life sciences. It also advertises itself as being more flexible than Boyd's (1990, 1991) Homeostatic Property Cluster (HPC) account—which has enjoyed a kind of orthodoxy in the life sciences—insofar as it is able to find stability without reference to mechanisms. It is also domain-relative, or able to be tuned to the specific requirements of different areas of inquiry. While this does not provide us with a universal notion of "natural kind," it does help to account for the plurality of natural kind concepts currently existing within the field; it does so by explicitly acknowledging that there are specific aims, interests, and norms being captured by the use of "natural kind" across different disciplines. These features, I argue, indicate that the account holds promise.

I ask, however, how it fares within the human and social sciences, especially in the face of Ian Hacking's (1995) "looping effects." Hacking's central worry about looping effects is that they are destabilising; because the targets of the human sciences change in response to classification, they are "moving targets." This undermines the potential for stable knowledge about them. Given that stability is the central feature of Slater's account, it seems as though the existence of such looping effects within the human sciences would preclude its applicability there. If human kinds truly are "moving targets" in Hacking's sense, they cannot count as natural kinds for Slater. I suggest that this is a potential pitfall of the account, since there are good reasons for resisting a sharp division between natural and human kinds (see e.g., Cooper, 2004). I then go on to consider ways in which we might deal with such looping effects.

First, I suggest that there is a way of thinking about looping effects such that they can be stabilising rather than destabilising. They can create kinds about which we can make reliable inferences, even if those inferences do not hold indefinitely (this situation holds in the biological sciences as well). Cases where stabilising looping effects occur pose no problems for the SPC account and indicate that it can be fruitfully applied within at least some areas of the human sciences. Mental disorders, I argue, provide good examples of SPC kinds; SPC kinds are roughly what the DSM is tracking.

I then consider cases where looping effects truly are destabilising. I argue that those cases should give us pause. Slater's SPC account shows us why. We should avoid making natural kind ascriptions in the face of destabilizing looping effects, but not for the reasons Hacking suggests. Looping effects do not mark a distinction between "natural" and "human" or "indifferent" and "interactive" kinds. Awareness or reflectivity, to which much attention has been given in discussions of looping effects, is not the crux of the issue. Looping effects are also not uniquely present within the human sciences—although instances in which they are generated by awareness or reflectivity might be. Rather, destabilizing looping effects should give us pause because they do not afford us reliable inferences. Things that undergo destabilising looping effects are not natural kinds precisely because they are unstable. In this way, Slater's account, by focusing on stability, gets it right. It helps us to sort out those kinds which are most relevant to our epistemic practices and those kinds about which we should be more tentative.

References:

Cooper, R. (2004). Why Hacking is Wrong about Human Kinds. British Journal for the Philosophy of Science, 55(1), 73-85.

Hacking, I. (1995). The Looping Effects of Human Kinds. In D. Sperber, D. Premack, & A. J. Premack (eds.), Causal Cognition: A Multi-Disciplinary Debate. Oxford: Clarendon Press, pp. 351-383.

Slater, M. (2015). Natural Kindness. British Journal for the Philosophy of Science, 66, 375-411.

Sufficiently Real? a Critical Review of the Theorems by Colbeck and Renner

Ronnie Hermens
Utrecht University

In a series of papers Colbeck and Renner have proven several results concerning the status of the quantum state [CR11,CR12,CR15a,CR15b,CR17]. Following their review paper [CR15a], their two main results, with some paraphrasing, may be stated as follows:

Claim 1: In any theory that is compatible with quantum mechanics and that satisfies the Free Choice assumption, the quantum state provides a sufficient description of the system.

Claim 2: In any theory that is compatible with quantum mechanics and that satisfies the Free Choice assumption, the quantum state is a necessary component of the description of the system.

These are strong conclusions and therefore the arguments leading up to them deserve careful analysis. Unfortunately, the arguments by Colbeck and Renner are not abundantly clear, as evidenced by the existence of a FAQ on the first paper [CR10], as well as an addendum on the Free Choice assumption [CR13]. Moreover, there have been noticeable attempts at clarifying/correcting/criticizing the work of Colbeck and Renner by Leifer [Lei14] on Claim 2 and Landsman [Lan15] and Leegwater [Lee16] on Claim 1.

In my talk I will not be giving a tedious in-depth technical analysis of all these theorems and their proofs. Instead, my focus is on understanding this current peculiar state of affairs. What is at stake? What has been achieved and what not? And what are the difficulties in making progress? These are the kind of questions I shall be dealing with.

The natural place to start is with the main assumptions to the claims. A Free Choice assumption, although common in the foundations of quantum mechanics, should be handled with care. But, as I will explain in the talk, this is not the main difficulty in analyzing the work of Colbeck and Renner. Instead, the main difficulty is with the notion of a theory being compatible with quantum mechanics.

If interpreted strongly, it may not be surprising that a theory that is “compatible with quantum mechanics” should use quantum states as necessary and sufficient descriptions of systems. Ideally, something weaker is intended, such as “ability to reproduce the predictions of quantum mechanics”. In fact, the truth lies somewhere in the middle. Most notably, what is assumed on top of empirical equivalence (on a selected domain of application), is that the theory “behaves like quantum mechanics” when one considers the possibility of coupling two or more systems and when one considers possible dynamics for the theory.

I shall argue that Leifer has presented a satisfactory reworking of Claim 2 that deals properly with these issues of compatibility with quantum mechanics. Leegwater’s reworking of Claim 1, on the other hand, shares some of the shortcomings of the original work of Colbeck and Renner, some of which also noted by Landsman. (Here my assessment differs from the one by Butterfield in [But18]).

The asymmetry in this assessment concerning claims 1 and 2 has nothing to do with the quality of the of the work on these claims. Rather, it is a reflection of the asymmetry in the nature of these two claims. The necessity of the quantum state can be understood as an ontological claim: necessity suggests a corresponding property in the system. This is the strategy Leifer adopts when reformulating Claim 2 as a ψ-onotlogy theorem (akin to the PBR theorem [PBR12]). The sufficiency of the quantum state can be understood as an epistemic claim: it poses a limit on the kind of knowledge one may obtain about the system.

The asymmetry in the two claims in itself is not problematic: there is no a priori reason why Claim 1 should be more difficult to prove than Claim 2. It only becomes problematic in light of the assumption on the compatibility with quantum mechanics. In particular, one may consider the assumption that the theory mimics quantum mechanics in a particular way with respect to its dynamics. For Claim 2 this assumption has to be reformulated to an assumption on how the so-called ontic state of the system evolves while for Claim 1 it has to reformulated in terms of a constraint on the possible updates of information about the system in light of a change of state of the system. How this exactly yields problems for the proof strategy for Claim 1 will be explained in the talk.

Apart from this negative claim I will end with a positive note. That the Colbeck-Renner strategy for the proof of Claim 1 is not entirely satisfactory of course does not imply that Claim 1 (or some proper reformulation thereof) is false. Pending such a satisfactory proof, I will put forward an argument (part of joint work) showing that the Free Choice assumption for Claim 1 may be relaxed. The argument rests again on the idea that Claim 1 is an epistemic one, and I shall explain why it does not work for Claim 2.

[But18] J. Butterfield 2018, https://doi.org/10.1007/978-981-13-2487-1_11

[CR10] R. Colbeck and R. Renner 2010, http://perimeterinstitute.ca/personal/rcolbeck/FAQ.html

[CR11] R. Colbeck and R. Renner 2011, https://doi.org/10.1038/ncomms1416

[CR12] R. Colbeck and R. Renner 2012,https://doi.org/10.1103/PhysRevLett.108.150402

[CR13] R. Colbeck and R. Renner 2013, arXiv: 1302.4446 [quant-ph]

[CR15a] R. Colbeck and R. Renner 2015, https://doi.org/10.1007/978-3-662-46422-9

[CR15b] R. Colbeck and R. Renner 2015, https://doi.org/10.1007/978-94-017-7303-4

[CR17] R. Colbeck and R. Renner 2017, https://doi.org/10.1088/1367-2630/aab328

[Lan15] K. Landsman 2015, https://doi.org/10.1063/1.4936556

[Lee16] G. Leegwater 2016, https://doi.org/10.1016/j.shpsb.2016.01.003

[Lei14] M. Leifer 2014, https://doi.org/10.12743/quanta.v3i1.22

[PBR12] M. Pusey, J. Barrett and T. Rudolph 2012, https://doi.org/10.1038/nphys2309

Support for Geometric Pooling

Jean Baccelli and Rush Stewart
Munich Center for Mathematical Philosophy

Forms of opinion pooling have been proposed as simpler alternatives to Bayesian conditioning in the context of learning from the opinions of others. But are there circumstances in which opinion pooling and Bayesian conditioning coincide? It has been established in the literature that, trivial cases put aside, linear pooling cannot give the Bayesian response. Our contribution is to show that there are circumstances in which geometric pooling, by contrast, can give the Bayesian response. More specifically, we show that, under certain simple and motivated assumptions, opinion pooling coincides with Bayesian conditioning if and only if it is geometric. The upshot, which proves robust to variations in the statement of the problem, is that geometric pooling enjoys a certain normative advantage over linear pooling as a recipe for social learning.

Long abstract:

Consider two or more Bayesians, endowed with a common prior. Assume that each privately gathers some evidence, updates her prior accordingly, and publicly announces her posterior beliefs. Now contrast the following two routes. The first corresponds to so-called supra-Bayesianism. It consists in treating the announcements of the posterior probability values as further

evidence on which the common prior is to be updated. The second route consists in pooling, more specifically, taking a weighted average of the announced posterior probability values. As the statement makes clear, this is simpler than supra-Bayesianism. One interesting question is whether there are pooling methods such that the two routes lead to the same result. This has been called the problem of Bayes-compatibility for pooling subjective probabilities. The problem has already been studied under linear pooling. An impossibility result has been established, to the effect that linear pooling cannot be (non-trivially) Bayes-compatible (Bradley, 2018). Under geometric pooling, the problem has been already touched upon (Dawid et al., 1995; Easwaran et al., 2016), but never systematically studied, to the best of our knowledge.

Our contribution is a step in this direction. We establish a possibility result, to the effect that geometric pooling can be Bayes-compatible. Indeed, our main result is that, under a simple construal of the problem, weighted geometric pooling is the only Bayes-compatible pooling method. To establish the result, we exploit certain recently-studied commutativity axioms of pooling functions (Dietrich, forthcoming). We also put our formal investigation in philosophical perspective by discussing how the Bayesian principle of total evidence translates in our setting. While geometric pooling respects the principle, linear pooling does not. Finally, we examine whether the comparative advantage of geometric pooling over linear pooling carries over to more general settings than the ones in which we investigate the Bayes-compatibility problem. We examine the effects of relaxing the common prior assumption into a common prior support assumption. We also examine the effects of appreciating the Bayes-compatibility of probability ratios, rather than single probability values. We find that, even when such generalizations have a significant impact on the Bayes-compatibility of geometric pooling, they do not eliminate its comparative advantage over linear pooling. This comparative advantage is the main take-home message of our study. It has implications in social epistemology for the problems of testimony, peer disagreement, and social learning.

References:

Bradley, 2018, “Learning from Others : Conditioning versus Averaging”, Theory and Decision, 85(1), 5–20

Dawid et al., 1995, “Coherent Combination of Experts’ Opinions”, Test, 4(2), 263–313

Dietrich, forthcoming, “A Theory of Bayesian Groups”, Noûs

Easwaran et al., 2016, “Updating on the Credences of Others : Disagreement, Agreement, and Synergy”, Philosophers’ Imprint, 16(11), 1–39

The Literary Form of Scientific Thought Experiments

Alice Murphy
University of Leeds

Thought experiments are a popular device in science used to justify, motivate, undermine or clarify theories. They take the form of short fictional narratives that have the purpose of instructing a scientific or public community to evaluate the described scenario in a certain way. In the philosophy of art, comparisons have been drawn between thought experiments and artworks, particularly works of literary fiction, as they share (at least some of) these key features, namely their fictionality—the events not actually taken place, or whether they have is inessential—and narrative form. Further, characterising fiction as a kind of extended, more complex thought experiment allows us to maintain that engaging with narrative art can lead us to new insights about the world and ourselves (Elgin, 2014).

Here, I want to discuss how comparisons can be drawn the other way as well, i.e. from aesthetics and philosophy of art in order to illuminate the science cases. My aim is to address how the aesthetic choices scientists make in the design of thought experiments contribute to the function of the thought experiment. The key issue is whether the aesthetic qualities provide anything beyond catching and maintaining our attention or at best, are a mere heuristic aid. There is a set of views that have argued this way, claiming that there are disanalogies between the art and science cases that undermine purported connections between how we learn from scientific and artistic representations, and the role of aesthetic considerations in science.

For example, Norton analyses thought experiments as arguments: all thought experiments can be reconstructed into argument form without any epistemic loss (2002, 50). Similarly for Egan (2016), thought experiments' typical narrative form and any appeal to concrete particulars are irrelevant to the conclusion and therefore dispensable. This presents a contrast with literary fictions, where the concrete particulars are an irreducible part of our engagement with the work. For Currie, models (and we can include thought experiments) 'are not dependent for their value in learning on any particular formulation' (2016, 305). Artistic fictions, on the other hand, do depend on their formal qualities in order to convince. We can add that we value such works for their formal properties and artistic skill, whereas models and thought experiments are not evaluated aesthetically.

A final difference is to do with what Frigg and Nguyen call 'the flexibility of interpretation' in artistic representations compared to scientific ones. In the case of models, they claim that the interpretation is 'usually fixed by the context and the interpretation highly regimented' (2017, 57). In works of literature, the interpretation is not fixed and attending carefully to the work and its features in order to come up with interesting and sometimes conflicting interpretations is part of engaging with those works. This has been discussed in the case of thought experiments. Hacking argues that unlike concrete experiments, thought experiments do not have a life of their own, they are 'fixed, largely immutable' (1992, 307).

I agree that there are significant differences between our engagement with art on one hand, and our engagement with science on the other. And these differences need to be taken into consideration when drawing comparisons between scientific and artistic representations and how we learn from them. But I want to resist the extent of their claims. I argue that while we can, of course, rationally reconstruct thought experiments into argument form, this will lead us to miss important features involved in their practice, and this is what I am interested in examining. As a consequence, I argue against Norton and Egan, and Currie's claim that formulation does not matter in scientific representations.

I demonstrate that there are important cognitive and practical considerations involved in the formulation of thought experiments, and that their demonstrative force is dependent upon their narrative form and appeal to ordinary, everyday objects. Further, I show the significance of utilising images alongside scientific thought experiments, and how this aids the imaginative process. Finally, I address the issue to do with flexibility of interpretation of artworks, compared to thought experiments and models. There is huge debate in the philosophy of art regarding interpretations of artworks, how flexible this is, and how much this can deviate from the artist's intentions. In addition, thought experiments are not limited to a single interpretation (Bokulich, 2001). There can be disagreement on what would happen in the scenario presented, or what conclusion we should draw, and this is dependent on theoretical commitments.

In summary, thought experiments are a good case study for thinking about aesthetic features in the scientific context. As a result, the difference between representations in art and science raised in current discussions is not as stark as it has been made out to be, and science is a more heterogeneous practice than has been allowed. Part of the value of thought experiments in scientific practice includes the qualities they share with literary works.

Bibliography

Bokulich, Alisa. "Rethinking Thought Experiments." Perspectives on Science 9, no. 3 (2001): 285–307.

Currie, Gregory. "Models As Fictions, Fictions As Models." The Monist 99, no. 3 (2016): 296–310.

Egan, David. "Literature and Thought Experiments." The Journal of Aesthetics and Art Criticism 74, no. 2 (2016): 139–50.

Elgin, Catherine Z. "Fiction as Thought Experiment." Perspectives on Science 22, no. 2 (2014): 221–241.

Frigg, R. and J. Nguyen, (2017). Of barrels and pipes: representation - as in art and science. In Otávio Bueno, George Darby, Steven French & Dean Rickles (eds.), Thinking about Science and Reflecting on Art: Bringing Aesthetics and the Philosophy of Science Together. Routledge, London and New York: pp. 41-61.

Hacking, Ian. "Do Thought Experiments Have a Life of Their Own? Comments on James Brown, Nancy Nersessian and David Gooding." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1992 (1992): 302–308.

Norton, John D. "Why Thought Experiments Do Not Transcend Empiricism." In Contemporary Debates in the Philosophy of Science, edited by Christopher Hitchcock, 44–66. Blackwell, (2002).

The Ontology of Patterns

Tiziano Ferrando
Université de Lausanne

Ladyman and Ross (2007, 2013) propose an ontology of real patterns based on previous work by Dennett (1991) and Ross (2004). Real patterns are supposed to give a precise way of understanding what emergence is and the relation between fundamental physics and special sciences. I argue that although the theory is tenable, it stands in need of elaboration with respect to some relevant issues: (1) Clarify the relation between the three existing approaches to describe real patterns: information-theoretic, statistic, and dynamic; (2) Establish the mind-independence of real patterns; (3) Introduce a notion of ontological dependence between emergent entities; (4) give an account of scale relativity that incorporates ways emergence occurs at scales others than size or time, particularly with respect to energy and complexity.

The aim of the paper is to address these issues:

(1) I argue that the information-theoretic definition of real patterns in Ross (2004) and Ladyman and Ross (2007) subsumes the ones in terms of non-redundant statistics and reduction of degrees of freedom, although depending on the context it may be useful to use one or the other. This is because in the end all three rely on information-processing, whether we are looking for statistical generalisations or the dynamics of a system in phase space. The information-theoretic setting may nevertheless be flawed as it stands, as suggested by Beni's (2017) criticism of the notion of projectibility.

(2) The ontological status of patterns and their connection to pattern-recognisers (agents, observers or information processors) can be understood as the manifestation of a power. Potential patterns manifest themselves as information when a system is coupled with a pattern-recogniser. Through the introduction of powers, we can say that the pattern can convey information if, at the right scale, it is coupled to a pattern-recogniser with enough computational resources. If there is no pattern-recogniser the pattern exists as unmanifested. This way we could say that although real patterns are indeed a product of data compression, there are patterns when no one looks at them. Also, this way of conceiving of "patterns in the wild" as potentialities may fit well with a proposal by McAllister (2011), who claims that each data set admits all possible patterns with a different amount of noise, and that the presence of a pattern when confronting datasets points to an existing structure in the outside world. The emergent and irreducible features of a real patterns give us new ways of addressing questions concerning identity, persistence and vagueness.

(3) Dynamics and degrees of freedom of a system play an important role in understanding how emergent patterns relate to other patterns at different scales. I grant Ladyman (2017) that an account of composition which fits actual science has to be dynamic and diachronic, and I argue that the same should hold for ontological dependence. I propose a notion of dynamical dependence for inter and intra level patterns, and explore whether it should be taken to be symmetric/asymmetric, transitive/intransitive, global/local. I argue that even if there might be no general notion that works for all cases of dependence between patterns, one could still benefit from considering dynamical dependence as an umbrella term, and address case by case the specific features of the relation according to the relative scientific domain.

(4) Ladyman and Ross (2007) claim that ontology is scale relative with respect to both space and time. I agree with the claim, but add that those are not the only scales we should look at when searching for "novel and robust behaviour" (Butterfield 2011). Interesting considerations about emergence, persistence and fundamentality are relative to the scale we are investigating. I will focus on the question of whether genuine emergence could occur with respect to some scales but not to others. Phase transitions occur at different levels of the energy scale, but extension in space and time fails to capture the relevant dynamical dependence. Complex behaviour could also count as emergent, but the complexity scale seems to be independent from space and time: the functioning of a star is simpler than a cat's digestive system. I will consider the interplay between different scales, and whether some scales could be taken to be dependent on others or redundant.

References:

Beni, Majid Davoody (2017). "Structural Realism, Metaphysical Unification, and the Ontology and Epistemology of Patterns". International Studies in the Philosophy of Science 31 (3):285-300.

Butterfield, Jeremy (2011). "Emergence, Reduction and Supervenience: a Varied Landscape". Found Phys 41: 920.

Dennett, Daniel C. (1991). “Real patterns”. Journal of Philosophy 88 (1):27-51.

Ladyman, J. (2017). “An Apology for Naturalized Metaphysics”. In: Matthew Slater and Zanja Yudell (eds), Metaphysics and the Philosophy of Science: New Essays. Oxford University Press.

Ladyman, James & Ross, Don (2007). Every Thing Must Go: Metaphysics Naturalized. Oxford University Press.

Ladyman, James & Ross, Don (2013). “The World in the Data”. In Don Ross, James Ladyman & Harold Kincaid (eds.), Scientific Metaphysics. New York, USA: Oxford University Press. pp. 108-150.

McAllister, James W. (2011). “What do patterns in empirical data tell us about the structure of the world?”. Synthese 182 (1):73-87.

Ross, D. (2004). “Rainforest realism: A Dennettian theory of existence”. In D. Ross, A. Brooks & D. Thompson (eds.), Dennett's Philosophy: A Comprehensive Assessment. MIT Press. pp. 147-168.

Time, Cauchy Problems and Physical Modality

Lucy James
University of Bristol

One aim of Callender's (2017) book, `What Makes Time Special?', is to unify those physical features of time which distinguish it from space, by means of physical modality, as understood according to the Best System Account of laws of nature (henceforth, the BSA). After giving an overview of Callender's argument, I give a general methodological critique which I respond to by suggesting a more conservative aim: to unify features of time in certain sub-classes of theories rather than across physics as a whole. I discuss how connections between mathematical form and physical modality relate to this project, and argue that the BSA is unhelpful here. My second criticism is that, with or without the BSA, Callender's focus is narrow. I respond to this by considering the form of a different set of laws, drawing comparisons between these and the laws identified by Callender.

To summarise Callender's argument very briefly, the BSA identifies as laws the axioms and theorems of the `best' deductive systems that describe the world. `Best' is understood as balancing a number of theoretical virtues including strength and simplicity. I acknowledge a difficulty in understanding the metaphysical significance of this seemingly pragmatic consideration (see Baron and Evans (2018)). Callender focuses on maximising strength without too heavy a cost to simplicity; this is spelt out as `informativeness', i.e. the system's ability to generate maximal information given minimal antecedent data. Informativeness in this sense, so the argument goes, is best achieved by partial differential equations (henceforth PDEs) which admit well-posed Cauchy problems. The general form of these PDEs is linear, second-order and hyperbolic. Their hyperbolicity ensures that, whenever they are applied to physical contexts, they take antecedent data on a spacelike hypersurface and evolve in timelike directions. The conclusion from this is supposed to be that there is a metaphysically deep asymmetry between time and space, which serves as a grounding for other special features of time we might identify.

I criticise this application of the BSA on the basis that it is not sufficiently general or unambiguous in its identification of physical laws, and so cannot achieve the unifying task required of it. I further criticise the BSA as a metaphysical account of physical modality, and my comments here lead me to resist the idea that we can find any account of modality which is both detailed enough to be useful and general enough to include all instances of what usually get called `laws of nature'. On a more positive note, I advocate adopting a flexible notion of law which arises directly from the way the term is used by scientists. To be clear, I do not attempt to give an account of laws of nature, or make explicit use of any existing account. I resist a general account altogether, but this resistance does not mean that modality is entirely mysterious - in fact, various forms of modality are investigated in a multitude of ways. I argue that mathematics can provide some ways to gain insight into physical modality. With this in mind, I investigate structural similarities between the laws identified by Callender's application of the BSA and another set of dynamical law represented by non-linear equations. This makes for a more general analysis (although not close to being maximally general) of how time is distinguished in physical theories, by the mathematics used to model them. This, I shall argue, is partly constitutive of physical modality.

Despite my criticisms of it, I do see some merit in the BSA and the philosophical principles it arose out of. In particular, I am not against the Humean spirit of requiring that a metaphysical account of modality be in harmony with scientific epistemology. A qualm I have with the account (and this is also what makes it unsuitable for unifying physical features of time) is that, as soon as we try to make precise the notions of strength, simplicity and balance, it becomes too restrictive and thus departs from actual scientific epistemology. Metaphysics ought to reflect good scientific practice - this I take to be a basic principle (see Ladyman and Ross (2007) for arguments to this effect) - and good scientific practice does indeed involve the balancing of strength and simplicity, among other theoretical virtues, when building and refining theories. However, these theoretical virtues are only ever given precise definitions in specific contexts, the details of which vary accordingly. Callender's unification project fails because it attempts to draw a general conclusion from a particular case study.

Expanding our investigations, whilst still not achieving the level of generality required to unify features of time across all of physics, allows us to tend towards a more unified picture. To this end, I investigate similarities between the way time is distinguished in both well-posed and ill-posed Cauchy problems. Mathematical form, I argue, does have a lot to do with the ways we understand physical modality. That is, in many cases, the mathematics used in a theory tells us what is necessary or impossible (usually, in practice, mathematics is used to constrain possibility - this is why I talk about possibility in the negative sense) according to that theory, which in turn provides us with our best guess as to what to expect from the physical world. More concretely, if some mathematical theorem (e.g. Bell's theorem) prohibits a certain physical scenario (e.g. causally local hidden variable theories), it is usual to deem that scenario impossible (obviously, provided the assumptions on which the theorem rests are accepted). By equal measure, if some mathematical result follows of necessity, then the corresponding physical result is also necessary - but only given the relevant antecedents. In Callender's case, it is necessary for those dynamical processes which lend themselves to being modelled by hyperbolic PDEs to make a distinction between time and space, of the sort presented. What I show in addition to this is that those processes modelled by non-linear PDEs have a different, although related, sort of asymmetry between time and space built into their geometric structure.

What Are We Pluralist About?

Franklin Jacoby
The University of Edinburgh

Pluralism about science has become a popular position that seems to do justice to the diversity we see in science while avoiding the over-simplification associated with monism. What, however, should be pluralist about? This paper sketches some possible answers to this question and argues that a common form of epistemic pluralism fails to provide sufficient criteria for identifying epistemic systems. I propose some additional criteria that, together with the old, provide a more complete picture of epistemic pluralism.

One common approach to pluralism is to be pluralist about epistemic systems or practices. Kusch (2017) and Chang (2012) take this approach and some members of the Minnesota school (2006) are also pluralists about epistemic practices and systems. Scientists, this pluralist stance suggests, are distributed into different epistemic systems or practices and the models, explanations, or evidence those scientist produce and the values they have are relative to those practices. This raises the following question: what is a practice and what makes one practice different from another? What criteria should we use to identify the practices to which scientists belong? This question is pressing for epistemic pluralists because without a definite answer, it is not only unclear who belongs to which practice, but also what practices there are. This lack of clarity makes it difficult to fully address some of the issues raised by Wylie (2015), such as how do we ensure a diversity of epistemic values are properly represented and how we should approach epistemic disagreements and conflicts. When should disagreements be resolved in favour of one side and when should we expect a more collaborative or pluralistic resolution?

One approach to the problem of identifying practices, which follows pragmatist lines, suggests a practice is defined by the goals or questions members have and the associated activities those scientist perform toward achieving goals or answering questions. Goals are the criteria by which we should identify practices. Chang (2015, 2012), Longino (2002), Danks (2015) are just a few who broadly follow this approach.

In this talk, I will argue, however, that goals on their own are insufficient criteria for identifying practices. One problem is that it is difficult to specify which goals are important for defining practices. This approach also suggests very little is shared between disagreeing scientists, which obscures the basis of disagreements and controversies. Goals are often very widely shared, construed properly, even between disagreeing scientists. It is also unclear how a robust notion of truth could play the kind of role in science that we intuitively expect.

I propose some additional criteria for identifying practices, which I call a perspectival approach, that draws on some analyses of practices in the philosophy of language, particularly those by Dummett (1993) and Ryle (1945). This approach goes some way toward mediating the conerns I raise against against using goals to define practices. A practice, my analysis suggests, requires a certain type of practical ability for membership. It is this ability in conjunction with goals that define a practice. Consequences of this view include allocating some disagreeing scientists to the same epistemic system and making room for more robust notions of truth. It also makes clear which goals are of central importance to a practice.

I argue the knowledge-how consists in the common use of a taxonomic system. As such, it follows in the tradition of drawing a close connection between meaning and use. The ability to use the conceptual taxonomy associated with a scientific inquiry is a precondition for engaging in that inquiry and thus being a member of the associated scientific practice. This view is perspectival because there can emerge different interpretations of a taxonomy that offer rival “perspectives.” Disagreements and controversies emerge because of these discrepencies. However, I will suggest that because these discrepencies are within a broadly shared system, they are resolveable. Although there can be multiple uses, not all uses are equally good.

This view of practices is not only compatible with a social epistemology reading of science, but also a more individual-based epistemology because it specifies what the epistemic standards are that individuals must meet in order to be members of a practice.

There are several upshots of this view. One is that epistemic practices can be more clearly delineated and analysed. It would be apparent which scientists are members of which epistemic practices or systems and, consequently, whether a diversity of views and values is well-represented. The role of goals in structuring a practice and guiding scientific work will be more perspicuous, which I believe addresses some of the worries that Boghossian (2007) raises. This view will also enable us to learn more about epistemic disagreements and conflicts by telling us when we have disagreements within a practice and when we have disagreements between practices.

Selected Literature

Boghossian, Paul. (2007). Fear of Knowledge: Against Relativism and Constructivism. Clarendon Press.

Chang, Hasok. (2012). Is Water H2O?: Evidence, Realism and Pluralism. Vol. 293. Springer Science & Business Media.

———. (2015). “The Chemical Revolution Revisited.” Studies in History and Philosophy of Science Part A 49. Elsevier Ltd: 91–98. https://doi.org/10.1016/j.shpsa.2014.11.002.

Danks, David. (2015). “Goal-Dependence in (Scientific) Ontology.” Synthese 192 (11): 3601–16.

Dummett, Michael. (1993). The Seas of Language. Oxford, New York: Oxford University Press.

Kellert, Stephen H, Helen E Longino, and Kenneth C Waters. (2006). Scientific Pluralism. Edited by Stephen H. Kellert, Helen E. Longino, and Kenneth C. Waters. Minneapolis: University of Minnesota Press.

Kusch, Martin. (2017). “Epistemic Relativism, Scepticism, Pluralism.” Synthese 194 (12). Springer Netherlands: 4687–4703. https://doi.org/10.1007/s11229-016-1041-0.

Longino, Helen E. (2002). The Fate of Knowledge. Princeton University Press.

Ryle, Gilbert. (1945). “Knowing How and Knowing That: The Presidential Address.” In Proceedings of the Aristotelian Society, 46:1–16.

Wylie, Alison. (2015). “A Plurality of Pluralisms: Collaborative Practice in Archaeology.” In Objectivity in Science, 189–210. Springer.

What Does It Take To Be a Psychological Primitive? Separating Innateness from Foundationalism

Domi Dessaix
Australian National University

This paper is part of a broader project on primitive concepts and whether they can play any role in a psychologically and biologically plausible theory of meaning in natural language. The first aim of that project is simply to get clearer on what it would take for statements like “x is a primitive concept” to be true. The second aim is to bridge the use of primitives in linguists’ theories of meaning with the psychological and biological facts about concept acquisition. This paper is centered on the first aim: getting clear on the view that there are primitive concepts. It focuses on Susan Carey’s (2009) view, according to which we have a stock of innate “conceptual primitives”, defined as primitive mental representations that are neither sensory nor perceptual, which lay the foundation for the rest of conceptual development. I will argue that analysing some of the key concepts here, especially the notion of “innateness” (widely acknowledged to be problematic, see e.g. Mameli & Bateson 2006), reveals that there are several independent claims at work in a proposal like that of Carey’s. In particular, I argue that some representation (or any psychological entity) having a foundational role in learning is distinct from it being innate, and that these two claims require partly independent kinds of evidence. For example, the first requires evidence for the representation emerging prior to others in the same domain. Yet evidence of this kind has no bearing on the second claim, which instead requires evidence that the representation is not acquirable solely via domain-general learning mechanisms. I also argue that in this case (if not more generally) the claim to innateness entails a claim about the specifically genetic contribution to the development of the trait.

The paper is structured as follows. I first introduce Carey’s (2009) proposal that there are innate “conceptual primitives”, and point to some of the developmental evidence she uses to support two particular proposed primitives, OBJECT and AGENT. I briefly set out Carey’s view that these primitives form part of “core cognition”, which in conjunction with a bootstrapping learning mechanism, can explain our acquisition of novel concepts. Then I move on to argue that the claim about the proposed primitives’ foundational role in learning is distinct from the claim to innateness, but that these are independently interesting. I also briefly point to uses of primitives in linguistic theories of meaning, such as in Jackendoff’s (2002) Conceptual Semantics, arguing that to the extent that such a theory succeeds in its aim to be a fully naturalistic – i.e. psychologically and biologically plausible – account of the semantics of natural language, it needs to be hooked up to some specific claims about what the primitives are, and thus must answer to the kinds of questions I discuss here.

References

Carey, S. (2009). The Origin of Concepts. Oxford: Oxford University Press.

Mameli, M. & P. Bateson. (2006). ‘Innateness and the sciences’. Biology and Philosophy, 21: 155-188.

Jackendoff, R. (2002). The Foundations of Language. Oxford: Oxford University Press.

What Is (successful) Extrapolation?

Donal Khosrowi
Durham University

Extrapolation plays important roles in scientific activities that involve studying some empirical system A with the aim of reliably drawing a conclusion about some distinct target system B. The extrapolation of causal effects and causal claims, in particular, has attracted the attention of methodologists, statisticians, and philosophers of science, and various proposals have been offered concerning how successful extrapolation may be achieved, with some authors suggesting that the problem of extrapolation has been successfully solved (e.g. Marcellesi, 2015 Philosophy of Science).

I suggest that such conclusions are unwarranted, and indeed that the very question of whether the problem of extrapolation has been solved is misleading. Two basic questions need to be addressed first: 1) What is an extrapolation? 2) What constitutes successful extrapolation? My aim is to make progress on these questions, showing that doing so can improve our understanding of extrapolation, can help us reach more nuanced conclusions about whether particular strategies for extrapolation are likely to be successful, and about whether any problems of extrapolation have been solved.

Addressing the first question, I begin from the observation that existing treatments have employed overly simplistic notions of extrapolation that unhelpfully gloss over the considerable diversity in real-world problems of extrapolation, leaving unclear which strategies are likely to be successful in overcoming these problems. To improve on this, I propose a more nuanced analysis of extrapolation as a highly heterogeneous collection of inferential activities targeting a highly diverse range of problems that, while amenable to a single, systematic analysis, can exhibit important differences. This analysis proceeds in two stages.

The first takes issue with important differences in the challenges posed by problems of extrapolation, i.e. variations in the kinds of causally relevant differences between study and target systems that make problems of extrapolation challenging. Here, I offer a systematic framework that distinguishes different levels at which study and target systems can exhibit such differences, and discuss different and heretofore unrecognized ways in which they can be realized at each level. Building on this, the second part of the analysis distinguishes different kinds of extrapolative inferences along several dimensions, including the envisioned mode of inference and its fidelity, the kinds of causal queries at issue, the availability of background theory and knowledge about study and target systems, the epistemic risks involved etc.

This two-stage analysis is useful in several ways: 1) It helps us recognize how problems of extrapolation vary considerably in difficulty, and how extrapolative inferences can significantly differ in their epistemic ambitions and important contextual features. 2) It helps distinguish existing strategies for extrapolation with respect to what kinds of problems they can in principle address and helps criticise them accordingly for their limitations. 3) It makes clear that there is no single answer to whether the problem of extrapolation has been solved, but multiple answers that depend on various contextual details.

In the second part, I aim to make further progress on providing such answers by clarifying what constitutes successful extrapolation. Here, an important challenge to be considered is Steel's (2008, OUP) extrapolator's circle: the knowledge about the target system required to arrive at an extrapolative conclusion must not be so extensive that this conclusion can be reached based on knowledge about the target alone. This would make the information from the study setting redundant. Recognizing the importance of this challenge, I make two proposals for how to improve our understanding of what constitutes successful extrapolation.

First, I add that the extrapolator's circle should not be understood as an all-or-nothing affair but rather in gradual terms, where the less significant the role played by the information obtained from a study in reaching a conclusion about a target, the more we fall prey to the extrapolator's circle.

Second, I argue that the extrapolator's circle should be incorporated into our understanding of what constitutes successful extrapolation. I propose a distinction between three dimensions of success:

The first part, justification, concerns how much empirical support an extrapolative conclusion enjoys. Here it seems that the more support there is, the better. At the same time, the support required will depend on important contextual features, including the kind of problem of extrapolation one faces and the kind of extrapolative inference one aims for. The second part, accuracy, concerns how accurate our extrapolative conclusion is. Again, the more accurate, the better, but the type of problem targeted, and the kind of extrapolative inference envisioned will bear importantly on how likely we are to achieve success. The third part, relevance, incorporates the extrapolator's circle. The less relevant the knowledge from a study setting is to our extrapolative conclusion, the less successful an extrapolation is. In the limit, when we can answer an extrapolative query based on information about the target alone, extrapolation fails entirely.

With these distinctions in place, I argue that general success in extrapolation should be understood as requiring a good mixture of justification, accuracy, and relevance, with all being necessary and none being sufficient for overall success. Moreover, I emphasize important tensions between different success criteria. For instance, the more extensive the empirical evidence about the experimental and target settings used to justify an extrapolation, the more accurate our extrapolative inferences will tend to be. But this will often come at the cost of falling prey to the extrapolator's circle. I illustrate how different strategies for extrapolation fare on the success criteria, and how they experience difficulties in responding to the tensions, leaving unclear whether they are, in general, likely to help achieve successful extrapolation.

I conclude that we should be sceptical about whether the problem of extrapolation has been solved. There is no single problem, but a multiplicity of problems, some of which are easier to solve, whereas others remain unlikely to be overcome by any strategy for extrapolation. Moreover, existing strategies for extrapolation may be able to address some problems but are unable to address a wide range of others, and some kinds of problems are unlikely to be successfully overcome by any strategy.

What We Cannot Learn from Analogue Experiments

Karen Crowther, Niels Linnemann and Christian Wüthrich
University of Geneva

Analogue experiments have attracted interest for their potential to shed light on inaccessible domains. Examples include the `acoustic horizons' of `dumb holes' in fluids and Bose-Einstein condensates, which are supposed to be analogues of black hole horizons (e.g., Euve et al., 2016, Steinhauer, 2016; Weinfurtner et al., 2013). These `tabletop' experiments produce effective phenomena that are described by modelling frameworks formally similar to those that are thought to describe black holes. Recently, several philosophers have argued that, under certain conditions, analogue experiments can provide confirmation of the existence of particular phenomena in their (inaccessible) target systems (Dardashti,Thebault & Winsberg, 2017; Dardashti et al., 2018; Thebault, Forthcoming). [Henceforth, I refer to these references as DTW].

DTW's main argument has the following form:

A system S provides an analogue experiment of system T when the following set of conditions obtain:

[Step 1:] For certain purposes and to a certain degree of desired accuracy, modelling framework MS is adequate for modelling system S within a certain domain of conditions DS.

[Step 2:] For certain purposes and to a certain degree of desired accuracy, modelling framework MT is adequate for modelling system T within a certain domain of conditions DT

[Step 3:] There exists exploitable mathematical similarities between the structure of MS and MT sufficient to define a syntactic isomorphism robust within the domains DS and DT.

[Step 4:] We are interested in knowing something about the behaviour of a system T within the domain of conditions DT, and to a degree of accuracy and for a purpose consistent with those specified in Step 2. For whatever reasons, however, we are unable to directly observe the behaviour of a system T in those conditions to the degree of accuracy we require.

[Step 5:] We are, on the other hand, able to study a system S after having put it under such conditions as will enable us to conclude a statement of the form:

[Claim S:] Under conditions DS and to degree of accuracy that will be needed below, we can for the purpose of employing the reasoning below assert that a system S will exhibit phenomena PS.

The formal similarities mentioned in Step 3 then allow us to reason from Claim S to [Claim T:] Under conditions DT, a system T will exhibit phenomena PT.

The relevant appearance of PS in an analogue experiment S is able, according to DTW, to provide confirmation of PT when we have `Model External Empirically Grounded Arguments' (MEEGA) that establish the universality of the particular phenomenon in question. The MEEGA are supposed to boost [Step 3] in DTW's argument, providing empirically grounded reasons to believe that PS and PT are in the same universality class.

The main example that DTW consider is the potential confirmation of the existence of Hawking radiation (PT) in black holes (T), by the relevant appearance of analogue Hawking radiation (PS) in dumb holes (S).

In this paper, we argue (against DTW) that analogue experiments cannot, in fact, provide confirmation of the existence of phenomena (PT) in their target systems. We first characterise analogue experiment by comparison with conventional experiment (which, we take it, can be confirmatory), arguing that the significant difference is the (relevant) inaccessibility of T in the case of analogue experiment. In all cases of experiment, S and T are presumed to be the same kind of system for the purposes of interest---i.e., they are supposed to produce phenomena in the same universality class (and are thus described by formally similar modelling frameworks). However, in the case of analogue experiment, the inaccessibility of T prevents the scientist from knowing whether or not S and T really are the same kind of system (in the sense just described).

We then outline the derivation of Hawking radiation, arguing that one of the many reasons why the confirmation of Hawking radiation would be of importance is that it would provide a crucial test of the framework (MT) that physicists use to describe black holes: quantum field theory (QFT) in curved spacetime. This is because scientists do not know if QFT in curved spacetime is in fact the correct framework to describe black holes (we find only one other potential test of this framework, being the prediction of a primordial specific primordial density perturbation spectrum associated with cosmic inflation scenarios, and argue that this is insufficient for establishing the applicability of QFT in curved spacetime).

Finally, we present our argument against DTW. In an analogue experiment, [Step 2] of DTW's argument cannot be established---because T is relevantly inaccessible, scientists cannot know that MT actually describes it. In the example of Hawking radiation, scientists do not know that QFT in curved spacetime actually describes black holes. The derivation of Hawking radiation is a consequence of QFT in curved spacetime (plus some assumptions), and thus, by assuming that QFT in curved spacetime does describe black holes, DTW are already assuming that there is Hawking radiation in black holes. By making this assumption in [Step 2], DTW are already assuming Claim T, which is what their argument is supposed to establish.

Analogue experiments are significant for a number of reasons. However, given their essential use of analogue reasoning, together with the inaccessibility of their target systems, we argue that they cannot provide any more confirmation of PT than other cases of analogue reasoning in science.

What's So Spatial About Time Anyway?

Peter Evans and Sam Baron
The University of Queensland (Evans) University of Western Australia (Baron)

In his recent book, Callender (2017) argues that time can be distinguished from space due to the special role it plays in our laws of nature: our laws determine the behaviour of physical systems across time, but not across space. In assessing the claim that the laws of nature might provide the basis for distinguishing time from space, this talk develops a radical reading of Callender's view and proposes a novel approach to differentiating time and space that we call temporal perspectivalism. This is the view according to which the difference between time and space is a function of the agentive perspective.

According to Callender, the feature that differentiates time from space is that time is the 'great informer': time is the direction in the manifold in which the greatest amount of information can be generated by the smallest set of antecedent conditions. This sort of informativeness is a hallmark of a good balance between strength and simplicity in a best systems account of laws: the laws arise as the most accurate description of as much of the world as possible (strength) in the most succinct manner (simplicity). Thus, for Callender, 'time is that direction in spacetime in which we can tell the strongest or most informative stories' (2017, p. 142).

Callender offers two distinct but related arguments in favour of his view. The first, call it the 'informal' argument (2017, Ch. 7), shows how the direction of informativeness that emerges from the process of systematisation that characterises the best systems account of laws 'binds' together a set of features ordinarily associated with time. Callender begins this argument espousing a conservative reading, in which the systematisation exposes 'an asymmetry in the distribution of events' (2017, p. 142) in the manifold, before progressing to a more radical reading, in which 'the choice of metric geometry hangs on systematizing too' such that 'the difference is not "out there" prior to systematizing' (2017, p. 151). The second argument, call it the 'formal' argument (2017, Ch. 8), shows that it is a formal property of the laws that uniquely distils the direction of informativeness, and so more tightly connects the set of temporal features with this direction. This talk outlines both arguments with a view to considering the consequences of combining the lessons of the formal argument with the radical reading of the informal argument. We suggest that, on the radical reading, the difference between time and space in the formal analysis may be due to an underlying pragmatic choice of natural kinds.

On the radical reading of the informal argument, it is not simply the laws that arise from the best systems systematisation, but the spacetime geometry, particularly the metric signature. Given one or more physical fields on a manifold with no metric structure, and the fact that the laws governing the fields admit of hyperbolisation, it can be shown that the laws define a metric structure on the manifold Geroch (2011). Thus systematisation on the radical reading defines the metrical geometry, and so the metrical difference between time and space. However, according to the better best systems account, this systematisation is carried out with respect to a particular choice of natural kinds, and it is the agents conducting the systematisation who make the choice of natural kinds in the interest of optimisation.

When we carry out the systematisation process, implicit in the process is that we are not simply trying to find the laws that best meet some absolute trade-off between simplicity and strength, we are trying to find those laws that are strongest in a manner that aligns with our particular predictive practices. Given this, it is plausible that our predictive practices, and so the particular trade-off that we make, are a function of our epistemic vantage point on the world. The information about the manifold to which we have access is exclusively in our pasts, and we are interested in using such data to model our unknown futures. Since we take our past to be a predictor of our future, we are naturally predisposed to take the boundary between our past and future as the antecedent boundary of our predictive practices, and to model our future as generated from our past (we have developed scientific practices that reflect this natural predisposition). Thus when systematising over the distribution of events to which we have access, we are pragmatically constrained to identify natural kinds living on the antecedent boundary separating our past from our future---that is, spacelike hypersurfaces---and that are best placed to allow efficient algorithms to take such antecedent boundaries as input and such kinds as dependent variables.

It is in this way that we can understand the lessons of the formal argument along the lines of the radical reading of the informal argument. According to Callender's formal argument the best, most efficient, and most informative algorithms developed by our scientific practices admit only antecedent data on spacelike hypersurfaces. But while a conservative reading of the formal argument renders these spacelike hypersurfaces as a property of a given metric signature inherent in the distribution of events on the manifold, a radical reading renders these spacelike hypersurfaces as effectively arising during the systematisation of the natural kinds and laws, which we claim is a function of the pragmatic concerns of the systematising agents: our pragmatic choice of natural kinds on spacelike boundaries reflects our interest in modelling our future as a function of our past. The pairing of natural kinds and laws brings with it a time/space split that reflects the epistemic constraints and pragmatic interests of agents. We call this view temporal perspectivalism.

The talk concludes with a toy example demonstrating the perspectival nature of the distinction between time and space according to temporal perspectivalism.

Why Symptom Based Approaches Are Not Enough: The Value of Psychiatric Diagnoses

Sam Fellowes

Lancaster University

Critics are concerned that psychiatric diagnoses fail to accurately describe patients and therefore should be abandoned. Most patients do not have all symptoms associated with their diagnosis and most patients have symptoms which are not associated with their diagnosis. Knowing someone has a diagnosis seems to convey much less useful information compared to knowing what symptoms someone has. It is certainly possible to learn both the diagnosis someone has and what symptoms they have but this seems to leave psychiatric diagnosis as superfluous (Boyle 1990, p.83; Cromby, Harper & Reavey 2015, p.116; Timini, Gardner & McCabe 2011, p.1). Additionally, critics often claim diagnoses can be harmful distractions. Diagnosed people are sometimes primarily seen as being their diagnosis whilst other, more helpful, ways of understanding the individual are not considered (Billingdon 2016, p.242; Cromby, Harper & Reavey 2015; p.115; Hodge 2016, p.197). If diagnoses are superfluous then there seems no good reason to employ them and if diagnoses can be harmful distractions then there seems good reason to abandon them. In this paper I will argue that psychiatric diagnoses have important benefits which are not recognised by these critics. In this paper I will employ Ronald Giere's account of scientific theories to show that those critics are mistaken to see psychiatric diagnosis as making no useful contribution.

Giere's account of scientific theories has previously been applied to psychiatry (for example, Murphy 2006) but one aspect has not been explored. Giere describes how scientific theories are abstract generalisations which lack specific detail. For example, Newton's laws, by themselves, make no claims about the world. Rather, they guide the building of more specific models and these specific models can be used to make claims about the world. He describe scientific theories as “recipes for constructing models” (Giere 1994, p.293). This notion of scientific theories as recipes which guide the building of less abstract models has not yet been applied to psychiatric diagnoses.

I argue psychiatric diagnoses guide the construction of models of people. They make contributions to understanding individuals which are absent when simply focusing upon what symptoms are being presented by specific individuals. Firstly, many symptoms can be subtle and difficult to spot. A patient may be unaware of the symptom and psychiatrists cannot practically investigate for every possible symptom. Psychiatric diagnoses can help guide investigation of symptoms. If an individual exhibits a few symptoms of a psychiatric diagnosis then there is reason to investigate for other symptoms of that psychiatric diagnosis. If an individual exhibits low social skills and low eye contact, both of which are symptoms of autism, then there is reason to investigate for other symptoms of autism. This may help spot subtle symptoms such as rigid thinking or difficulty accommodating to changes. Thus the diagnosis guides investigating for the presence of symptoms. Secondly, patients fluctuate in the symptoms they present over time. The symptoms which are presented to a psychiatrist at time of interview may not cover symptoms previously exhibited or those exhibited in the future. However, knowing the individual has a diagnosis which is associated with a range of symptoms, more than any one diagnosed person actually exhibits, guides awareness towards a range of possible symptoms not present in a diagnosed person at one specific time. The diagnosis guides awareness towards alternative symptoms that may present at other times within diagnosed individuals. Thirdly, symptoms themselves have a level of generality and may manifest in quite different ways. For example, the low social skills of autistic individuals are typically quite different to the low social skills of schizophrenic individuals. Thus knowing the diagnosis of an individual can lead to greater understanding of how specific symptoms manifest. The diagnosis guides building more realistic models of ways individuals manifest symptoms.

I have shown how psychiatric diagnosis make a contribution to understanding individuals. Critics of psychiatric diagnosis are mistaken to believe psychiatric diagnosis make no contribution and should be abandoned.

References:

Billingdon, T. (2015). Critical Autism and Critical Neuroscience: Towards a Science of Research and Practise. 239-251. in Runswick-Cole, K., Mallett, R. & Timini, S. Re-Thinking Autism (London: Jessica Kingsley Publishing).

Boyle, Mary. (1990). Schizophrenia: a scientific delusion? (London: Routledge).

Cromby, John., Harper, Dave. & Reavey, Paula. (2015). Psychology, Mental Health and Distress. (Basingstoke: Palgrave-McMillian).

Giere, Ronald, N. (1994). The Cognitive Structure of Scientific Theories. Philosophy of Science, 61/2, 276-296.

Hodge, N. (2015). Schools without labels. 185-203. in Runswick-Cole, K., Mallett, R. & Timini, S. Re-Thinking Autism (London: Jessica Kingsley Publishing).

Murphy, Dominic. (2006). Psychiatry in the Scientific Image (Massachusetts: Massachusetts Institute of Technology).

Timini, Sami., Gardner, Niel. & McCabe, Bain. (2011). The Myth of Autism (Palgrave-McMillian).