British Society for the Philosophy of Science

Annual Conference Abstracts

Online

7–9 July 2021
Individual Papers .................................................................................................................. 3
Bakshi, Kabir ............................................................................................................................ 3
Baldissera, Marina with Suraje Dessai, Seamus Bradley, and David A. Stainforth .......... 5
Beasley, Charles ...................................................................................................................... 10
Belaustegui, Javier .................................................................................................................. 12
Bocchi, Federica ..................................................................................................................... 14
Boespflug, Mark .................................................................................................................... 16
Bondarenko, Olesya ............................................................................................................... 18
Bruckner, Michael ................................................................................................................. 20
Brzović, Zdenka and Predrag Šustar ..................................................................................... 22
Canali, Stefano with Simon Lohse ....................................................................................... 24
Chellappoo, Azita .................................................................................................................. 26
Chin-Yee, Benjamin .............................................................................................................. 28
Chua, Eugene Y.S. ................................................................................................................. 30
Clarke, Christopher ............................................................................................................... 32
Coraci, Davide with Gustavo Cevolani ................................................................................. 34
Curtis-Trudel, Andre ............................................................................................................. 36
Daoust, Louise ....................................................................................................................... 38
De Benedetto, Matteo with Michele Luchetti ..................................................................... 40
Duguid, Callum ..................................................................................................................... 41
Dupre, Gabe .......................................................................................................................... 43
Farr, Matt .............................................................................................................................. 45
Feldbacher-Escamilla, Christian J. with Alexander Gebharter ....................................... 48
Findl, Johannes and Javier Súarez ...................................................................................... 50
Fletcher, Samuel .................................................................................................................... 52
García-Lapeña, Alfonso ....................................................................................................... 53
Geddes, Alexander ............................................................................................................... 54
Greenwood, John ................................................................................................................... 56
Guralp, Genco ........................................................................................................................ 58
Heilmann, Conrad with Marta Szymanowska and Melissa Vergara Fernández ................ 61
Ioannidis, Stavros ................................................................................................................. 63
Irvine, Elizabeth .................................................................................................................... 65
Jolobeanu, Dana ................................................................. 67
Jarnicki, Pawel with Hajo Greif .............................................. 68
Jefferson, Anneli .................................................................. 70
Jhun, Jennifer .................................................................... 73
Jurjako, Marko .................................................................. 75
Kenna, Aaron .................................................................... 77
Khosrowi, Donal .................................................................. 80
Köiv, Riin ........................................................................... 82
Kononova, Viktoria .............................................................. 84
Larroulet Philippi, Cristian ................................................... 86
Leretere, Laurie .................................................................. 88
Lisciandra, Chiara ................................................................ 90
Lohse, Richard .................................................................. 92
López, Cristian .................................................................. 94
Lorenzetti, Lorenzo .............................................................. 95
McCullough-Benner, Colin ..................................................... 97
Pantazakos, Themistoklis ....................................................... 99
Pelayo, Areins .................................................................. 101
Pellet, François .................................................................. 104
Poth, Nina ........................................................................... 106
Pulkkinen, Karoliina ............................................................. 108
Rathkopf, Charles ................................................................ 109
Sagrafena, Cristina .............................................................. 110
Sarwar, Ameer .................................................................. 112
Stanley, Shaun .................................................................. 114
Stuart, Michael .................................................................. 116
Suárez, Mauricio ................................................................ 117
Thoma, Johanna .................................................................. 120
Tinklenberg, Brandon ............................................................ 121
Vassallo, Antonio with Pedro Naranjo .................................. 123
Veigl, Sophie Juliane ............................................................ 124
Vergara Fernández, Melissa with Conrad Heilmann and Marta Szymanowska .......................................................... 125
Vos, Bobby ........................................................................ 127
Zach, Martin ...................................................................... 129

Symposia .................................................................................. 132

The Aesthetics of Scientific Experiments ................................. 132
Analogy and Unity in Chemical Ontology ................................................................. 134
Conceptual and Evidential Issues in Contemporary Psychotherapy ...................... 136
Dependence in Physics ......................................................................................... 139
Epistemic Diversity and Argumentative Practices in Science ................................. 141
Evidential pluralism and Causality in the Special Sciences .................................... 143
Fictions in Science and Metaphysics .................................................................... 146
Model Transfer in Science .................................................................................. 148
Models in Particle Physics: Three Challenges to Realism ...................................... 150
Scientific Experts and the Pressures of Pandemic Policy Advice ............................ 152
SU(n) Surplusage: Symmetry, Gauge, and Equivalence ......................................... 154
Symmetry Principles ......................................................................................... 156
What We Have Learned about Memory (and What Remains to be Learned) ........ 158
The Better Best Systems Account (BBSA) is proffered as a refinement and improvement of the Best System Account (BSA) of laws. On the BSA, a (contingently) true generalization is a law of nature just in the case it is an axiom or a theorem of a deductive system which encodes the knowledge of an omniscient being and which a) (robustly) best balances strength, simplicity, and fit and b) contains only predicates that refer to perfectly natural properties (Lewis 1994). The proponents of the BBSA claim that it improves on the BSA on three major counts: the BBSA i) provides an anti-reductionist account of special science laws; ii) accommodates ceteris paribus laws in the special sciences; and iii) circumvents the commitment to Lewisian perfectly natural properties (Cohen and Callender 2009, Schrenk 2017).

Although the BBSA is an umbrella term for a variety of different approaches, the core idea undergirding all the approaches is similar: instead of running a best systems competition on a vocabulary containing only perfectly natural predicates, multiple best systems competitions are to be run, each corresponding to the vocabulary of a special science. The laws of biology are the axioms or theorems of the best deductive system which admits only biological properties. Similarly for laws of chemistry, laws of ecology, and so on. To accommodate ceteris paribus laws, Schrenk (2017) argues for a completer account of ceteris paribus laws wherein weakened strict generalizations, which exclude exceptional individuals in their antecedent, are allowed to compete in the best systems competition. Thus, on the BBSA, a (contingently) true generalization \((x)(Ax \land (x\neq e_1)\land (x\neq e_2)\land \cdots \land (x\neq e_n)\supset Bx)\) is a law of a special science \(S\) just in case it is an axiom or a theorem of a deductive system which encodes the knowledge of an omniscient being and which a) (robustly) best balances strength, simplicity, and fit and b) contains only predicates that refer to \(S\)-properties.

In this paper, I critically evaluate the BBSA and raise two worries for the account. The first is that the BBSA falls ill to a modified version of Hempel’s problem of provisos (Hempel 1988). I argue that on the BBSA, a special science law:

Horn 1: is either strictly false; or

Horn 2: if qualified as Schrenk suggests,

2.1) is either trivially true, i.e it is an identity statement; or

2.2) is vacuously true, i.e it is true in virtue of impossible antecedents

I show that taking either of the horns is unfavourable to the proponent of the BBSA. Take Boyle’s Law which Cohen and Callender (2009) consider. Read as a strict universal quantification, Boyle’s Law is false because there are many instances when the volume of a gas is not inversely proportional to its pressure (H1). Qualified by a cp clause, Boyle’s Law is either vacuously true because the cp conditions are never satisfied (H2.2) or it is trivially true because it states that the pressure and volume of a gas are inversely proportional unless the pressure and volume of a gas are not inversely proportional (H2.1). The proponent of the BBSA can neither accept H1 since laws are true
generalizations; nor can she accept H2.1 (or H2.2) because an identity (vacuous) statement might be simple but it cannot be strong.

The second worry is that the BBSA, as currently construed, is unreceptive towards the interdisciplinary nature of scientific practice. I argue for this failure by showing that the BBSA:

i) is unable to accommodate special science laws with cp clauses which are formulated in the vocabulary of other sciences. Such laws are the norm rather than the exception in scientific practice. Consider, for example, Engel’s Law in economics. In addition to cp clauses in the vocabulary of economics, Engel’s Law also contains cp clauses in the vocabulary of other sciences: absence of a sudden shift in human psychology (a cp clause in the vocabulary of psychology), absence of second law of thermodynamics violating behaviour in humans (thermodynamics), to name a few.

ii) provides incorrect analysis in cases where laws are formulated (exclusively) in the vocabulary of some other science. Consider, for example, Poiseuille’s law in physiology which relates blood flow to the difference between the arterial and the venous pressure. Poiseuille’s law of blood flow plays all the aspects of a law of physiology (predictive, inductive, explanatory, and counterfactual support), but is an instantiation of the general Poiseuille’s law in fluid dynamics.

I submit that the BBSA cannot accommodate such laws on pain of contradicting its central tenets.

However, I think that these worries are not impossible to overcome for the proponent of the BBSA. I suggest a modification of the BBSA that assuages these worries. In particular, I submit that an appeal to normality considerations while running the best systems competition for a special science not only makes it possible to correctly confer lawhood in cases such as Engel’s Law and Poiseuille’s Law, but also does not fall ill to the modified version of Hempel’s problem of provisos. I close the paper with a comment on the epistemology of special science laws by noting that normality consideration is a part of, what De Regt (2017) calls, meso-level intelligibility standard and that such considerations play an important part in scientific understanding.

References


Evaluating the quality of model-based regional climate information: the case of the UK Climate Projections 2018

Adapting to a changing climate is an increasingly urgent necessity. Anthropogenic greenhouse gas emissions have already caused about 1 °C of global warming, and even for the most optimistic mitigation scenarios, we are likely committed to 1.5 °C by 2030-2050 (IPCC, 2018). Informing the preparations needed to manage the risks, limit the damages and take advantage of the opportunities that arise in light of this changing climate is a grand challenge of climate change science (Moss et al. 2013).

The kind of long-term regional climate information that is increasingly important for making adaptation decisions varies in temporal and spatial resolution, and this information is usually derived from Global Climate models (GCMs). However, information about future changes in regional climate also comes with high degrees of uncertainty—an important element of the information given the high decision stakes of climate change adaptation.

Given these considerations, Baldissera Pacchetti et al. (in press) have proposed a quality assessment framework for evaluating the epistemic quality of regional climate information that intends to inform decision making. Evaluating the quality of this information is particularly important for information that is passed on to decision makers in the form of climate services. Baldissera et al. suggest that one desirable characteristic of this information is “epistemic reliability”, i.e. that the information about future climate and related probabilities suitably represent the likelihood of different realizations, and that there is an explanation of why this is the case. This understanding of reliability becomes important when statistical-empirical characterizations of reliability are not available to scientists, as is the case for long term climate projections (see e.g. Winsberg 2006 and Baldissera Pacchetti 2020).

Baldissera Pacchetti et al. identify five dimensions along which quality can be assessed: diversity, completeness, theory, adequacy for purpose and transparency. Diversity indicates that different types of evidence should be taken into account. Completeness indicates that diversity of evidence should be maximized. Theory evaluates the strength of the theoretical underpinning of the statement about future climate. This dimension is particularly important for giving epistemic reliability to a statement. Adequacy for purpose refers to the empirical adequacy required for the stated purpose of a model and its output (e.g. making predictions, assessing uncertainty, etc.). This dimension also explicitly requires that the purpose of the information should be made explicit. The last dimension is transparency, which requires that both the evidence and methodology be accessible enough for the other quality dimensions to be used to evaluate the information, even by non-experts.

In this paper, we critically evaluate this framework by applying it to one example of climate information for adaptation: the UK Climate Projections of 2018 (UKCP18). There are two main motivations for the choice of UKCP18. First, this product embodies some of the main modeling strategies that drive the field of climate science today. For example, the land projections produced by UKCP18 provide probabilistic uncertainty assessments using multi-model and perturbed physics ensembles (MME and PPE), use locally developed GCMs and the models from the international Climate Model Intercomparison Project (CMIP), perform dynamical downscaling for producing information at the regional scale and further fine grain information with convection permitting models (Lowe et al. 2018). Second, the earlier version of the UK Climate Projections (UKCP09) has received criticism from philosophers of science (Frigg et al. 2013, 2015). The quality assessment
framework proposed by Baldissera Pacchetti et al. partly aims to reveal whether the pitfalls identified by philosophers in UKCP09 persist in UKCP18.

The UKCP18 land projections have four main “strands” for which quality can be assessed, and each strand is representative of the driving methodologies in climate science described above. The probabilistic projections combine MME and PPE to provide a probabilistic estimate of the uncertainties tied to future changes in the climate. The global projections are a set of model projections that aim to maximize the range of possible future responses to anthropogenic forcing. The regional projections share some methodological details with the global projections but include dynamical downscaling using a PPE of regional climate models. Finally, the local projections further downscale the information produced by the regional projections with convection permitting models.

We apply the quality assessment framework to each strand of UKCP18 and illustrate whether and to what extent each of the above strands of the land projections satisfies the quality dimensions of the framework. When appropriate, we show whether quality varies depending on the variable of interest within a particular strand or across strands. For example, the theory quality dimension highlights that epistemic quality along this dimension is better satisfied for estimates about variables that depend on thermodynamic principles (e.g. global average temperature) than fluid dynamical theory (e.g. precipitation) (see, e.g., Risbey and O'Kane 2011) independently of the strand under assessment. We conclude that for those dimensions that can be evaluated, UKCP18 is not sufficiently epistemically reliable to provide information of high quality for all of the products provided.

The framework provided by Baldissera Pacchetti et al. falls short on one dimension. The adequacy for purpose dimension does not sufficiently account for the importance of overall adequacy for purpose, especially as formulated by philosophers of science (e.g. Parker 2020), and we suggest distinguishing between adequacy for purpose and empirical adequacy. Nevertheless, the framework is valuable in providing insights into the epistemic reliability of model-based climate information. It demonstrates, for example, that some of the criticisms raised by Frigg et al. (2015) are still valid for the probabilistic projections of UKCP18. Moreover, it highlights the importance of producers of information about future climate, especially when this information is model-based, giving users the tools to evaluate the epistemic reliability of the information.

References


Barack, David

Computation with Neural Manifolds

Recent research in cognitive neuroscience has uncovered neural manifolds in many tasks that play a central role in explanations of behavior. Revealed through the use of a range of dimensionality reduction techniques, these manifolds are entities in low-dimensional spaces embedded in high-dimensional neural spaces. Different studies provide different interpretations of these manifolds. In this paper, I explore a possible computational interpretation for the role of manifolds in cognition. I begin with a discussion of computation in the brain. Computation herein is understood as the reliable transformation of inputs into outputs by a processor in order to perform a function for a system. Particular computations are defined in terms of a series of inputs and states to states and outputs. Drawing on ideas prevalent in cognitive neuroscience, I present the neuronal model of neural computation. On this model, inputs and outputs in the brain are electrical signals sent between neurons, and the processors that transform these inputs into outputs and take on different states are the neurons themselves. I next review two recent studies that have revealed manifolds using distinct techniques. The first study investigated the neural mechanisms of motion and color perceptual decisions. This study revealed orthogonal neural manifolds in the prefrontal cortex in a low-dimensional task variable space. The reliance on presupposed task variables, however, might suggest that manifolds are the result of experimenter-imposed constraints on the interpretation of neural activity. The second study dispensed with such an assumption and instead a data-driven dimensionality reduction was performed on neural activity in the prefrontal cortex during a temporal interval estimation task. This study revealed a curved neural manifold that played a central computational explanatory role in how animal estimate such intervals. On the basis of these two studies, I argue that manifolds provide evidence for neural computations. Evidence for neural computations takes the form of the description of a series of states in the system that can be mapped on to the computation. I argue that neural manifolds are just such a series of states and, hence, constitute evidence for computation in parts of the brain. I then turn to an examination of what such evidence implies about the transformers in the brain, the parts of the system that transform inputs into outputs. I consider three such possible transformers: single neurons, neural populations, and manifolds themselves. The first two possibilities are consistent with the neuronal model of neural computation whereas the last is not. I argue that, granted manifolds as evidence for neural computation, neither neurons nor neural populations can be the parts that transformer inputs into outputs. I defend this surprising conclusion against several objection. I next turn to a discussion of the third possibility, that neural manifolds themselves are the relevant transformers. I describe one way that the position along a manifold might be decoded by another part of the brain to provide the output of a computation. However, despite this proposal for how a part of the brain might decode the output of a computation from the position along a manifold, manifolds do not possess the right properties to transform the inputs to outputs themselves. Manifolds are abstracta that do not possess spatiotemporal properties that persist over time to play the role of a transformer. Instead, the role of manifolds in transformations highlights the need for development of a concept of computation in the brain that relinquishes a role for transformers. This analysis matches the proposed role for manifolds in some studies and explains the evidential role of manifolds for computation in cognition.

Representative References:

https://doi.org/10.1007/s11023-017-9442-5


Beasley, Charles

Replication pluralism

While the literature around the replication crisis has seen an explosion over the past decade, the debates have been overwhelming focused on formal issues such as p-value reforms, or questionable practices such as p-hacking or HARK-ing (Colling and Szűcs 2018). Recently, however, there has been a clear turn towards a more foundational question: What is a replication? Answers to this question move beyond the mere categorization of practices of replication (Stroebe and Strack 2014) or arguments for the superiority of an existing version of replication, such as direct or conceptual replication (Pashler and Harris 2012, Simons 2014), and instead venture to provide novel and general accounts of what replication amounts to.

In section (1) I introduce the standard accounts of direct and conceptual replication before transitioning to an analysis of two more recent and novel accounts: the diagnostic account (Nosek and Errington 2020) and the resampling account (Machery 2020). I argue that, ultimately, both accounts reduce our ability to assess the quality of our evidence, as well as our ability to produce more reliable evidence moving forward. For this reason they should both be rejected.

On Nosek and Errington’s (2020) diagnostic account, one experiment is a replication of another if it is able to both increase and decrease one’s credence in a hypothesis from a prior experiment. If it can’t decrease one’s credence in a hypothesis, then it is considered to be a generalizability test; that is, a test of the scope of the application of the hypothesis that is being tested.

While the simplicity of this account is appealing, I argue that it should be rejected for the following five reasons. First, it captures cases that are far outside of what would normally count as replications, and thereby diminishes the significance of the claim that an experiment has been ‘replicated’. I illustrate this point with a case from analytic chemistry. Second, even if this issue was bracketed, exactly what types of evidence should be considered to be able to both increase and decrease one’s confidence in a prior evidence remains underspecified and potentially open to misuse. Third, the account is grounded in a distorted view of experimentation on which the scope of a hypothesis is determined in the process of replication rather than in process of hypothesis formation prior to an experiment being carried out. Fourth, it remains silent on cases in which one experiment is only able to reduce our confidence in the results from a target experiment. Fifth and finally, there are reasons to think that the ability of one experiment to only confirm another hypothesis is not reducible to mere generalization.

I then turn to Machery’s (2020) resampling account, on which experiments are divided into the following components: treatment, measurement, setting, and experimental units. In carrying out a given experiment, each component has a set of members that can be randomly sampled from. This captures the sense in which every experiment is fundamentally unique, while also allowing an identity relation between experiments to exist on a more coarse-grained level of description. An experiment is replicated on this account when at least one component is randomly ‘resampled’.

While, again, this account is appealing, it should be rejected for the following reasons. First, the independence of the experimental components, and the centrality of randomly sampling members of those components in the process of replication, places the account in a fatal bind. Either it excludes instances of replication that it should include, excludes instances of replication that it should include, or both. I illustrate this point by means of two case studies: one from developmental psychology and one in neuroscience.
I then go on to argue that it is not clear that the account is able to adequately capture the practice of experimentation, in so far as every component of an experiment is ‘randomly sampled’. This neglects the very real and relevant practice of determining features of an experiment beyond the population based upon convenience alone. I highlight the dangers of adopting the resampling account with a recent paper by Halina (2020), who uses it to argue post hoc that experiments in the non-human animal mindreading research program, have in fact been replicated. This is held to be the case even though most comparative psychologists working in the field would say that these studies have not been replicated.

In section (2) I introduce replication pluralism as a progress-generating alternative to all previously discussed accounts that is rooted in experimental practice. In doing so, I defend the in-principle impossibility of ever isolating a singular account of replication. I begin by disambiguating pluralism and highlighting three aspects of replication that one could be pluralist about: the content of replication, the practice of replication, and the normative aspect of replication. This paves the way for a specified introduction of replication pluralism. On the account which I endorse, one experiment X is a replication of an experiment Y, iff X is relevantly similar to Y, where relevant similarity is determined within the context of evaluating experiment X alone. I go on to further specify the account by making use of the previous disambiguation. In doing so I endorse and defend the following three claims:

First, the content of every aspect of an experiment can be varied in a replication. While this clashes with the standard accounts that were centered on direct versus conceptual replication, this claim is consistent with both the diagnostic and the resampling accounts.

Second, there is no singular practice of replication that is shared across contexts. This clashes with all accounts of replication discussed thus far. I support this claim by appealing to micro-replications and a case study on the in vitro binding assay (Güttinger 2018).

Third, there should be no singular practice of replication that is shared across contexts. In so far as experimentation and the practice of replication across the sciences has been shown to be various, and constraining it to a singular practice would be clearly deleterious, I argue that replication pluralism should be normatively adopted.
Belastegui, Javier

*Kant’s Duality Principle for Kinds*

Kinds, be they vernacular or scientific, seem to be hierarchically arranged. This order is plausibly due to the specificity relations that hold between kinds. It is an interesting question whether there are any informative principles that describe these specificity relations. One of the principles that can be found in the literature is the Hierarchy condition (Ellis, 2001), (Tobin, 2010). It says that kinds are arranged forming a tree-like pattern:

\[(H) \text{ If kinds } K \text{ and } K' \text{ overlap, then either } K < K' \text{ or } K' < K \text{ or } K = K'\]

This condition goes back to the famous Porphyry tree. It was reintroduced by (Thomason, 1969), who proposed that lattices satisfying it would provide an adequate algebraic model for kinds. However, several philosophers of science (Tobin, 2010), (Ruphy, 2010), (Hendry, 2015) have suggested counterexamples to this principle.

The aim of this talk is to show that Thomason’s proposal does not require appealing to (H) and that, moreover, dispensing with it provides us with a more plausible principle. In order to argue for this claim, I will make use of a more general algebraic model for kinds based on Wille’s Theory of Concept Lattices (Ganter; Wille, 1999). The basic assumption to be made is:

\[(\text{Minimal}) \text{ Each kind corresponds to a set of objects, the extension, and a set of natural attributes, the intension. The objects in the extension are exactly all of those that share all the natural attributes in the intension, whereas the natural attributes in the intension are exactly all of those shared by all the objects in the extension.}\]

This model for kinds will be argued to have several advantages over the one just mentioned. First, the hierarchy condition only holds in concept lattices under very special constraints. Generally speaking, the condition fails because most concept lattices are not trees. Second, the class of concept lattices is included into the class of structures that provides a complete semantics for Corcoran and Martin’s syllogistic logic (Martin, 1997). This implies that the basic relations described by the Aristotelian square of opposition, which plausibly apply to kinds, hold in the model. Finally, although hierarchy does not hold in these structures, it will be shown that an arguably more illuminating principle does, namely Kant’s Law of the Duality between Extension and Intension (Swoyer, 1995). This one says that:

\[(\text{Kant’s Law}) \text{ The extension of a kind is inversely related to its intension}\]

In other words, the more general a kind is, the fewer attributes will be common to its instances, and viceversa, the more specific a kind is, the more attributes its instances will share. Thus I suggest to replace the hierarchy condition by Kant’s Law as a more plausible principle concerning the specificity relations between kinds.

REFERENCES


In my talk, I argue that more successful conservation actions result when ecologists and policy makers endorse two theoretical and operational principles. 1. Processual ontology; 2. Deep Time literacy. In conservation biology and restoration ecology, thinking in terms of processual ontology amounts to conceptualizing ecosystems as unbalanced and nonstatic systems. In practice, this translates into specific protection strategies called “adaptive management”. Processual ontology clashes with the more traditional metaphysical view that dominated the field of ecology until recently, that has been called “balance of nature” paradigm. Deep time literacy, also known as “timefullness” (Bjornerud, 2018) refers to the awareness of the geological timescale. In the context of ecology, this can involve, among other things, the integration of paleodata in ecological models. That the past might help predict how the future will look like is not a triviality in conservation, a field traditionally future-oriented. Being ecologically time-literate translates into putting palaeontology, palaeocology and deep time research at work in investigating states of variability and resilience that an ecosystem can support and formulating thresholds of potential concerns for losing ecosystem services of interest. In a nutshell: After showing why processual ontology and deep-time literacy are irreducible to one another and complement each other, I will claim that endorsing 1 and 2, as both a conceptual and operational framework in biodiversity assessment and ecosystem management, results in a better conservation praxis. I will support my claim using a fascinating case study, the case of the Kruger National Park in South Africa. More in detail: In the first part of my talk, I identify the Kuhnian paradigm shift that the field of conservation biology has been undergoing since the ’70s. I claim that shift in paradigm is motivated by 1 and 2. Before the ’70s, conservation was conceived of as a science attempting to avoid species extinction and maintaining ecosystem stability by keeping humans out of the equation. The conceptual framework has been called the “balance of nature paradigm” (Gillson, 2015). The types of intervention admitted were centered around avoiding exceeding the carrying capacity limited by primary producers in the system of interest. Theoretical worries and practical considerations lead to a new conception of conservation’s study subject and its role, namely understanding and promoting the adapt-ability and functional integrity of whole ecosystems (Barnosky et al., 2017). Non-equilibrium ecology emerged as paradigm from realizing that no absolute steady state exists, but rather a morphospace of possible outcomes each favored by a set of unique variables. I therefore argue that conservation biology is becoming a process-centered enterprise, with a focus on measuring and conserving evolutionary potentialities. Using deep time data is increasingly helpful in identifying tipping points and thresholds that might raise concerns for ecosystem services of interest. The most effective associated conservation strategy is adaptive ecosystem management, in a framework that could be called a “process-based approach” to conservation ecology, in line with Daniel Nicholson and John Dupre’s (2018) biological processualism. In the second part of my talk, I turn to the case study of the Kruger National Park in South Africa, as evidence that conservation efforts are more successful if ecologists and policymakers endorse 1 and 2. The Kruger National Park was instituted in 1926 to reintegrate and protect the decimated elephant population in East South Africa. The most compelling aspect that the restoration strategy targeted was achieving the optimal balance between the elephant population and vegetation abundance. The initial management strategy demanded intensive human intervention, such as culling and fire suppression. This interventionist strategy soon revealed its limitation, requiring a cascade of additional intervention to maintain the desired balance. The main issue, as it turned out, was that the desired steady state imposed on the elephant population and based on the carrying capacity of the extant vegetation was arbitrary and did not account for long scale environmental dynamics. Since the ’90s, the Park administration has adopted the non-equilibrium ecology paradigm integrated with research in deep time ecology using proxies such as
fossil pollen and radiocarbon dating. The results, so far, are encouraging. I conclude by identifying areas for conservation and restoration that still underappreciate—but could greatly benefit from—the non equilibrium ecological paradigm, besides National Park management. My talk also reveals how conservation biology, philosophy and deep time research are intertwined and how environmental policy-making in general can greatly benefit from this interplay.
Herein, I argue that the scientific beliefs of non-scientists routinely amount to an exercise of faith; I take belief in anthropogenic climate change as a case in point. Yet, I suggest that perhaps there is a story to tell about why such beliefs are nonetheless epistemically permissible.

Even if it is not news to most philosophers that every lay person’s beliefs about climate change must somehow be based on the testimony of experts if they are to possess positive epistemic status at all, this tends to come as a shock to most people, given that such beliefs are scientific and, therefore, presumably based in some way on perception and/or reasoning. One interesting consequence of this is that beliefs about climate change technically qualify as a case of the exercise of faith—at least as faith has been used as a term of art among most philosophers for the 1,300 years spanning Augustine to Hume. For faith, classically speaking, is believing what one does not see to be the case for oneself, typically on the basis of testimony.

While the fact that the classical conception of faith may be applicable to belief in climate change may on its own be rather uninteresting, I go on to offer further reasons for taking belief in climate change to often embody other features commonly associated with faith. I utilize survey data and research in social psychology to show that, in addition to involving trust in the reports of scientific experts, beliefs about climate typically include: (1) a belief in propositions which could be independently verified to some degree, but such verification is typically not pursued, (2) often little corroborating evidence concerning reporters is sought prior to believing the deliverances of science, (3) consequently, lay trust in science would appear to manifest a degree of credence that markedly outstrips the believer’s evidence and (4) lay belief in climate change turns out to be a resilient kind of belief in that it is resistant to counterevidence. Here, I appropriate a helpful set of criteria offered by Elizabeth Anderson for making a “sound judgment” regarding expertise concerning climate change.

In underlining these features of lay-scientific belief in climate change, I am by no means attempting to cast doubt upon the deliverances of science. Rather, I am merely attempting to bring a surprising characteristic of belief about climate change to light—one which may be of considerable interest to the field of science communication as well as to the task of engendering public trust in scientific experts. If beliefs about climate change are more like faith commitments than beliefs carefully calibrated to the evidence, then our approach to getting skeptics to believe in climate change should perhaps involve methods that go beyond the mere presentation of scientific evidence.

Beyond this, I highlight a family of further features of belief in climate change that cause it to further resemble a characteristically religious kind of belief, and respond to two objections. The first objection is that I have unfairly countenanced the issue in rather stringent internalist, reductionist and individualist terms. The second is that I respond to the claim that a consensus among scientific experts is adequate evidence for belief in climate change. I conclude that non-scientists’ typical degree of credence in the reality of climate change—the mode credence of my study was 1, while the mean was .93—is often not warranted by their evidence. I close by sketching a case for why such belief may be rational after all. I suggest that a conviction that belief in climate change must be rational would seem to be more a consequence of moral sensibilities than sensibilities concerning epistemic entitlement, and that faith in science may, on moral grounds and in some qualified sense, possess warrant after all. This may be viewed as one sort of application of what Annette Baier called for nearly a half a century ago, namely, “a faith in the human community and its evolving procedures—in the prospects for many-handed cognitive ambitions and moral hopes.”
References


The recent history of social sciences is rife with examples of integrative projects involving various areas of biological research, such as “neuroeconomics”, “genopolitics” or “sociogenomics”. One feature of such projects is that insights about the underlying biology of individual behaviour are brought to bear on the aims of classification and explanation in the social sciences, occasionally leading to the revision of existing knowledge about behavioural phenomena (see Craver & Alexandrova, 2008). Recently, however, there have been attempts to leverage biological discoveries for investigation not only of behaviour itself but of the very social environment in which it is embedded. This paper focuses on the epistemic strategy which deploys the toolkit of genome-wide association studies (GWAS) in order to re-evaluate some of the available scientific knowledge about socio-environmental causes of behaviour in the light of genetic findings. Such a strategy has been proposed in the context of what is known as “sociogenomic” research, which seeks to address research questions in the social sciences with the help of cutting-edge genomic designs. The GWAS tools are thought to be instrumental for investigating the environment in three main ways: (1) identifying intermediate characteristics which are relevant to the outcomes of interest and may offer new clues as to the role of the environment; (2) shedding light on the endogenous (i.e. genetically mediated), rather than exogenous, character of specific environmental influences; (3) increasing the accuracy of causal estimates in the social sciences. It is expected that, by helping to achieve these three aims, the use of GWAS toolkit will strengthen causal inference about the contribution of environmental factors to a range of individual outcomes – from educational attainment to subjective well-being.

Even though all of the three aims are important for the understanding and evaluation of this strategy, in this paper I primarily examine the issues which pertain to the realisation of aim (1). In doing so, I make several claims, of which some are interpretive and some are critical. On the interpretive side, I highlight how the shift from “hypothesis-driven” candidate gene studies to “hypothesis-free” GWAS has afforded opportunities for the type of exploratory research that is agnostic not only to the biological mechanisms of gene action but also to the mechanisms of environmental influence at work in the outcomes of interest. As such, the sociogenomic research does not rely on the existing theoretical frameworks for its exploratory stage, which is seen by some (e.g. Freese 2018) as aiding the discovery of new knowledge about the environment. Secondly, I show how the project of learning about the environment through GWAS is tightly linked to the mapping of the identified genetic variants to intermediate psychological and cognitive phenotypes of individuals – a challenging task for which various methods, including the so-called phenotypic annotation, have been proposed.

I then proceed to make my main critical claim. I argue that, in trying to link the GWAS findings to intermediate characteristics of individuals, sociogenomic researchers have often relied on psychometric constructs which are unstable (sensu Sullivan (2016)) and subject to revision. Among other things, this instability means that such constructs, depending on how they are operationalised and measured, can support very different inferences about environmental factors, both those that contribute to the characteristics themselves and those that make them relevant to the studied outcomes. Therefore, the sociogenomic scientists should proceed with caution when bringing the information about intermediate characteristics to bear on their project of investigating environmental causes, as there may be alternative inferences that can be drawn from the same sets of genetic data. I conclude that sociogenomics would do well to articulate a strategy for dealing with the instability of the utilised constructs and to recognise the implications it may have for further research on the environment.
GWAS studies of individual outcomes are part and parcel of the scientific field of behavioural genetics which has received some attention within philosophy of science (Longino 2013; Plaisance & Reydon 2012; Schaffner 2016; Tabery 2014). Furthermore, philosophers and scientists have examined epistemic questions related to the use of GWAS as a scientific method (Reimers et al. 2019; Craver et al. 2020). I aim to contribute to this literature by documenting a new scientific application of the GWAS toolkit and opening up a new line of discussion, namely: what are the promises and the pitfalls of the epistemic strategy that uses genomic discoveries in order to gain insights into the environment?

References


The goal of this paper is to call into question a certain piece of Conventional Wisdom in the philosophy of cognitive science. It states that the research programs of extended cognition and embodied cognition are allies. This commonplace might seem surprising at first. After all, on a rough-and-ready characterization, the goal of embodied cognition is to situate cognitive processes firmly within the cognizer’s body, and in some cases the cognizer’s immediate environment. By contrast, on an equally rough-and-ready characterization, extended cognition seeks to find cognitive processes in vehicles outside the cognizer’s body that (thanks to the technological advances of the digital age) could be just about anywhere in physical space. Yet, conventional wisdom views them as allies.

One explanation for Conventional Wisdom is that extended cognitivists and embodied cognitivists alike have displayed a tendency to lead cognitive science away from the brain as the locus of cognition. They do so by endorsing versions of the claim that (as stated by Larry Shapiro) “the constituents of cognitive processes may extend beyond the brain, to include features of the body or world.” The starting point of my paper is the assumption that this shared commitment to extracranial constitution is the main motivation for Conventional Wisdom. From there, I cast doubt on Conventional Wisdom by arguing that extracranial constitution is only a superficial similarity, outweighed by dissimilarities. I offer two arguments to this effect.

My first, relatively tame argument against Conventional Wisdom aims to show that extended and embodied cognitivists are talking past each other, when it comes to extracranial constitution. This argument rests on two observations.

Observation 1 is that the extended cognitivists’ transcranial trajectory considerably overshoots that of the embodied cognitivists. In principle, their trajectory reaches however far from the cognizer’s brain they are able to plausibly locate vehicles (e.g. smartphones, personal assistants) that might still play functional roles in the cognitive process. While certain precursors and strands of embodied cognition also locate parts of the cognitive process in the cognizer’s immediate environment, they are much more limited than that.

Observation 2, which is more damning, is that extended cognitivists and embodied cognitivists are using different senses of “constitution.” When extended cognitivists say things like, “Otto’s notebook is a constituent of his cognitive system,” what they mean is, “Otto’s notebook is a part of his cognitive system.” But when embodied cognitivists say things like, “Sensorimotor activity is a constituent of concept possession,” what they mean is, “Sensorimotor activity is essential to concept possession.” The two senses differ most clearly in their modal profiles. (“A is essential to B” entails that B cannot exist without A. “A is a part of B” does not entail such dependence.) I further argue that they also convey different ideas.

Conjoining observations 1 and 2, I conclude that their respective commitments to extracranial constitution differs between extended and embodied cognitivists in both degree and kind. Thus, we should take this apparent similarity with a grain of salt.

My second, more confrontational argument against Conventional Wisdom aims to create tension between two claims: claim 1 states that certain concepts are essentially embodied, while claim 2 states that those same concepts are potentially outsourced. The reason why it matters that claims 1 and 2 are at tension is that certain embodied cognitivists explicitly argue for claim 1, while extended
cognitivists are under pressure to accept claim 2. This pressure stems from Clark and Chalmers’s own “parity principle,” which has served as a major asset for extended cognitivists. The principle states that we ought to treat any extra-cranial goings-on as part of the cognitive process, if we would treat them as such if they were done in the head. I argue that, by this very principle, concepts have an equally good claim to be capable of being outsourced into external vehicles as Clark and Chalmers’s own examples of memories and beliefs. Thus, extended cognitivists should accept that concepts are potentially outsourced. But if so, then they are not essentially embodied (pace embodied cognitivism of the stripe I have just described).

My ambition for neither argument I offer in this paper (the relatively tame one and the more confrontational one) is to knock down the extended/embodied alliance entirely. Rather, I am offering them both as templates for generating centers of friction. My only hope is that this will open up avenues of research that will help to further illuminate the lay of the land in post-standard cognitive science.

Bibliography


At which point does a nonfunctional genomic segment become functional? Or, more precisely, when do its activities acquire a fully functional character? That type of questions, to which we will refer in the present paper as temporal questions (TQs), are closely related to more discussed ones in the recent function debate; namely, the type of questions concerned with criteria for singling out functional elements of a genome, as instantiated by the “ENCODE controversy”.

TQs particularly pervade the biological theories of new genes origin: the gene duplication-divergence theory and, currently, the theory of de novo gene origin and functionality. In brief, according to the latter theory and its explanatory models, an initially noncoding, intergenic genomic sequence can eventually develop into a “proto-gene” as a result of noise in the standard functioning of the mechanism of protein synthesis. Proto-genes are those genomic entities that exhibit stable expression and/or translation, but not amounting to a proper function. The underlying idea here is that in order for a genomic entity to be classified as a gene, it has to display a determined function. The transition from proto-functional to functional status starts to take place when the produced peptides begin to provide adaptive advantage, and the new genomic entity starts to spread through a population. This new genomic entity is considered as functional together with the corresponding genomic regulatory networks.

As to our answer to TQs, we argue that ascribing a function to a genomic activity is not an all-or-nothing affair, but, rather, a transitioning one, involving distinct steps in a continuum between a mere fortuitous benefit and a fully functional activity of a genomic segment. Natural selection can start to act on the activity exhibited by a proto-gene, given that its protein-product provides adaptive advantage to the containing organism, as claimed by the de novo explanatory models. However, we are here still at the level of a proto-functionality. Namely, we are not entitled to ascribe a function to an entity at the moment in which a very first adaptive advantage is displayed, contrary to what the ahistorical accounts of functions claim. At this stage, this is still a ‘private’ genomic entity, because it belongs to a single individual organism. In order to consider something as a functional response to selection pressures, it must exhibit its activity with some degree of regularity. Think of a molecular object, most usually, a gene, RNA or a protein. Unless it repeatedly and with a high-level accuracy performs its activity in a cell it cannot be considered as a functional response, but a purely fortuitous, one-off event.

Even when a new gene has emerged, we do not ascribe to it an all or nothing functionality. Rather, the functionality of a certain gene varies depending on the robustness of mechanism of their expression. Regulation and expression of some genes are highly robust, because they are controlled by invariable expression programs. However, there are genes with higher levels of noise and different expression patterns from cell to cell and from individual to individual (see, for instance, MacNeil and Walhout 2011). Thus, the extent of functionality ascribed to a genomic segment corresponds to the robustness of the mechanistic arrangement leading up to the functional product.

This brings us to MQ, i.e., the general question of how we should best disentangle the relationship between the function and mechanism notions. As shown, the function ascription is directly related to the robustness of mechanisms producing the functional outcomes in question. Here it interesting to examine how features of mechanisms involved in the production of new genes, most notably, the mechanism of protein synthesis, are directly related to the above discussed transitioning functionality framework. The emergence of proto-functional elements in the genome is the result of a certain amount of stochasticity in the mechanism of protein synthesis, i.e., of transcriptional and
translational noise. But we can start to talk about functional outcomes only at the point when a relatively robust, new mechanistic arrangement has been established.

We argue, accordingly, that there is an intimate connection between functions and mechanisms. Not only, as claimed by Garson (2019), in the sense that mechanisms are for functions. In the general scientific area here under consideration, in order to ascribe a function, we should establish whether a process leading up to the corresponding functional outcome, is a mechanistic one. In sum, the degree of functionality ascribed to a genomic segment depends on how structured and resilient is the mechanism that leads to a determined protein product. In establishing a close connection between mechanisms and functions we do not have in mind the minimal, all-encompassing notion of mechanism. We characterize them, instead, as a distinct subset of causal processes buffered from external perturbations, thereby gaining causal structure and ensuring regularity in the outcome production.

Our proposal avoids one difficulty that etiological accounts of function generally face. Namely, how to account for the puzzling explanatory contribution of function ascriptions. According to critics’ portrayal of the etiological approach, two causally indistinguishable structures can differ with regards to their functionality. However, what is, then, explanatory about a certain functional attribution if its explanatoriness does not follow from the entities’ causal powers. We argue that the causal structures in question are not indistinguishable. First generation causal processes leading to a certain, potentially functional, trait are unstructured and fortuitous, because they have not yet been embedded into the existing genomic architecture, such as regulatory networks. The embedding starts with the working of natural selection and ensures the increase of causal insensitivity of the process from external perturbations. Thus, becoming a structured part of a larger organized system ensures the stability of the overall process and the regularity of the outcome.

References

Garson, J. (2019), What Biological Functions Are and Why They Matter (Cambridge, CUP);

Canali, Stefano with Simon Lohse

Interdisciplinary Knowledge Integration in Public Health Policy

During the COVID-19 pandemic, science has taken a central role as the provider of data and advice for policy-making. Several scientists have become key expert advisors and media pundits and the general public has become familiar with various tools and notions from epidemiology and biomedicine, such as epidemiological computer models, the basic reproduction number R0 and the phrase “flatten the curve”. Accordingly, scientific evidence and advice have had a direct influence on the implementation of lockdown measures to prevent the spreading of COVID-19 and on other long-term strategies to deal with the pandemic (Adam, 2020). Yet, the prominent role played by science has also been met with scepticism and critique, especially in relation to issues such as the objectivity of experts, the lack of independent scrutiny of scientific evidence, and the quality and reliability of available data (see, e.g., Ioannidis, 2020). Additionally, the extent to which policy interventions have been and are politically legitimate has been questioned in a variety of science-related ways. Some critics have pointed out that policy-making relied far too much on epidemiological modelling despite uncertain model assumptions. Others have criticised scientific taskforces for providing one-sided expertise and neglecting social aspects of the pandemic. Furthermore, competing and controversial views within science have been highlighted as a problem for the idea of “following the science”.

In this talk, we will contribute to these discussions through the lens of philosophy of the social sciences in practice. Our aim is to explore the potential of, and challenges to, a tighter integration of the social sciences in dealing with public health threats based on the example of the COVID-19 pandemic. More specifically, we will defend the twofold claim that the social sciences were marginalised in evidence-based public health policy in 2020 and that we should attempt to actively facilitate a more prominent role of the social sciences in similar public health crises in the future.

We will start by showing that public health measures and strategies that were meant to manage the pandemic in the EU and the UK were almost exclusively based on biomedical evidence, in particular epidemiological data and model projections (Manzo, 2020). We will draw on examples from Italy, Germany and the UK to show in what ways and to what extent evidence-based policy was dominated by biomedical expertise and evidence. This applies to the composition of pandemic task forces, policy recommendations, and the public justification of public health measures by politicians and public health experts, among other things. Most importantly, we will show that the social sciences – with the notable exception of economics – played only a marginal role in contributing to the evidence-base for dealing with the pandemic.

In part II of this talk, we will discuss aspects that the social sciences could have brought to the table to improve the management of the COVID-19 pandemic by public authorities and, in turn, should bring to the table to improve the management of future pandemics. We will discuss three ways by which the social sciences could contribute essential elements to evidence-based pandemic management. (a) The social sciences could improve pandemic surveillance, e.g. by deploying representative surveys that inform us about the influence of social structure on disease transmission dynamics. (b) They could contribute to better predictions of the likely effects of pandemics and public health strategies (lockdowns etc.), e.g. by enriching epidemiological projections with information about behavioural contact networks and behavioural patterns in different countries. (c) The social sciences could be highly useful in devising and increasing the effectiveness of public health measures, e.g. by using knowledge about socio-economically disadvantaged groups to think about stratified policy measures and communication strategies.
Our discussion in part II will provide a strong rationale for a stronger involvement of the social sciences in public health crises in the future. In part III, we will turn to key challenges to be tackled on the way to the realisation of such an involvement of the social sciences and to interdisciplinary knowledge integration in evidence-based public health policy. We will briefly highlight *sociological inhibitors*, such as epistemic hierarchies in science and different styles of public engagement of social and biomedical scientists. We will then focus our analysis on important *epistemic challenges* for a tighter integration of the social sciences. First, we will discuss conceptual issues concerning the notion of causality in public health and highlight an imbalance in taking into account material and social causes in this context. Next, we will focus on epistemological differences between biomedicine and the social sciences (such as the latter’s multidisciplinarity) that inhibit knowledge integration in evidence-based policy-making. Finally, we will explore methodological problems in selecting, gathering and integrating data on social phenomena for policy-making as an additional challenge.

In part IV of this talk, we will address a number of empirical and practical objections to our analysis that will at the same time help to further clarify our main claims. In the concluding part of our talk, we will summarise our key findings and suggest (tentative) normative implications of our discussion for improving policy-directed science and evidence-based policy, thereby contributing to the discussions in philosophy of science on the matter (Cartwright & Hardie, 2012).

References


Chellappoo, Azita

Can a microbiome be ‘obesogenic’?

Research into the human microbiome, the collective term for the communities of microorganisms that live inside and upon our bodies, has proliferated over the past two decades. The advent of techniques such as metagenomics that allow for analysis of microbial communities, including taxa that cannot be easily cultured, has enabled this rapid scientific growth. This has led to an increasing understanding of the ways that microbes contribute to and are intertwined with the development and function of a range of bodily processes, including digestion, metabolism, immune system function, and cognition. Coupled with these developments is huge translational potential in terms of tackling questions of health and disease, including the promise of effective therapeutic interventions addressing complex conditions or diseases, such as obesity, depression and autism.

However, there has been growing debate over how we should interpret the findings of this rapidly progressing body of work. Several scholars have critiqued the appropriateness of causal claims pertaining to the whole microbiome, arguing that understanding the entire microbiome, rather than particular microbial taxa, as a causal agent leading to particular health outcomes is not supported by the current evidence, and frequently fails to meet criteria of strong causal explanations under interventionist accounts of causation in particular (Bourrat 2018; Lynch et al 2019). Others have pushed back against this, pointing to reasons for considering the entire microbial community and its potentially emergent properties as causally relevant, at least in certain contexts.

This debate shows that the question of what precisely is the causal agent in microbiome research is central to how we should interpret the findings of this research, what future avenues of inquiry should be opened up, and how these findings should be translated into interventions. Here, I focus on research concerning the microbiome and obesity, and the construction of the notion of the obesogenic microbiome in particular. Obesity has been seen by many in public health and biomedical research as a growing ‘epidemic’ with increasingly severe health consequences for the world’s population, although other scholars have been critical of the medicalisation of obesity and the ways in which mainstream approaches contribute to weight stigma. Regardless of perspective, it is certainly clear that developing therapeutic interventions that effectively target obesity are in high demand, and could be extremely lucrative. It is therefore no surprise that a significant strand of human microbiome research investigates the causal role of the microbiome in generating obese or lean phenotypes, with the hope of developing such interventions.

Early findings from explorations of the relationship between the gut microbiome and obesity in both mouse models and humans indicated that differences in microbiome composition at the phylum level (the Firmicutes to Bacteroidetes ration in particular) play a role in determining whether the host phenotype is obese or lean (Turnbaugh et al 2006). This work prompted the construction of the notion of the ‘obesogenic microbiome’. Further research has explored the effects of more fine-grained differences in community composition, the causal role of particular microbial taxa, and mechanistic evidence for microbial effects on metabolism (Tseng et al 2019). However, the concept of the obesogenic microbiome persists in both scientific research and popular science narratives, and in the early development of interventions that target the microbiome in its entirety.

I argue that the concept of an ‘obesogenic microbiome’ is problematic in two key respects, and further clarification and conceptual modification is needed to understand the role that microbiota play in obesity, and therefore what this might mean for microbiota-targeted interventions. Firstly, there is a problem with the way in which obesity is conceptualised and operationalised in microbiome research. I offer an account of this operationalisation, where either BMI and/or a
particular kind of diet are frequently used as proxies for obesity, despite issues with the relationship between these proxies and the purported microbial target of intervention. A deeper problem comes from the way in which obesity is treated as a single unified category in this research, which inherently carries increased risk of disease, or as a disease in itself, despite the heterogeneity of its biological causes and consequences. I argue that it is a mistake to treat obesity in this way, and this compromises the capacity of microbiome research to provide strong causal connections and develop effective interventions. I suggest we should separate obesity, or fatness, understood as a somewhat unified social category, from the diverse and heterogeneous nature of the biological phenomena that are relevant to health consequences. Furthermore, the way in which obesity is understood in these studies affects the quality of explanations offered up by microbiome research insofar as it affects their proportionality (in both the first and second sense as described by Lynch et al 2019).

The second problem I highlight is the way in which the microbiome (typically the gut microbiome in particular) is implicitly treated as an organ of the body in this area of microbiome research, which can be manipulated with a significant degree of independence from the host and the environment in order to produce particular host phenotypes. I suggest that this framing of causal independence precludes a more integrated understanding of the relationship between the environment, host and microbiome, that encompasses social relations and individual agency. In particular, it becomes a barrier to conceptual integration with analyses of the social environments that contribute to disease risk, including the effects of discrimination.

I therefore argue that these issues with the concept of the ‘obesogenic microbiome’ mean that we should set it aside entirely. Rather, we should embrace a finer-grained understanding of the causal relationship between the microbiome and obesity, as well as allowing for conceptual integration between the microbiome, host, and environment. Furthermore, this case has broader implications both for microbiome research that seeks to develop insights into and therapeutics for poorly understood states, conditions or diseases, and for our understanding of what makes something ‘obesogenic’ more generally, including how we should think about obesogenic environments.

References


Chin-Yee, Benjamin

On the Uses and Abuses of Biomarkers in Genomic Medicine

Consider a list of typical sentences that might be heard during clinical rounds on the haematology-oncology service: ‘This is a 70-year-old male with myelodysplastic syndrome with 5q deletion;’ ‘this is a 50-year-old female with acute myeloid leukaemia with complex karyotype;’ ‘this is a case of chronic lymphocytic leukaemia with TP53 mutation.’ Each ostensibly thin description contains key genomic information for clinical judgment, carrying specific diagnostic, prognostic and therapeutic considerations. Together, these sentences signal the arrival of genomics in routine medical practice, particularly in oncology where genomic technologies are having a profound impact in reshaping diagnostic categories and therapeutic approaches.

Genomic medicine, along with related movements in precision medicine, have lately attracted considerable philosophical attention. Philosophers have identified a range of problems, from defining ‘precision’ medicine (Lemoine, 2017) to evaluating mechanistic evidence (Clarke et al., 2014; Williamson, 2019) and biomarkers (Hey, 2015; Plutynski 2020). The ontology of genomic biomarkers and their role in clinical epistemology, however, remains underexplored. Some have raised concerns over issues of ‘epistemic capture’ and ‘value partitioning’ arising from genomic technologies, whereby ambiguous information is treated as definite knowledge and classified along inappropriately simplistic scales, namely ‘positive’ versus ‘negative’ (Reynolds, 2020). The result, it is argued, is a form of epistemic injustice, harming patients in their capacity as knowers (Fricker, 2007; Carel and Kidd, 2014). This paper seeks to extend this analysis, focusing on three interrelated philosophical problems for genomic medicine: (1) the ontological status of genomic biomarkers; (2) their evidentiary role in clinical epistemology; and (3) the implications for epistemic injustice in patient care.

My first argument addresses problem (1), the ontology of genomic biomarkers. I argue that the phenomena of ‘epistemic capture’ and ‘value partitioning’ arising from genomic technologies (Reynolds, 2020) are better understood as outcomes of ‘pernicious reification,’ according to the abstraction-ontologizing account recently put forward by Rasmus Winther (2014; 2020), inspired by the pragmatist writings of James and Dewey (1929). This account offers important advantages for understanding genomic biomarkers, highlighting how biomarkers form abstract representations, which are rendered concrete and ontologized in clinical practice. These processes of abstraction and ontologizing can be inferentially generative and play important roles in clinical reasoning. Ontologizing, however, carries an inherent risk of reification. Following Winther (2020), I argue that reification is best avoided through careful attention to context and consideration of a concept’s history, function, and appropriate analytical-level, which together can help underwrite ‘contextual objectivity’ for a given ontology. Applying this view to the ontology of biomarkers, I contend, allows us to move beyond a one-sided focus on the epistemic harms posed by genomic technologies, maintaining sensitivity to the risks of reification while enabling epistemically fruitful abstraction-ontologizing practices.

My second argument focuses on problem (2), the evidentiary role of biomarkers. I examine epistemic claims invoked on the basis of biomarkers in clinical practice. For example, the above-stated sentence, ‘this is a case of chronic lymphocytic leukaemia with TP53 mutation,’ might be followed by the claim ‘TP53 mutation is associated with poor outcomes in chronic lymphocytic leukaemia.’ Although such claims are often interpreted as objective, value-free statements about a patient’s disease, I argue that they are better understood as mixed claims as defined by Alexandrova (2018), combining empirical hypotheses about causal or statistical relations with value-laden variables. While empirically supported biomarker claims can serve as important evidence in clinical
judgment, attention must be directed at their value-laden content, which arises from various sources, from defining methods of detection and thresholds of ‘positivity’ to correlation with ‘clinically significant’ outcomes (Chin-Yee et al., 2020). I consider Alexandrova’s (2018) proposed ‘rules’ for evaluating mixed claims, which call for close examination of value presuppositions in measurement and emphasize robustness across epistemic communities. I adapt these considerations for the appraisal of biomarker claims in clinical practice, and defend them against other approaches to ‘biomarker qualification’ (Leptak et al., 2017), which I argue are primarily designed for regulatory purposes and carry limited clinical import.

My final argument follows from the first two, and addresses problem (3), the implications of genomic biomarkers for epistemic injustice in patient care. While I agree with Reynolds (2020) that genomic technologies can contribute to epistemic injustice, I argue that this outcome is not inevitable, and a closer analysis of the sources of epistemic injustice might help remedy the situation. I locate two main sources of epistemic injustice in genomic medicine: firstly, inattentive ontologizing practices resulting in pernicious reification of genomic biomarkers, as discussed in problem (1); secondly, failure to recognize biomarker claims as having mixed empirical and evaluative content, and/or lack of sufficient appraisal of value-laden measures, as discussed in problem (2).

I conclude with a case study from haematology-oncology which illustrates these sources of epistemic injustice and shows how misplaced concreteness through reification combined with concealment of value-laden dimensions of biomarker claims can result in significant epistemic harms for patients. I demonstrate how a ‘contextually objective’ biomarker ontology and awareness of the value-ladenness of biomarker evidence might help mitigate these harms. Through this case study, I further clarify my position and address potential counterarguments, in particular causal and mechanistic interpretations of biomarkers. Although at times helpful in clinical reasoning, I argue that these accounts also encounter limitations, which I characterize as stemming from abstraction without ontologizing. I contend that the pragmatic approach to biomarker ontology and epistemology advanced in this paper is not only descriptively accurate of expert clinical judgment but moreover offers a normative account for using, and not abusing, biomarkers in the era of genomic medicine.
Chua, Eugene Y.S

Degeneration and Entropy

I. Degeneration

Lakatos’s (1976/2015) goal in Proofs and Refutations (P&R) was to argue that the comprehension of mathematical concepts must be accompanied with a clear understanding of how and why the concept came about. We must understand the concept’s problem-situation, the questions which led to the concept’s genesis and evolution, alongside an understanding of the concept itself. A concept is not an atom. Instead, a concept is a temporally extended process through which the initial, primitive, concept is continually refined through past iterations. To rip a concept apart from its context of discovery is to miss a complete understanding of the concept.

All of the above – concerning how to comprehend a concept – has been much discussed over the last few decades. What has been less discussed, to my knowledge, is how to evaluate a concept according to P&R: how do we know whether a concept is problematic, needs rehabilitation, or, worse still, must be abandoned, given the heuristic approach? In short, how do we know whether a concept is degenerating?

The primary goal here is to provide and motivate an account of degeneration based on P&R, and to show how this account of degeneration differs from – but can be augmented with – the account found in Lakatos’s later work in the Methodology of Scientific Research Programmes (MSRP). Based on my reading of P&R, I propose two criteria for degeneration: (i) superfluity, which involves the introduction of trivial extensions or terminology into a theory or concept, and (ii) authoritarianism, the introduction and employment of concepts into discourse without justification while ignoring the problem-situation of said concept.

On the one hand, superfluity reflects a lack of concern for a concept’s problem-situation. Terminology is produced because simply one can. Instead of considering what problems the terminology is meant to resolve, we are instead pursuing ‘cheap, shallow generalisations’ (1976/2015, fn. 160) because it would be nice even if the result is ultimately trivial. By failing to grasp what is trivial, research degenerates by either treading trodden grounds or extending a concept to domains which are simply unfruitful.

On the other hand, authoritarianism fails to account for past problems and errors by tearing apart present discussions from past problems – the discussion is presented mystically ‘as is’, without context. This obscures the errors and problems that were crucial in generating the proof-generated concept: we are “rewriting history to purge it from error” (1976/2015, 49) and hence “the zig-zag of discovery cannot be discerned in the end-product”. (1976/2015, 44) The degenerate concept becomes atomized as a result.

Hence, I propose an extended account of Lakatosian degeneration accounting for both content and depth. Much ink has been spilled over Lakatos’s account in MSRP, which understands the progress of a theory in terms of theoretical and empirical progress. In my view, this account of progress perfectly complements my new account based on the P&R. While the account of degeneration in MSRP focused on content, the account of degeneration in P&R focused on depth – how deep or trivial the research is, how it connects with its predecessors, the potential or actual fruitfulness of the research, and so on (discussed above in II) – which in turn hinges on methodology.
II. Entropy

I turn to the second goal, which is to apply this account of degeneration to the task of evaluating the historical trajectory of entropy. The concept of entropy has a tumultuous past. This is coupled, however, with its extensive usage (contentiously, I must add) in countless sciences. This makes entropy an interesting case study for degeneration. Particularly, I evaluate the transition from thermodynamic to information-theoretic interpretations of entropy by critiquing Jaynes’s (1957) landmark paper on thermodynamics and information, where he famously proposed the adoption of maximum entropy (MAXENT) as an information-theoretic principle of rationality. This, he claims, requires a subjective interpretation of statistical mechanics and entropy.

I argue that this transition suffered from both superfluity and authoritarianism.

Firstly, the discussion of interpretation is simply superfluous in the context of Jaynes’s paper. Shannon information, and hence MAXENT, is neutral between the objective and subjective interpretations of probability. Furthermore, the MAXENT proposal is just a useful piece of computation for calculating predictions: the actual goal of the proposal isn’t even about interpretation or the metaphysics of statistical mechanics. This is exemplified by the fact that while Jaynes purportedly proposes a subjectivist ‘interpretation’ of statistical mechanics, he still claims that ‘objective statistical mechanics’ is still needed for actual interpretation. In fact, his take on the interpretation of probabilities about the actual physical system remains an objectivist one, one seemingly adopting some version of the ergodic hypothesis! (627) This goes to show that the probabilities prescribed by statistical mechanics about actual systems – and their interpretations – are not even in question here.

Secondly, Jaynes displays an authoritarian attitude when simply presenting the information-theoretic interpretation, the subjectivist interpretation, and MAXENT, as though they must be taken altogether. Throughout the paper, Jaynes insists that the subjectivist approach is necessary. However, two questions arise: first, why the downplaying of ‘physical hypotheses’ used by ‘objective statistical mechanics’ and why do we need to ‘free’ statistical mechanics from them? (621) Second, why the focus on prediction and information theory, and the downplaying of interpretation? Both questions are unfortunately unanswered. The paper, in introducing these concepts and ignoring others, thus demonstrates a clear ignorance of the problem-situation of statistical mechanics set up by Boltzmann (e.g. (1896/1995, 5) and Gibbs (e.g. 1902, ix), with a clear emphasis on interpreting physical systems, and interpreting the connection between statistical mechanics and thermodynamics.

Finally, since Jaynes himself acknowledged that MAXENT can provide no new predictions (624–625), the proposal is both theoretically and empirically degenerative. Overall, then, Jaynes’s paper is degenerative both with respect to depth (falling afoul of superfluity and authoritarianism), and content (empirical and theoretical degeneration). It is thus degenerative tout court.
Clarke, Christopher

How does process tracing work in political science?

King Keohane and Verba (1994) was a field-defining manifesto about causal inference in political science. In it the authors argued for what I will call monism: although there are two traditions in political science---quantitative and qualitative---there is only one logic of causal inference, namely that uncovered by quantitative methodologists. Qualitative research can be good, but only insofar as it mimics quantitative research.

In response, most qualitative scholars in political science embraced dualism: there are two distinct logics of causal inference in political science---one quantitative and one qualitative---and perhaps even two wholly distinct types of causation (Brady and Collier 2010); though see Runhardt (2015) for disagreement. And their exemplar of a qualitative method with a distinct logic is process tracing. So process tracing sits at the centre of one of the big methodological disputes within political science.

It is commonly accepted that the logic behind process tracing in political science is given by Bayes rule (Fairfield and Charman 2017). But Bayes rule is arguably the logic behind *all* good scientific inference, be it quantitative or qualitative, contrary to what “frequentists” about quantitative methods claim (Howson and Urbach 1989/2006). And advocates of process tracing claim that process tracing is a qualitative method whose logic is distinct from that of standard quantitative methods (Brady and Collier 2010). So what then is this distinctive logic? Given three assumptions, I argue that there are only three ways in which process tracing could possibly distinguish itself from standard quantitative methods:

1) Provide a distinct way of using within-case evidence to learn about the “propensities” attached to a single case;

2) Provide some distinct principles / ways in which detailed causal hypotheses (about a single case) make predictions about the propensities attached to that case;

3) Provide a distinct way of leveraging cross-case evidence when discriminating between causal hypotheses about a single case.

I then contend that (1) and (2) do not seem to be what process tracers do. In relation to (2), for example, process tracers do not seek to deny (or to provide alternatives to) quantitative principles such as the principle of the common cause, or the principle of the common effect, or the principle of faithfulness. This leaves only (3) as the way in which process tracing can distinguish itself. This is ironic, because the methodologists who have done most to clarify and to defend the logic behind process tracing usually insist that process tracing is a “within case” method, thus playing down the reliance of process tracing upon cross-case evidence (Beach and Pedersen 2016).

The assumptions I need to make this argument are Bayes rule itself, plus:

- David Lewis’ Principal Principle (1980): For any set of variables, given perfect knowledge of the propensities governing these variables (in a particular case), all other available information is redundant to predicting in advance the values that these variables will take (in this particular case). This principle thus relates credences (rational degrees of confidence) to propensities (probabilities as features of the world).

- Causation can be inferred, but it cannot be directly observed.
These assumptions are minimal. To make this argument, I will not need to take a position on:

- the nature of variable causation (Interventionist / Difference Making / Mechanistic / Productive / Energy Transfer / Primitive)

- whether the cases one examines are governed by the *same* propensities, and the *same* causal relations between variables (Homogeneity / Heterogeneity)

- what hypotheses about variable causation imply about propensities (see above).

Thus this argument still applies even if (as qualitative methodologists suggest) there are two distinct types of causation in the world, one type “difference making” that is best studied by quantitative methods, and another type “mechanistic production” that is best suited to being studied by qualitative methods.

I then illustrate this abstract argument with a concrete suggestion about how process tracing might leverage cross-case evidence in a distinctive way: process tracing makes a much more subtle and sophisticated set of “causal homogeneity” assumptions than standard quantitative methods do. This is not obvious because methodologists have played down the reliance of process tracing on background knowledge about other cases, I suggest.

The upshot is that at a very high level of abstraction there is indeed (as King Keohane and Verba suggest) only one logic of causal inference, but that at a lower level of abstraction there are multiple logics of causal inference.

References


Reverse inference (RI) is a crucial inferential strategy widely employed in cognitive neuroscience to derive conclusions about the engagement of cognitive processes from patterns of brain activation. Despite its central role, in recent years RI faced increasing skepticism, especially after the influential critique advanced by leading neuroscientist Russell Poldrack (2006). Poldrack’s paper triggered a hot debate, but no convincing solution seems still on offer. In this paper, we assess the discussion so far and propose Bayesian confirmation theory as a new way to advance the debate.

In a typical neuroimaging study, the experimenter observes a pattern of neural activation $A$ in subjects performing a task and infers the likely engagement of some cognitive process $C$. This “reverse” inference is usually based on the (meta-)analysis of previous studies showing that in experimental tasks engaging $C$, activation $A$ is systematically observed. However, the status of RI is controversial for several reasons. First, RI is deductively invalid, reflecting the well-known fallacy of “affirming the consequent”; at best, it can be defended as an instance of abductive reasoning (Poldrack 2006). Second, a crucial weakness of RI has to do with the problem of (lack of) selectivity of the neural response: i.e., the fact that the same brain region is usually activated by different cognitive processes. Therefore, a RI relying on the activation of a lowly specific region will provide only weak evidence in favor of its conclusion. Third, the current lack of a satisfying cognitive ontology, i.e., a complete and non-problematic mapping among observed brain activations, engaged mental processes and experimental manipulations make RI still more problematic.

Both philosophers and neuroscientists have recently advanced a few proposals to address these weaknesses and to make RI more reliable. For instance, Poldrack (2006) himself reconstructs and defends RI in a broadly Bayesian framework, while Machery (2014) addresses the problems of RI by formalizing it in purely “likelihoodist” terms. Others focus on task-relativization as a way of dealing with the problem of the lack of selectivity of neural responses (see Nathan and Del Pinal 2017 for a survey). In the meantime, projects like BrainMap and Neurosynth offer a tool for the practicing neuroscientist to improve the reliability of RI through systematic meta-analyses and the use of machine learning techniques. Still, such techniques also have shortcomings and often leave the basic conceptual issues essentially unresolved.

In this paper, we argue that Bayesian confirmation theory, as developed within formal philosophy of science, offers the right conceptual framework to deal with the issues raised by RI. As suggested by Poldrack (2006), one can apply the Bayesian machinery to compute the probability $p(C|A)$ of engagement of cognitive process $C$ conditional on some activation pattern $A$. In this connection, it is however crucially important to recall a well-known, if often overlooked, distinction: that between posterior probability and confirmation (or inductive support). In general, a hypothesis $C$ is confirmed by some evidence $A$ when $p(C|A)$ is (much) higher than $p(C)$, i.e., when observing $A$ makes the probability of $C$ increase. Importantly, assessments of high confirmation and of high posterior probability are independent from each other, and may indeed point in different directions. As we argue, paying attention to such distinction can illuminate the role of RI in neuroscientific research, as we show both with theoretical arguments and with case-studies from neuroscientific practice. We proceed as follows.

First, we recap some basic results from the Bayesian analysis of scientific inference and of abductive reasoning in particular (Niiniluoto 2018). Let us say that the engagement of cognitive process $C$ inductively explains the pattern of neural activation $A$ if $A$ is more expected when $C$ is taken into account than if not; then one can prove that if $C$ inductively explains $A$, then $A$ confirms $C$. (Niiniluoto
Accordingly, RI can be construed as an abductive inference which lends credibility to its conclusion (the engagement of process C) given the success of C in explaining the observed activation A. By applying different measures of “explanatory power” recently discussed in the philosophical literature, we assess the conditions under which RI can be indeed defended in this way.

Second, we analyze and discuss current neuroscientific practice of using Bayes Factors (BF) for quantifying the selectivity of competing hypotheses and hence the strength of the corresponding RIs (Poldrack 2006; Cauda et al. 2019). Interestingly, BF is equivalent to some well-known confirmation measures discussed in the logico-philosophical literature. This suggests that neuroscientists actually tackle the selectivity issue by choosing the most confirmed among the competing cognitive hypotheses. We test this hypothesis by discussing real case-studies from current neuroscientific research, thus disentangling the different conceptual roles that posterior probability and confirmation play in RI.

Finally, we show how our analysis also helps in systematizing and improving current discussion concerning robust inferential strategies in cognitive neuroscience besides RI. We conclude by suggesting that Bayesian confirmation theory is a promising way to address a number of methodological issues in this and related fields.

References


Much recent philosophical work on physical computation addresses questions about computational individuation. Of particular interest is how to type individuate physical computing systems according to the mathematical tasks they perform. But efforts in this direction are stymied by the fact that physical systems seem to simultaneously perform a wide variety of apparently incompatible mathematical tasks. This has come to be known as the problem of simultaneous implementation.

In a bit more detail, the problem can be characterized as an inconsistent triad. The first claim is an observation about practice in the computational sciences (chiefly computer and cognitive science): (1) Computational scientists type individuate computing systems, in at least some cases, according to the tasks they perform. Typical tasks are things like computing a function or executing an algorithm. Computational tasks can be characterized in terms of the notion of implementation. Very roughly, implementation is a relation between an abstract mathematical computation and a physical system, where a physical system implements a computation just in case the computation accurately describes that system. For instance, a system might implement a Turing machine for addition, and we say that the device performs the task of addition.

In light of this, (1) this can be reframed as: (2) Computational scientists type individuate physical systems, at least in some cases, according to the computations they implement. The next claim connects computational individuation to computational explanation: (3) Successful computational explanation requires that a physical system fall under a unique computational type/implement a unique computation at a given time. This is plausibly supported by explanatory practice in the computational sciences, at least on the face of it. Trouble emerges however with the following claim: (4) Physical systems simultaneously implement a variety of distinct computations. (4) is an apparent consequence of many plausible accounts of implementation.

By (3), computational explanation requires that we be able to uniquely type identify a physical computing system. By (1)/(2), this involves identifying a unique computation implemented by that system. But by (4), there is no such computation. Under the plausible assumption that computational explanations of the sort mentioned above do *not* fail --- if they did, the computational sciences would presumably be in a much worse way than they are --- it seems that one of the above claims has to go.

Perhaps the most popular response to the problem is to deny (4). According to philosophers who deny (4), all that the problem shows is that we must tighten up the definition of computational implementation so as to avoid simultaneous implementation. There are different proposals about how to do this. Some philosophers endorse a semantic account of implementation. Others endorse a mechanistic account. Yet others endorse an account that foregrounds teleological functions. And of course mixtures are possible. However, these approaches face challenges of their own. One point to notice is that they avoid simultaneous implementation by imposing uniform implementation conditions for every computation. Yet it is far from clear that such a uniform account can capture the variety of computing systems investigated in contemporary computer science. This issue has generated substantial debate, and little in the way of agreement (see Shagrir (2020) for the state of the debate).

In any event, whether or not that debate is resolved, I think that these responses are ultimately beside the point. For the problem of simultaneous implementation is no problem at all. In my view, (4) is true, but it is true in a way that doesn’t undermine (3).
Central to computational implementation is the notion of a grouping scheme. Grouping schemes capture how the abstract, mathematical states of a computation are applied to the states of physical systems. Strictly speaking, we shouldn't say that a system implements a computation simpliciter. Rather, we should say that a system implements a computation relative to some grouping scheme.

How does this help? When a system simultaneously implements different computations, it does so relative to different grouping schemes. But computational explanation requires only that we explain with respect to a particular implemented computation, or at least a class of related computations. It doesn't also require that that computation be the only one implemented by a system at a time. In a way, this is not so surprising. Description of physical systems in terms of implemented computations is an instance of the more general practice of applying mathematics to physical systems. It is a familiar observation that one and the same physical system simultaneously falls under a wide variety of different mathematical descriptions. To take Frege's well-worn example, a single hunk of matter may simultaneously constitute one deck and fifty-two cards. Applying some branch of mathematics to a physical system, such as arithmetic or, in our case, computability theory, requires that we specify *how* that branch is to be applied. Finite cardinals apply under sortal concepts, such as 'deck' or 'card'. And computations apply under grouping schemes. In both cases, different ways of applying mathematical objects may yield different mathematical descriptions of physical systems. But descriptions which are on the face of it incompatible turn out to be compatible, once we realize that they arise from different ways of applying the branch of mathematics in question. Distinct mathematical descriptions of physical systems can apply simultaneously, yet in harmony.

(4) is true, but it is true only in the sense that systems simultaneously implements different computations relative to different grouping schemes. This is no threat to (3). Computational explanation requires only that we specify a computation relative to some grouping scheme. It does not require that there be only one computation or grouping scheme possible. So the problem of simultaneous implementation is no problem at all.

Bibliography


- Fresco, Nir, and Milkowski, Marcin. 2019. "Mechanistic Computational Individuation without Biting the Bullet."

- Shagrir, Oron, "In defense of the semantic view of computation," Synthese 197:4083-4180.
Daoust, Louise

Representationalism, Size Perception, and the Challenge from Deep Fovea

Standard representationalist views of perception purport to cast off the intellectualist baggage of twentieth-century thinking, and to address perception in its own distinctive terms. Focusing on spatial perception, I show that it is standard for these approaches to aim to reduce spatial aspects of the percept, such as apparent size or shape, to mind-independent geometrical facts about the object-perceiver relation, such as viewing distance and angle of elevation of an object viewed. This focus on reduction is tied to an assumption that successful perception involves a matching or correspondence between percepts and the physical environment represented. Drawing on a case from comparative psychology, I argue that, in aiming to reduce spatial aspects of the percept to mind-independent geometrical facts about the object-perceiver relation, representationalists continue unwarrantedly to model perception on belief. Many raptors have much greater visual acuity than humans. For instance, an eagle can detect an insect 0.23cm long under superb viewing conditions from 35 meters up (approximately the height of a 10-story building). And, at least among many types of raptors, movement of the target can allow for detection of an even smaller object. Such acuity is partially explained by the typically large tubular eyes of raptors, creating large retinal projections, and the densely-packed photoreceptors in their retinas. However, impressive acuity alone does not challenge the view that the percept can be explained in terms of mind-independent geometrical facts about the relation between the bird and its physical environment. What is more problematic for the representationalist is the raptor’s fovea (called a convexiclicate fovea), a deep convex depression in the retina of the eye where there is the highest density of photoreceptors. In raptors, the fovea is not shallow as it is in primates, but is a deep depression with steep walls that bulge. It is likely that the raptor’s fovea works as an internal telephoto lens, an optical element that projects a magnified image on the receptors at the center of the fovea, an area responsible for the center of the field of vision. In a human-sized falconiform eye, for instance, the magnification factor at the center of the field of vision is estimated to be approximately 1.45. Moreover, vertebrates are not the only class of organisms in which deep fovea create magnification effects. Similar magnification effects are well-documented in the eyes of jumping spiders (Salticidae), for instance. I contend that variation in size percepts caused by these physiological examples challenge the representationalist view that spatial aspects of the percept can be explained in terms of mind-independent features of the geometrical relation between perceiver and object. I conclude with a discussion of what it would mean to model perceiver-environment relations without a commitment to account for spatial aspects of the percept in terms of mind-independent facts. That is, I conclude with a discussion about what it would be to model perception in its own distinctive terms.

References


In this talk, we present a novel perspective on recent applications of social choice theory to the problem of theory choice. This literature maintains that an analogue of Arrow’s theorem is a serious threat to the rationality of theory choice (cf. Okasha, 2011). We claim that epistemic values are themselves historical entities, co-evolving with scientific theories, as already suggested by Kuhn. Hence, accounts of theory choice must consider the impact that previous choices have on the weight of epistemic values. We provide several possible characterizations of this phenomenon through a simple dynamic model of theory choice, where the social choice apparatus is augmented with weights dependent on past choices. We show how this dynamic model gives a less idealized picture of scientific theory choice, applying it to the controversy between Mendelians and biometricians in early 20th century biology.
In addition to the role they play in both aesthetics and the study of geometric shapes, symmetries play a central role in contemporary physics. It follows from Noether’s (first) theorem that the well-known conservation laws are tied to principles of symmetry, in the sense that they are interderivable. Many authors have connected physical symmetries to explanation, claiming, for instance, that Permutation Invariance explains the nature of quantum statistics. The status of symmetries within physics is not just theoretical, however, as they also play a heuristic role. The best-known example of this is in the Eightfold Way classificatory scheme of the 1960s. On the basis of symmetry considerations, Murray Gell-Mann and Yuval Ne’eman predicted the properties of an undiscovered particle that would complete the baryon decuplet – a prediction that was soon after experimentally verified. That the physical laws abide by symmetry principles is seen today as no mere accident: the latter are treated as guiding principles which we expect future laws to accord with.

A prominent advocate for philosophical consideration of the role of symmetries is Marc Lange who treats them as ‘metalaws’ (that is, second-order laws which are related to the first-order laws in the same way that the latter are related to the events). The purpose of this paper is to suggest a novel use of that treatment beyond the realm of physics. In particular, the present suggestion is that we can utilise this conception to make sense of the connection between certain kinds of biological principle and the associated patterns in the biological phenomena. Put simply, I argue that we should make room for metalaws within the philosophy of biology.

This will strike some as an immediately unappealing move. It is a matter of some controversy whether there are any laws within biology or whether the regularities studied fail to meet the criteria for lawhood (on account, perhaps, of their instability and restricted scope). John Beatty has influentially argued that there are no such genuinely biological laws. Various proponents of laws – such as Sandra Mitchell, Elliot Sober and James Woodward – have pushed back against this reasoning, arguing that if we are willing to be flexible in our definition of lawhood then we will find laws in biology. Although interesting, the question of whether biological ‘laws’ are genuinely laws is tangential to this paper. Supposing for the sake of the argument that there are generalisations that we’d be willing to call laws, the question is whether there is any role to be played by something like the metalaws one finds in physics.

One example of a biological candidate for lawhood is Kleiber’s ‘law’, which relates an organism’s metabolic rate to its mass. The relevant equation contains an exponent term which, on the basis of geometric considerations regarding the relationship between surface area and volume, one might expect to be 2/3. It turns out that such considerations are misleading. Instead, Kleiber’s law is a formalisation of the experimental observation that the value for the term is actually closer to 3/4.

Given the difference between the armchair prediction and the observational result, one might wonder what the explanation for the 3/4 term is. One answer has been suggested by West, Brown and Enquist (WBE). The WBE model treats organisms as having a transportation problem to solve: what is the best system for transporting materials to the cells that require them? For reasons concerning energy efficiency, they argue that organisms will possess a fractal branching transportation system. This then allows them to mathematically derive the quarter-power law. The same considerations are then appealed to in order to explain a broad swath of similar allometric power laws.
Treating the WBE model as capturing a metalaw concerning allometric power laws provides a place for it within an established theoretical framework. WBE’s reasoning shows both the explanatory and predictive features that metalaws possess: they take themselves to have demonstrated why Kleiber’s law contains a 3/4 term and they use their reasoning to predict the values of exponents in other equations. Flatworms, for example, are predicted to have a 2/3 term in their mass–metabolic rate relationship on account of being effectively two-dimensional organisms. Such predictions demonstrate the intended counterfactual robustness of WBE’s reasoning: they take the principles embedded in their model to remain constant over a broad range of different circumstances. These usages are justified if we are dealing with a second-order law concerned with first-order biological regularities.

Kleiber’s law has been subject to a variety of criticisms, including the criticism that there is no generally applicable exponent: the 3/4 term is merely a rough and sometimes misleading approximation. The WBE model too has been criticised for its mathematical assumptions. However, even if these criticisms are entirely accurate, there is still something of philosophical value to be extracted from this example. As the literature on the WBE model attests, it has been taken seriously by working biologists. Our philosophy of biology should be capable of accommodating this without pre-judging its success from the armchair. Whether there are in fact any metalaws in biology should be determined by the relevant science, the role of philosophers here is to build an account where the goals of WBE’s project make sense. Allowing for biological metalaws and treating the WBE model as aiming at capturing one does exactly that.


It is widely agreed by both philosophers and cognitive scientists that an explanatory cognitive theory will be committed to positing representations. In addition to folk psychological states such as beliefs and desires, robustly confirmed theories of language use and acquisition, motor control, perception, planning, and numerous other cognitive domains all seem to work by identifying the representational structures and processes used by the mind/brain. However, philosophical accounts of what these representations are have proved far more controversial. One reason for this is that typical philosophical theories of representation ground this notion in relations between representations and the mind-external entities they represent. However, work in the cognitive sciences suggests that many such posited states seem to bear, at best, highly complex relations to anything in the environment. As such, it is difficult to see what they represent. Mental states of various sorts have been argued to lack external correlates. This creates a tension. On the one hand, representation seems to play a key role in cognitive science. On the other, many paradigmatic purported representations seem, on close inspection, not to represent anything.

In this paper, I shall aim to resolve this tension with a pluralistic account of representation. I shall argue that the difficulty arises by conflating two different questions we can ask about representations: (i) what are the representations involved in a given cognitive process? and (ii) why does the mind utilize these representations rather than others? Computational cognitive science aims to answer the former question, and will include in the kinds it identifies anything which plays a certain internal role, whether it does so in order to track the external environment or not. Anti-individualist theories in philosophy aim to provide an answer to the second, and will thus apply only to a subset of the former kind.

Foundational work (e.g. Burge 2010, Shea 2018) has argued that representation is required in explanations in psychology when an internal state, e.g. a pattern of neural firing, functions to serve as a ‘proxy’ for some external state to which the system does not have direct access. For example, when the visual system generates an internal representation of the nearby predator, the properties of the former can be made to correlate with those of the latter in ways that enable the system to keep track of the predator even though the predator itself is causally remote. Downstream cognitive systems can then use this representation to guide decisions whether to flee or fight. We can only fully understand this system when we realize that the internal state represents the predator.

On the other hand, one central result in modern psychology is that the mind seems to generate its own classificatory schemes, which need not correlate to any environmental distinctions. Pain (Akins 1996), Colour (Chirimuuta 2015), Language (Rey 2020), Smell (Barwich 2020) and numerous other cognitive phenomena have been argued to work in this way. Some of these systems seem to impose distinctions that aren’t found in the objective phenomena, while others seem not to correspond to any objective phenomena at all.

I believe these opposing forces can be reconciled by appealing to a distinction drawn by Marr in his foundational work on psychological explanation. At the computational level, the level at which representations are traditionally posited, Marr claimed that psychologists can ask two questions: what is being computed? and why? What the above cases show is that not all representational systems will receive the same answer to the ‘why’ question. The cases the anti-individualists focus on, such as depth-detection in vision, or ant navigation, are explained by appealing to the usefulness to the system of tracking some objective environmental property. However, tracking the differences in the environment is not the only useful way for a mental system to classify. Internal pressures such
as computational efficiency, ease of recall, informativeness, etc. can all contribute to the representational scheme used by cognitive systems.

From one perspective, the crucial notion of representation is simply a repeatable mental symbol. However, some kinds of mental symbols are explained in externalist ways, and thus form a natural sub-kind relative to the larger set of representations. From a processing perspective, the reason why a distinction is drawn doesn't matter much. But for certain philosophical purposes, identifying this sub-kind is crucial. This thus exemplifies the relativity of classificatory scheme to theoretical domain that has been widely discussed in the philosophy of the special sciences.

References:

Akins 1996- Of Sensory Systems and the 'Aboutness' of Mental states

Barwich 2020- Smellosophy: What the Nose Tells the Mind

Chirimuuta 2015- Outside Color: Perceptual Science and the Puzzle of Color in Philosophy

Burge 2010- Origins of Objectivity

Rey 2020- Representation of Language: Philosophical Issues in a Chomskyan Linguistics

Shea 2018- Representation in Cognitive Science
Farr, Matt

What's so special about initial conditions?

SUMMARY

The early universe is thought to be extremely low probability in a way that calls for explanation. Some have used the ‘initialness defence’ to argue that since initial (as opposed to final) conditions are intrinsically special in that they don’t require further explanation. Such defences commonly assume a primitive directionality of time to distinguish between initial and final conditions, and so rely on the time-directed B-theory of time. I outline and support a deflationary account of the initialness defence in the context of the temporally adirectional C-theory of time, and argue that although there is no intrinsic difference between initial and final conditions, once we have sufficient structure to discern them we should not seek explanations of low-probability initial conditions.

1. The initialness defence

The standard statistical mechanical account of why macroscopic systems obey the second law of thermodynamics relies on the assumption that in the past, entropy was far lower than it is now. Indeed, contemporary cosmology posits an extremely low entropy and thus low probability initial state of the universe, which following Albert (2000) is termed the Past Hypothesis. There has been much disagreement as to whether the extremely low-probability initial state posited by the Past Hypothesis stands in need of explanation (e.g. Price (2004) argues it does; Callender (2004b) argues it does not), and what would constitute an explanation of such a thing (e.g. Callender (2004a) questions whether there could in principle be such an explanation). One thought implicit in a number of considerations of this issue is that there is something intrinsically special about initial conditions, as opposed to final conditions, in that they are the earliest or first state of the universe, and that they thus are not produced by nor depend upon any earlier set of conditions in any causally relevant sense. Call this the ‘initialness defence’.

2. Maudlin’s version of the initialness defence

Maudlin (2007) makes a novel case for the initialness defence, arguing that an underlying directionality of time itself does explanatory work regarding the past hypothesis. Though the state posited by the past hypothesis is macroscopically atypical in that it is very low entropy, later states of such a universe are also dynamically atypical in that the molecular positions and momenta are such that, if evolved backwards in time, they’d lead to ever lower-entropy macrostates. Maudlin argues that his own time-directed, B-theoretic, metaphysics explains such dynamical atypicality away by holding systems to really evolve forwards and not backwards in time: dynamical atypicality is ‘completely accounted for by how it was generated or produced [. . . via] evolution from [the] initial state’ (p. 133), something not available within the context of the temporally adirectional C-theory, since ‘[i]t has its sort of explanation requires that there be a fact about which states produce which[, which] is provided by a direction of time’ (p. 134; emphases added).
3. The C-theory

Central to Maudlin’s argument is the premise that whereas a B-theory of time can account for dynamical atypicality, a temporally adirectional C-theory cannot. I show this to be false by giving an account of the commitments and entailments of C-theories and thus explicating what it means for time to be adirectional. The key difference between Band C-theories is that on the C-theory, there can be no two worlds that differ solely in terms of the direction of time. In other words, time-reversal on the C-theory is a passive transformation, and forwards and backwards time can be understood as giving equivalent descriptions of a single world. A B-theory, on the contrary, takes forwards and backwards time to describe different possible worlds.

4. What’s wrong with the B-theoretic initialness defence?

What the C-theory lacks that the B-theory allows is for there to be a fact as to whether the world is (1) an entropy-increasing world (i.e. the low-entropy end is earlier than now), or (2) an entropy-decreasing world (i.e. the low-entropy end is later than now). On Maudlin’s account, only (1) and not (2) gets around the problem of dynamical atypicality. But I argue that to hold that classical statistical mechanics only successfully accounts for one (1) and not (2) is to make a distinction between the two to which the physics itself is insensitive, and to place an arbitrary restriction on the explanatory power of the physics. There’s no quantifiable sense in which the second B-world is unphysical, or physically impossible. Rather, I suggest that the C-theory gets things better: the relevant statistical mechanical explanation is of the higher entropy state in terms of the lower entropy state, and this is so regardless of which is taken to be the ‘initial’ state.

5. The C-theoretic initialness defence

Although the C-theory takes the two directions of time to offer equivalent descriptions, it does not follow that forwards time is not preferable to backwards time. There are important empirical reasons for preferring this direction; (a) evolution from a low-entropy state leads with high likelihood to a higher-entropy state, but not vice versa; (b) relatedly, causal explanations invariably are from lower entropy states to higher entropy states. These two facts, I argue, are sufficient to show that given the Past Hypothesis, we should prefer to take the low-entropy end to be to temporally and causally precede the present state, and moreover, I show that the hypothesis of a low-entropy temporal end to the universe plays a constitutive role in providing explaining other states in general, and as such it is misguided to look for a further explanation of the low-entropy end itself. I demonstrate this via reflection on the role of exogenous variables in causal models.

References


A hypothesis' ability to unify and systemize different and diverse pieces of evidence is generally seen as an epistemic virtue in philosophy of science. Unification is also often associated with other core concepts of philosophy of science such as abduction, confirmation, causation, prediction, and explanation. In this talk, we bracket abduction, confirmation, and prediction and rather focus on two different views of unification and their connection to explanation from a causal perspective. Taking such a causal perspective will allow us to see that causal structure matters for how probabilistic measures of unification perform and relate to explanatory relevance. Many philosophers of science hold the view that the better a hypothesis h unifies a body of evidence e1,...,en, the better it can explain it. Some even go as far as proposing that explanation can be reduced to unification (see, e.g., Kitcher 1981, Kitcher 1989). In this talk, however, we rather focus on the more general question of whether unification is a good indicator for explanatory relevance.

Which account of unification gets things right and how exactly unificatory power can be measured is still controversial. In this talk, we are especially interested in the following two prominent approaches to unification:

MIU -- Mutual information unification: A hypothesis h has the more unificatory power with respect to pieces of evidence e1,...,en the more it renders these pieces of evidence (more) informative about each other.

COU -- Common origin unification: A hypothesis h unifies a body of evidence e1,...,en in so far as it posits a common origin for these pieces of evidence.

MIU has been defended by Myrvold (2003, 2017), and COU by Lange (2004). Both authors have criticized each other's account. Myrvold stressed that according to a Bayesian decomposition, only "MIU contributes to incremental evidential support, and there is no scope, within Bayesian updating, for COU to add to the evidential support of the theory" (p.93). And Lange claimed that "genuinely to unify [pieces of evidence], a theory must reveal them to have some deep common explanatory basis" (p.208) and that Myrvold's account is inadequate because it "sets the bar too low to distinguish genuine from bogus unification" (ibid.). So, whereas Myrvold claims that COU is inadequate in terms of confirmation, Lange claims that MIU is inadequate in terms of explanation. Against Myrvold's claim about COU, Niiniluoto (2016) argued that COU also plays an important role for confirmation if more broadly conceived as including abductive confirmation. However, this makes COU not automatically a better candidate for accounting for the different roles typically associated with unification.

The main goal of this talk is to shed new light on both MIU and COU from a causal perspective. To this end, we draw on Reichenbach's (1956) insight that common causes screen off their effects (or render them less informative about each other in the presence of additional causal connections) that became a crucial assumption in causal modeling. Based on this simple idea, we propose a probabilistic measure for COU that in some sense complements Myrvold's probabilistic measure for MIU: According to this first probabilistic take on COU, a hypothesis has the more unificatory power the more it renders pieces of evidence uninformative about each other. As a next step, we will use causal Bayesian networks (Pearl 2000, Spirtes 2000) to represent different patterns of how a hypothesis h can be causally connected to a body of evidence. As we will see, already focusing on the simplest causal patterns suffices to make some relevant observations. We apply Myrvold's (2003, 2017) measure for MIU and our first take on measuring COU to each of these structures. The
The upshot of this will be that causal structure heavily constrains the performance of these probabilistic measures. Next, we use the basic causal structures and our results about how the measures for MIU and COU behave to shed new light on the connection of unification and explanation in causal settings. It will turn out that both probabilistic measures of unification do a bad job as indicators for explanatory power. While the measure for MIU underperforms when applied to the elementary causal structures we discuss, the probabilistic measure for COU is too permissive. Based on this observation we further develop our probabilistic measure for COU by adding a causal constraint, which will improve its ability to indicate explanatory power significantly. We modify the measure for MIU in a similar way and compare it with the causal measure for COU.

The structure of the talk will be as follows: First, we will introduce Myrvold’s (2003, 2017) measure for MIU. We will then propose a complementary probabilistic measure for COU and discuss its relation to Myrvold’s measure. Afterwards, we will investigate how these measures for MIU and COU perform given different elementary causal structures. In the final part we will explore what we can learn from our earlier results about the relation between unification and explanation and how we can account for these insights by adding a causal constraint to our probabilistic measure for COU and how this causal measure for COU performs compared to a similarly modified version of the measure for MIU.
Findl, Johannes and Javier Súarez

Prediction, understanding and COVID-19: Early epidemiological models of COVID-19 shed light on the relationship between prediction and understanding

In an unprecedented manner, the COVID-19 pandemic has caused rapidly growing rates of viral infections all over the world, threatening the lives of many people and putting hospitals in acute danger of becoming overwhelmed. This emergency situation has accelerated scientific research, and many resources have been dedicated to understanding the pandemic, one key aspect being how it would spread. By developing epidemiological COVID-19 models that predict the rates of infections, mortality, and hospitalizations for scenarios in which countermeasures such as social distancing policies are introduced or removed at a time, scientists have sought to provide an urgently needed basis for political decision-making.

COVID-19 models can be distinguished by three different types: statistical models that derive their estimations from a regression analysis that fits a curve to empirical data such as the number of infections or deaths, mechanistic models that simulate disease transmission between (groups of) persons on the basis of empirical data such as the virus’s spread and the onset of disease symptoms, and hybrid models that combine both approaches.

Since little was known about the spread of the disease at the beginning of COVID-19, predictions were mostly generated with statistical models that did not include causal-mechanistic knowledge. While initially criticized for the large discrepancy between predicted and actual deaths, scientists have further developed those early statistical models and have eventually released increasingly accurate versions of them. Through continuous updates and modifications, some of these models have been improved up to the point where they can be said to have acquired genuinely predictive capacity, becoming indispensable tools for understanding how COVID-19 would spread in different scenarios. Arguably, the improvement also suggests that by and by, epidemiologists have achieved a better understanding of the main variables determining the spread of the disease.

This raises important philosophical questions about how purely statistical models can yield understanding, and what the relationship between prediction and understanding in such models is. In this paper, we investigate them by analysing how the statistical model from the Institute of Health Metrics and Evaluation (IHME model) was developed and modified between March and April 2020. The IHME model projects the mortality and hospitalization rates resulting from COVID-19 and has been used by US policy makers from early on in the pandemic.

We believe that by way of studying the details of the IHME model building process over time, we will be in a good position to analyse how the concepts of prediction and understanding interact and affect each other in the epidemiological practice related to COVID-19, which may also give us valuable insights into the more general nature of this relationship.

The first result of our analysis is that a key epistemic virtue of the IHME model is its ability to generate regularity patterns through predictions, thereby enabling scientists to understand how COVID-19 would spread in different scenarios. More precisely, we observe that the model creates regularity patterns by combing a simple technical framework that depicts a mathematical function (i.e., a Gaussian error function) with a series of assumptions about the set of variables that would affect the results of the technical framework (i.e., the introduction or the removal of countermeasures) and, in a sense, affect the phenomenon (i.e., the death and hospitalization rates). Understanding, we argue, results from considering and comparing the model’s visualizations. As we show, this is in line with de Regt’s intelligibility requirement, according to which scientists can
acquire understanding only if they recognize qualitative consequences of their theory. We claim that the IHME’s model visualizations should be thought of as regularity patterns that apply to possible scenarios in which countermeasures are either introduced or removed at a time.

A second result of our analysis is that predictions have another important epistemic function over and above providing the regularity patterns that makes understanding of the phenomenon feasible. Focusing on how the IHME scientists developed their model between March and April 2020, we find that when comparing the model’s predictions with the actual evidence, the scientists were prompted to reconsider their starting assumptions. By doing so, they eventually understood which of them were correct, and which ones were mistaken. Concretely, our analysis shows that certain important local conditions that severely affected the death rates from COVID-19 in different countries had not been included in early versions of the IHME model (such as population density, urban area vs. countryside, or level of compliance with the mobility restrictions) and thus were incorporated through a number of updates. This observation confirms our hypothesis that predictions facilitate understanding of a phenomenon by pinpointing the error and success in the model building process. That is, scientists use the model’s predictions in order to test the assumptions and the technical framework that had generated them. This crucial step allows them to see where the model gets the phenomenon right, and where it needs to be improved.

Finally, we show that in the case of the IHME model’s development, the concepts of understanding and prediction are intimately linked in a dynamic and dialectical way which results from the generation of a regularity pattern that is then compared with actual data. As Figure 3 illustrates, at the first stage, the model generates predictions that provide a regularity pattern that allows scientists to understand the phenomenon (in our case, COVID-19-derived mortality rate). At the second stage, though, the predictions work backwards: they are contrasted with the evidence and, if scientists observe large divergence, they are forced to reshape their model by modifying the assumptions upon which it relies. This second step may require altering some of the prediction-generating assumptions, as it happened in the IHME model’s development. Importantly, after this second stage has taken place, the predictions are themselves altered, i.e., they are freshly produced. This is because the change in the assumptions reshapes the model, hence the prediction-generating process starts again, and, in that vein, the whole assumption-prediction-understanding process is taken to the next stage.

References:


Landauer’s Principle (LP) proposes a connection between thermodynamics and physical computation: each “erasure” that a computer implements incurs a minimum of energy dissipation into the environment, with implications for engineers’ analysis of the energy use of the computational economy. Yet the truth of LP and the soundness of purported proofs of it have been attacked on various conceptual grounds. The debate, it seems, has reached an impasse.

Drawing on recent technical work on LP from quantum statistical mechanics, this presentation aims to make progress on this problem by recasting LP and reinterpreting its relationship to computation. This interpretation surprisingly vindicates both the skepticism about LP as providing an essential connection between computation and thermodynamics, on the one hand, and the relevance of LP to the practical concerns in engineering to which it has been applied. Thus, it is a case where a fundamental idea works in practice, but not in principle!
Truthlikeness is a property of a theory or a proposition that represents its ‘closeness’, ‘similarity’ or ‘likeness’ to the truth. Since Popper’s (1963) failure to give a satisfactory definition of the concept, the notion has been a topic of intense discussion by philosophers of science and logicians, particularly as it seems an indispensable element for a plausible formulation of scientific realism and scientific progress.

The literature on truthlikeness has focussed mainly on qualitative and quantitative languages where truth was assumed to be deterministic. So far, there are no proposals on how to deal with probabilistic truths (with the exception of Niiniluoto (1987)). However, scientific theories seem to present genuinely probabilistic laws. In this regard, we will try to present an appropriate definition of truthlikeness for probabilistic laws (PL henceforth). This will be done in two main parts.

(1) On the one hand, Niiniluoto (1987) suggests to use the Kullback–Leibler divergence (KL henceforth) to define the distance between a probability law X and the true probability law T. However, he does not provide an extensive argumentation on why KL would represent an appropriate similarity metric for PL. In the literature of probability there is a host of available distances. We will select a small sample of the most representatives, according to the categorization developed by Cha (2007). We will confront them in different scenarios, finally concluding that KL represents the better notion of similarity to the truth for probabilistic systems.

(2) On the other hand, Niiniluoto (1987) defines truthlikeness for deterministic laws (DL) as a function of accuracy. García-Lapeña (2020) expands Niiniluoto’s proposal, showing that accuracy represents a necessary but not sufficient condition, and defines truthlikeness for DL as a function of two factors, accuracy and nomicity (value similarity and shape similarity). Analysing the exponential law of radioactive decay and some probabilistic cases regarding a normal distribution, we will argue that (as in the case of DL) KL is not sufficient to properly define truthlikeness for PL. Another factor, probabilistic nomicity, needs to be added to the definition.

The final proposal will define truthlikeness for PL as a function of two factors, p-accuracy and p-nomicity, in intimate connexion with García-Lapeña’s (2020) proposal for DL.

References


Geddes, Alexander

The Limits of Immunology: Pradeu on Organisms and Organismic Parthood

In a series of works, Thomas Pradeu (2010, 2012, 2013, 2016) articulates and defends an immunological account of the individuality of organisms. The picture he offers is empirically sophisticated, but not always fully transparent in its metaphysical commitments and ambitions. In this paper, I bring out and critique some of the metaphysical details of Pradeu's account, and argue that there are limits to the work it is able to do in accounting for the existence, persistence and composition of organisms.

In §1, I distinguish between four closely connected but conceptually distinct questions. The first concerns *organismality*: what it is to be an organism. The second concerns *organismic realisation*: the conditions under which a plurality realises (or grounds the existence of) an organism. The third concerns *organismic composition*: the conditions under which a plurality constitutes an organism. And the fourth concerns *organismic parthood*: the conditions under which something is part of an organism. These distinct questions are often run together. But despite this, or perhaps because of it, different accounts of biological individuality can fruitfully be read as varying in which question they treat as prior. And I argue that Pradeu's account is in the first instance an account of organismic parthood.

In §2, I set out what I take to be Pradeu's account of organismic parthood. (Or rather, I present what I take to be the *letter* of his account. I go on to offer refinements aiming to maintain its spirit, and which plausibly reflect aspects of his thinking that are not quite spelt out in the text.) Letting *e* be any entity and *o* be any multicellular organism, he seems to hold that:

1. *e* is part of *o* iff *e* is immunologically tolerated by *o*

Immunological tolerance is then understood in terms of immune interactions:

2. *e* is immunologically tolerated by *o* iff there are immune interactions (of an appropriate sort) between *e* and *o*

Much of his (2012) concerns itself with spelling out what the relevant sort of immune interaction is, and providing a theory (the 'continuity theory', contrasted with the 'self/non-self' theory) about when and why immune interactions constituting tolerance and rejection occur. But abstracting away from these details, we can say:

3. There are immune interactions (of an appropriate sort) between *e* and *o* iff there are interactions (of an appropriate sort) between ligands expressed by *e* and *o*'s immune receptors

Putting these together, we get:

4. *e* is part of *o* iff there are interactions (of an appropriate sort) between ligands expressed by *e* and *o*'s immune receptors

Pradeu takes this to provide a satisfying, theoretically grounded and empirically adequate picture of what an organism comprises, albeit one with some potentially counterintuitive results. That is, he takes it to deliver the verdict that all of an organism's normal organs and tissues are parts of that organism, but also that fetuses, bacteria and parasites are, when not rejected, also parts of the organisms hosting them.
In §3, I raise issues for this account of organismic parthood. The issues arise from the fact that (1)-(4) are formulated primarily with cells in mind as candidates for parthood. So to arrive at a fully adequate and general account of organismic parthood covering all entities, we both need to ‘build down’, to cover sub-cellular components, and to ‘build up’, to explain how Pradeu’s account can handle entities larger than cells: organs, tissues, parasites, fetuses. And while the former can be accommodated with minor tweaks, the latter, I argue, present slightly more serious difficulties.

Now, Pradeu’s (2012) discussion does focus on these larger entities at numerous points. And when it does, it proceeds as if he accepts the following principle:

(P) if some part of *e* is immunologically tolerated by *o*, then *e* is immunologically tolerated by *o*

However, I raise two sets of issues for (P) in combination with (1)-(4). The first set concerns candidates for organismic parthood that are themselves organisms, such as fetuses. I argue that we get conflicting verdicts about whether fetuses are parts (some fetal cells are parts according to (4), and so, by (P), fetuses are parts; but some fetal cells are *not* parts according to (4), and so, by the transitivity of parthood, fetuses are not parts), and as well as implausible verdicts of mutual parthood ((4) and (P) together imply both that the fetus is part of the maternal organism and that the maternal organism is part of the fetus.)

The second set of issues concerns candidates for organismic parthood that are *not* organisms, such as tissues and organs. Of particular interest are functional tissues that include dead cells—such as hairs, nails, (the outermost layer of ) skin—and organs at immunoprivileged sites. The problem, I suggest, is that something like (P) is needed to deliver the verdict that these are parts, as they contain cells that either cannot or do not routinely interact with the organism’s immune system; and yet (P) can only deliver this verdict if there is some *further* account of parthood for such organs and tissues, besides the present account of organismic parthood.

Finally, in §4, I discuss how Pradeu ought to respond to these issues. I point out that they pull in different directions: the first set seems to require abandoning (P), while the second set seems to require keeping it. In the end, I argue, the only reasonable way to resolve the tension is to take immunological tolerance to provide, not a full account of organismic parthood, but simply a sufficient condition for it. I conclude by considering whether this is likely to end up undermining itself, in requiring supplementation by some other account of organismic parthood which might then end up generating conflicts with the claim, and by tracing some of the implications it may have for the other questions identified in §1.

Short Bibliography:


Greenwood, John

**Durkheim’s Dogma**

Emile Durkheim is known to philosophers of social science as a paradigmatic ontological holist, who maintained that social groups or collectives exist in some sense ‘over and above’ the individuals that compose them. Thus Durkheim claimed of social groups that ‘The whole does not equal the sum of its parts; it is something different, whose properties differ from those displayed by the parts from which it is formed’ (1895a, p. 128) and maintained that social facts are distinct from psychological facts:

Social facts differ not only in quality from psychical [psychological] facts; they have a different substratum, they do not evolve in the same environment or depend on the same conditions....The mentality of groups is not that of individuals; it was its own laws.

(1901, p. 40).

Durkheim’s apparent commitment to a radical ontological discontinuity between the stratified subject matters of sociology and psychology also explains why many have also treated Durkheim as a paradigmatic explanatory holist, who held that social phenomena could only be explained in terms of social facts. Thus Durkheim famously—or infamously—claimed that:

There is between psychology and sociology the same break in continuity as there is between biology and the physical and chemical sciences. Consequently, every time a social phenomenon is directly explained by a psychological phenomenon we may be sure the explanation is false.

(1895a, p. 129)

I call this latter claim ‘Durkheim’s Dogma.’ In this paper I suggest that this claim was only apparently dogmatic and that Durkheim was not opposed to psychological explanations of social phenomena, but only to a particular form of psychological explanation.

With respect to sociological and psychological explanation, Durkheim maintained that:

I have never said that sociology contains nothing that is psychological and I fully accept...that it is a psychology, but distinct from individual psychology.

(1985b, p. 249).

All Durkheim seems to have meant in claiming that “the mentality of groups is not that of individuals; it has its own laws” is that those forms of mentality that are a function of “the way in which individuals associating together are formed in groups”(1897, p. 171) are likely different from those that are formed independently of such association (via other reasons or causes). Alternatively put, those forms of “acting, feeling and thinking” (1895a, p. 52) that are socially engaged, that are oriented towards the acting, feeling and thinking of members of social groups, are different from those forms of acting, feeling and thinking that are individually engaged, for reasons or causes independent of an individual’s orientation to social groups.

This would appear to be the distinction Durkheim was making when he maintained that sociology was not distinct from psychology, but was a different form of psychology, namely social psychology. While he recognized that many sociological explanations are social psychological explanations, he
claimed that such explanations do not reduce to individual psychological explanations, or at least could be presumed to do so: “one has no right to treat collective psychology as an extension, an enlargement or a new illustration of individual psychology” (1895b, p. 250).

In making these claims, Durkheim was also making the critical point that whether or not social explanations do reduce to individual psychological explanations is an empirical question to be determined by sociological and psychological science, and not by philosophical fiat. Thus, to take two examples of Durkheimian social-statistical facts, the best explanation of differential suicide rates between Catholics and Protestants may be social psychological, in terms of the different beliefs about the moral acceptability of suicide socially engaged by Protestants and Catholics, whereas the best explanation of differential suicide rates between the employed and the unemployed may be individual psychological, in terms of differences in rates of major depression between the employed and the unemployed.

Finally, Durkheim’s oft-quoted pronouncement that when a psychological explanation of a social phenomenon is offered we may be sure it is false can be interpreted not as a dogmatic statement but as a reasonable prediction about the explanatory potential of the individualist associationist psychology of his own day. Durkheim actually allowed that social explanations might ultimately reduce to individual psychological explanations, but insisted that such questions could not be answered until an adequate social psychology was developed, which was not the case in his own day:

Social psychology, whose task it should be to determine them is hardly more than a term which covers all kinds of general questions, various and imprecise, without any defined object.

(1901, p. 42)

Whether the situation has changed in our own century remains an interesting and open question.

References


Beginning with the seminal work by William Wimsatt, the concept of robustness has received considerable attention in philosophy of science. In his original conception, Wimsatt characterized robustness as the “use of multiple means of determination to ‘triangulate’ on the existence and character of a common phenomenon, object, or result...” (Wimsatt, 2007, 43). According to Wimsatt, scientists achieve robustness by introducing variations into scientific procedures, which may include data analysis, theoretical derivation, or experimental technique. Following this characterization, many authors put the concept of multiple determination to philosophical use in different contexts, including accounts of model-confirmation (Weisberg, 2006; Lloyd, 2015), establishing the reality of entities (Hacking, 1983), or evidencing hypotheses (Staley, 2004).

However, the value of robustness as an epistemic virtue has also been criticized. Using a formal Bayesian framework, François Claveau and Olivier Grenier recently argued that robust evidence, conceptualized as a variety of evidence thesis, may in fact lead to a decreased confirmation of a scientific hypothesis (Claveau and Grenier, 2019). From a different perspective, Robert Hudson argued, on the basis of case studies taken from various sciences such as biology, molecular physics, and astronomy, that robustness considerations do not actually play any epistemic role in evidential reasonings in science (Hudson, 2013). Among the critical approaches to robustness, a more specific argument is developed by Jacob Stegenga, who focuses on the problem of discordant evidence in the context of robustness analysis in experimental science. In his criticism, Stegenga observes that when multimodal evidence, i.e., evidence from multiple techniques, for a given hypothesis is discordant, “it is unclear what support is provided to a hypothesis...” (Stegenga, 2009, 654).

According to Stegenga, the reason for this is that we lack “principled methods of quantifying concordance and assessing and amalgamating multimodal evidence...” (Stegenga, 2009, 656). If there are multiple experiments that disagree among each other, without a clear scheme of how to amalgamate them, the result would necessarily be inconclusive. Stegenga observes that the issue of relevance arises in this context: One way to overcome the problem of discordance could be to argue that some of the evidential modes that constitute the discrepancy are in fact not relevant (or less relevant) to the hypothesis. In this way, by disregarding the problematic results, one could still achieve genuine robustness by regaining concordance. However, Stegenga doubts this possibility as “scientists lack universal criteria for making decisions regarding relevance...” (Stegenga, 2009, 658). There may be local criteria in particular disciplines, such as the particle physicists’ treatment of “golden events” as high-quality evidence, yet this does not solve the problem, as high-quality does not necessarily mean most relevant (Stegenga, 2009, 658). Stegenga concludes that “the more a particular body of evidence is discordant, the less useful the methodological strategy of robustness is” (Stegenga, 2009, 656).

The aim of my paper is to examine Stegenga’s objection to robustness analysis in the context of a case study in experimental cosmology. The case I focus on is the measurement of the Hubble constant. A significant problem in contemporary cosmology is that current measurements of the Hubble constant, which measures the rate of expansion of the universe, give discordant results. Furthermore, in the publications of the major experiments on this issue, we find that the scientists appeal to robustness analysis as a justificatory tool. Thus, the issues raised by Stegenga are immediately pertinent to contemporary cosmology. Yet, as I aim to argue, contrary to what Stegenga maintains, contemporary practice in experimental cosmology shows that discordant multimodal evidence does not “diminish the epistemic value of robustness” (Stegenga, 2009, 658). The key fact to note is that multimodal measurements of the Hubble constant involve experiments
that depend on the standard model of cosmology in distinct and precise ways. This dependency network, as one may call it, opens up the possibility that the discordance may in fact be an indicator of the presence of new physics, provided that the discordance is robust. As Allan Franklin observed, robustness can be provided by “variations on a single experiment” as well as by the more common multimodality of a “sequence of experiments” (Franklin, 2002, 38). The robustness of the Hubble discordance is established by showing that each of the individual measurements in the multimodal evidential basis are themselves robust. Thus, contemporary practice in cosmology shows that the strategy of robustness is not only useful in the context of discordant evidence, it is in fact a methodological requirement for establishing an evidential basis for new physics. In other words, in order to show that the contemporary cosmological model needs to be modified or amended—and make the case for hitherto unknown fundamental physics (for example, various exotic dark energy scenarios, dark matter–dark energy coupling, or a new relativistic particle) beyond the standard cosmological model—scientists need to establish the discordancy between the different measurements of the Hubble constant in a robust way. This is done by introducing several types of variations (in Wimsatt’s sense) into each measurement, to rule out possible systematic errors that may be responsible for an erroneous result.

Currently, there are multiple experimental programs that aim at measuring the Hubble constant. In the paper, I consider four of these to illustrate and evidence my claim. The first evidential mode is the early universe, or the indirect, measurement by the ESA Planck satellite that uses cosmic microwave background observations to obtain a value of the Hubble constant. The key aspect of the Planck measurement is the fact that it is obtained by assuming the standard ΛCDM model to be valid: the value of the constant is inferred from the measured model parameters. Thus, the early universe determination of the constant is ΛCDM model-dependent. The current Planck value for the Hubble constant is \( H_0 = (67.4 \pm 0.5) \text{ km s} \text{ Mpc} \) (Planck Collaboration et al., 2020). The remaining modes are all late universe or “local” probes (which are independent of ΛCDM), that rely on various types of distance measurements to determine the Hubble constant. It is this differential dependence of the distinct evidential modes on the base model that enables the discordance to function as a probe for new physics. One such method uses Type Ia supernova distances, calibrated by Cepheid stars. The current best estimate of the supernova measurement is \( H_0 = (74.03 \pm 1.42) \text{ km s} \text{ Mpc} \) (Riess et al., 2019). This implies a 4.4σ tension with the Planck result. A third project, which uses red giant branch stars as distance indicators, obtained a value of \( H_0 = (69.8 \pm 1.9) \text{ km s} \text{ Mpc} \), which “sits midway in the range defined by the current Hubble tension” (Freedman et al., 2019, 1). Finally, a fourth mode that employs “time-delay distance” measurements of gravitationally lensed quasars obtained the value of \( H_0 = H_0 = (73.3 \pm 1.8) \text{ km s} \text{ Mpc} \) (Wong et al., 2019). This measurement is 3.1σ discordant with the Planck result, but a “combination of time-delay cosmography and the distance ladder results is in 5.3σ tension with Planck” (Wong et al., 2019, 1420). As the authors of a recently published report on the Hubble discrepancy write, the discord is “robust to the exclusion of any one method, team or source” (Verde et al., 2019, 892).

For Stegenga, the fact that no universal criteria for determining relevance is available means that discordance implies inconclusiveness. However, we see that no such relevance determination is necessary, as all evidence can be taken into account, and a robustly established discordance would conclusively imply the presence of new physics. Similarly, the lack of a systematic means of evidence amalgamation does not mean no “local” criteria can be found. Experimental cosmologists use several strategies, including formal methods, and normative “best practices” approaches, to successfully amalgamate the evidence.

To conclude, the fact that we can obtain both model-dependent and model-independent measurements of the same quantity, supported by properly (and locally) constructed practices of
evidence amalgamation, enables us to use discordance to test the model. A robust and discordant measurement is not necessarily a contradiction in terms, but a method for knowledge growth.

References


Heilmann, Conrad with Marta Szymanowska and Melissa Vergara Fernández

Testing concerns about financial economics

Introduction

Financial economics -- or finance for short -- studies decision-making under uncertainty in the financial sector, i.e. financial markets, institutions, and agents. The models and theories in asset pricing typically centre around one question: how is the price of financial assets determined -- why do assets differ in returns? Asset pricing has been hampered by severe doubts about its testability. In particular, finance scholars have raised doubts about the testability of two most important building blocks of asset pricing: the Efficient Market Hypothesis (EMH) and the Capital Asset Pricing Model (CAPM).

In this article, we analyse the problems of testability of both EMH and CAPM. Doing so is important as it not only sheds light on an under-researched area of economics in the philosophy of science. It also is a crucial step towards critically appraising the scientific status of financial economics from a philosophical perspective. Specifically, it will help to analyse the enormous proliferation of models in modern asset-pricing theory, as there is at present no consensus about the correct asset pricing model, and indeed also a debate about whether progress is being made in asset pricing research (Fama 1991, Cochrane 2011, Campbell 2014, Harvey et al. 2016).

We show that the problems of testability that finance scholars have identified are not as principled and severe as claimed by them. We do so in two steps: One, we characterise the sources and precise nature of the challenges in testing EMH and CAPM. Two, we argue that the empirical challenges of testing the CAPM amount to a straightforward case of underdetermination and those of testing EMH are a much more difficult case of it.

CAPM and Roll’s critique

The capital asset pricing model (CAPM) of Treynor (1962), Lintner (1965), Sharpe (1964), and Mossin (1966) is one of the major frameworks in financial economics for analysing investor behaviour under the conditions of risk. The essence of the CAPM is the statement that the market portfolio is mean-variance efficient. Famously, Roll (1977) has argued that the CAPM cannot be tested directly, as the market portfolio is unobservable. What can be observed are proxies, such as the S&P 500 index. This leads to the following challenge: whenever one tests the CAPM, one also tests whether the proxy used for the market portfolio was correct. Any confirmation or falsification of the CAPM might well be due to the proxy not reflecting the market portfolio accurately enough, rather than the market portfolio (not) being mean-variance efficient.

In the literature, there have been several responses to Roll. Some authors have offered multiple tests with different proxies for the market portfolio, e.g. Jagannathan and Wang (2007). There are also multiple approaches that focus on testing mean-variance efficiency, e.g. Levy and Roll (2010), and multiple tests of linear relation between beta and expected return with varying degree of attention towards market efficiency, e.g. Fama and French (1992). In short, there is a large literature that has developed different variants of the CAPM in order to accommodate the problems of testability. We argue that there is thus nothing special, or severe, about these kinds of testability challenges. Broadly speaking, asset pricing is doing the right thing in trying to get better at finding an
appropriate proxy for the market portfolio, varying the proxy, and also varying how exactly mean-variance efficiency is explained.

Efficient market hypothesis (EMH)

The EMH is perhaps one of the most important building blocks of asset pricing (Fama 1970). The EMH can be stated as follows: market prices equal the conditional expectation of information. Importantly, the EMH is a ‘joint’ hypothesis: (H1) about the equality between prices and the conditional expectation of information. And (H2) about how the conditional expectation of payoffs and returns relates to all the available information. Now, testing the H1 directly is impossible. The only possibility is to test H2. That, however, is impossible without observing market prices. For one, payoffs and returns are prices. For another, the information set is also best observed, in large parts, through prices. In other words, one needs to assume H1 to test H2. This problem is thus considerably more severe than the one discussed for CAPM.

The asset pricing literature offers tests focused on H1, where impact of H2 is minimized, such as by introducing event studies, random walks, serial autocovariance, or short-term dependence. There are also tests focused on H2, such as the Fama French three-factor model, and tests with intertwined H1 and H2, such as those focusing on return predictability, e.g. the rational/irrational time-variation and the variant of conditional model. There is thus fragmented evidence: the idea is to accumulate partial tests to provide evidence for or against EMH. Testing the EMH is thus a much more severe case of underdetermination.

Conclusions

Asset pricing, in particular EMH and CAPM, faces underdetermination. We have shown that the CAPM is a straightforward case. General worries about the lack of progress of the science of asset pricing in light of Roll’s (1977) critique of the CAPM are thus overblown. The EMH is however a much more severe case of underdetermination, as it involves a ‘joint’ hypothesis. These insights into testability challenges in asset pricing offer a much more nuanced picture than has been proffered by financial economists, in particular by the highly influential Roll (1977) and Fama (1970, 1991). They also offer a suitable starting point for analysing which theoretical virtues might be pursued by finance scholars in order to maximise the success of their testing, improve model selection, and further scientific progress.

Key references


Reconsidering multi-level mechanistic explanation

According to a widespread ontological view associated with new mechanism, levels of nature typically invoked in explanations in life sciences are levels of mechanisms. On Craver’s (2007) popular account, the relation between mechanistic levels is the relation between the mechanism as a whole and its components; in turn, this relation is considered to be a non-causal dependency relation, and is to be viewed in terms of mutual manipulability. Craver analyses mutual manipulability in terms of Woodward’s concept of ideal intervention. So, Craver’s mutual manipulability account serves both to give an account of the non-causal relations between the components and the whole mechanism, and to ground a hierarchy of mechanistic levels.

However, while this account seems to capture the practice of interlevel experiments in life sciences, it has recently come under criticism, as it has been argued that the concept of ideal intervention cannot be applied to constitutive relations (Harinen 2018, Kӓstner & Andersen 2018). While philosophers have responded to this problem by trying to reconcile interventionism with mechanistic constitution, or by finding an alternative way to characterise mechanistic constitution, there exists a third, and more radical, option: namely, to reject the view that constitution is needed to understand the notion of a mechanism. This option gives rise to the question, how shall we understand mechanistic levels and the relations between them.

In this paper I will first briefly review the main reasons for introducing constitution in the analysis of mechanism and the problems associated with accounting for constitution in terms of interventionism. I will then present a new argument that undermines the motivations behind Craver’s account of constitutive mechanisms and mechanistic levels. This argument is based on the claim that typical and paradigmatic biological mechanisms are causal pathways. Lastly, I will sketch an alternative account of multi-level mechanistic explanation, by discussing various biological examples.

Craver describes the components of the mechanism as x₁-φing, x₂-φing etc, where the xs refer to the entities that comprise the mechanism and the term ‘x-φing’ refers to an entity engaging in an activity. He describes the phenomenon as S-ψing, where the S is a structure that ψs, and where S-ψing is the phenomenon that is taken to be constituted by the mechanism. Craver requires that all xs (i.e. all components of the mechanisms) be parts of S. For example, when a neuron generates an action potential, S is the neuron, S-ψing the neuron generating an action potential, and the x that φ are the various components of the mechanism for the generation of the action potential.

The main criticism against Craver’s account will be that there is no biological motivation for supposing that in cases of biological mechanisms, the organised entities and activities are always parts of a larger entity, whose behaviour the mechanism underlies. Firstly, we cannot always find natural boundaries around mechanisms. Secondly, we cannot expect this: typical and paradigmatic biological mechanisms are causal pathways, which are not confined within biological objects/structures. In other words, in many cases there is no biological S, where every X (component of mechanism/pathway) is part of S. Thirdly, take a mechanism that occurs within a biological object, for example protein synthesis which occurs within cells. What is the S that ψ-s in this case? It cannot be the cell—the reason is that the mechanism can exist outside the cell. but then the parts of the mechanism (the Xs) can exist without the S. But whenever we have the mechanism, we necessarily have S-ψing. So, the S cannot be the cell. Can the S be the mereological sum of all the components of the mechanism? I will claim that this sum is not a ‘natural’ biological object, and it is better seen
as an occurrent and not a continuant. All this shows that it is better to view typical and paradigmatic biological mechanisms as etiological (rather than ‘constitutive’).

I will then present an alternative view of multi-level mechanistic explanation. The main ingredients of this alternative view are three basic claims: (i) biological mechanisms are causal pathways, (ii) levels and mechanisms are distinct notions and (iii) levels of nature and of multi-level explanations are levels of composition. A key claim of this account is that whatever contributes to the phenomenon is part of the same pathway; but causal pathways can contain entities at just one level of composition, or they may contain entities from multiple levels of composition. We therefore need to distinguish between ontological levels (i.e. levels of composition) and explanatory levels. While multi-level explanations contain entities from various ontological levels, ontological levels as such do not matter for explanation; what matters is the existence of a causal pathway.

So, according to this account, multi-level mechanistic explanations are causal explanations that identify particular causal pathways; but the components of the pathway are at different levels of composition: for example, a causal pathway may involve cells, hormones, and behavioural outcomes. This view of multi-level explanation is in stark contrast to the widespread view of multi-level mechanistic explanation as presented in Craver (2007), according to which ‘levels’ in multi-level explanations are not levels of composition but levels of mechanisms and mechanistic multi-level explanations are instances of constitutive explanations.

In order to motivate and illustrate the view, I will use particular biological examples: the pathway of visual perception, mechanisms of cell death, developmental mechanisms (axis formation and determination mechanisms, mechanisms for the generation of the tetrapod limb) and mechanisms underlying behavioural responses. The main aim of this paper will be to show that this account of multi-level mechanistic explanation applies to many different biological examples.

References


Irvine, Elizabeth

**Back to basics: Prediction and predictive processes**

While predictive processing (PP) accounts of the brain remain popular in cognitive science (Clark 2016, Hohwy 2020), there is surprisingly little in this literature on what it means for brain processing to be ‘predictive’, compared to anything else. In this paper I analyse what it means for a process to be predictive, argue that predictive processing (despite the label) is not best characterized as involving prediction, and that prediction may not play a significant role in perception/cognition more generally. This is because not all mechanisms that make it possible to infer hidden causes using stored knowledge, and to ‘look forwards’, are genuinely predictive.

To start, I identify key shared features of definitions of predictions (or predictive processes) from philosophy of science (e.g. Barrett & Stanford, 2006) and the case study of language comprehension (Pickering & Gambi 2018), where linguists were debating the role of prediction well before PP came on the scene. These features are:

1) Prediction involves making a claim (or e.g. activating a representation) about a matter of fact.

2) Predictions are generated by ‘indirect’ methods that make use of (pre-established) theories, statistical models or associations. ‘Direct’ methods are those based on, for example, observation and measurement, and sensory input more generally.

3) Predictive claims about a matter of fact are made before it occurs, or independently of claims made using direct methods.

I next analyse how these features apply to apparent instances of semantic prediction and prediction of speaker meaning, and then generalize to perception/cognition more generally.

I argue that semantic prediction, where recipients identify an upcoming word in a sentence, looks genuinely predictive. Before a word is uttered, there is plausibly no way that a direct method could be used to identify what that word is. So, where linguistic representations of a word are activated before it is uttered, perhaps using the kind of mechanisms associated with PP, or perhaps through priming, this looks genuinely predictive.

Identifying speaker meaning, the real target of language comprehension, seems to involve prediction too. Evidence from turn-taking (Stivers et al. 2009) strongly suggests that recipients routinely ‘guess’ at the meaning of an utterance well before the utterance is finished. Yet I claim that this is not best thought of as being predictive. To identify speaker meaning, recipients draw on a range of available information, both stored and current sensory evidence, in order to infer what meaning the producer is trying to get across. Even half-way through an utterance, a recipient may have a lot of evidence to draw on. In this sense, although the recipient identifies speaker meaning before the end of the utterance, they are usually heavily relying on direct methods to do so. Therefore, identifying speaker meaning is not (usually) predictive.

I suggest that a useful alternative for modelling this process is an accumulator model of decision making. In this kind of model, incoming evidence (e.g. stored or incoming sensory information) is evaluated to see how well it supports a range of decision options. As soon as one option is supported above an evidential threshold, the model outputs that option as the one to act on. Crucially, accumulator models that output decisions before ‘all’ available evidence has been.
processed cannot be characterised as making predictions: they generate well-supported decisions about matters of fact, based on direct methods.

From here, I argue that the question of whether perception/cognition in general can be described as predictive comes down to whether processing in the brain is more like semantic prediction or more like evidence accumulation used to identify speaker meaning. That is, it comes down to whether perception/cognition is more geared towards correctly understanding and navigating a discrete series of states or events in time, or whether it is more geared towards understanding and navigating continuous events or states. I will give some arguments to think that the latter is more likely. I also suggest that PP may be compatible with a modified version of an accumulator model of processing. Together, this means that while there may be a range of mechanisms that make it possible to infer hidden causes using stored knowledge, and to ‘look forwards’, neither PP, nor perception/cognition in general, should be characterised as primarily predictive.

References


Jalobeanu, Dana

Experimenting with artificial life: Francis Bacon’s *Historia de animato et modellino*

Among the unfinished projects one can find among Francis Bacon’s posthumous and fragmentary works there is an extremely interesting and under-investigated short fragment of a projected history of “animated” and “unanimated” beings. It is called *Historia et inquisitio de animato et inanimato*. The very title sets this fragment aside from Bacon’s other natural and experimental histories. *Historia* and *inquisitio* are, for Bacon, technical terms indicating two levels of philosophically informed experimental investigations; a preliminary gathering and organizing of observations and experiments and a more sophisticated, topically organized, experimental inquiry directed towards formulating (preliminary) causal correlations (Jalobeanu, 2015, 2016). The purpose of these experimental inquiries seems to be, on the one hand, to understand what distinguishes animated from inanimate bodies and, on the other, to create life in the laboratory. Unfortunately, the surviving fragment entitled *Historia et inquisitio de animato et inanimato* is insufficient to provide, by itself, a comprehensive picture of Bacon’s inquiry into artificial life. However, in this paper I claim that we do have means to reconstruct this project if we gather the recipes, experiments and preliminary inquiries on the matter one can find scattered in other Bacon’s works, most particularly in the posthumous *Sylva Sylvarum*. I will show how, taken together, these fragments outline a bold and innovative approach which blends natural history, experimental investigations and appetitive metaphysics into one of Bacon’s closest approximations of the workings of a properly organized, properly conducted “new science.” My general claim is that a better understanding of this example will shed new light on Bacon’s remarkable experimental methodology, while also unearthing useful insights into the more general, philosophical problems regarding the various stages of theory construction in experimental science.

Bibliography


Guido Giglioni (2016) Lists of motions: Francis Bacon on material disquietude. In Francis Bacon on Motion and Power (pp. 61-82). Springer, Cham.


Dana Jalobeanu (2016) Disciplining experience: Francis Bacon’s art of experimentation, in Perspectives on Science, 24 (3) p.324-342


Doina Cristina Rusu (2018). Same spirit, different structure: Francis Bacon on inanimate and animate matter. Early Science and Medicine, 23(5-6), 444-458.
Jarnicki, Paweł with Hajo Greif

The Aristotle Experience and the Genesis of the Kuhnian Revolution

Thomas Kuhn nourished a legend about himself according to which he single-handedly changed the way we think about science. This legend is founded upon the “Aristotle experience,” an event in 1947 that Kuhn retrospectively characterised as a revelation that made him instantly understand the fundamental difference between two scientific paradigms or thought styles. However, Kuhn formulated this narrative only in 1976, when he was informed about an English translation of Ludwik Fleck’s 1935 book “Entstehung und Entwicklung einer wissenschaftlichen Tatsache” being prepared. The first mention of the Aristotle experience is in the Foreword (signed in 1976) to the English translation of Fleck’s book (1979). The second description is in the introduction to The Essential Tension (1977) and the third in an article published in 1987, which was probably written in 1981 (lecture given at the Center for Cognitive Science, MIT). Later he discussed it in talks and interviews (1990, 1991, 1995). All (except 1981/1987) of the statements in which the Aristotle experience is mentioned are of an occasional or quasi-scientific nature.

Most of Kuhn’s readers, who, before getting to know Fleck’s book, got to know the Aristotle experience story, still believe this story, although the similarity between Fleck’s and Kuhn’s concepts is striking almost to everyone and it is well known that Kuhn read Fleck’s book as early as in 1949 (and thus before the Lowell Lectures, which are considered the beginning of his work on The Structure of Scientific Revolutions). It is most often assumed that Fleck’s book anticipated many central Kuhnian claims, so there is no or at most weak genealogical relation between the theories of Fleck and Kuhn. Usually, Kuhn’s arguments are repeated: that his understanding of Fleck was limited because of the (German) language barrier and that he found in Fleck what was already on his mind. We argue that there is circumstantial evidence that compels one to question the nature of the relationship between these theories.

First of all, Kuhn began to ascribe a special revelatory and inspirational meaning to Aristotle experience only around the time when he learned that the translation of Fleck’s book into English was being prepared. We know about the reaction of one person who first learned about Fleck and then about the Aristotle experience story – Wilhelm Baldamus’ early reaction to the similarity of both conceptions was to implicitly suggest plagiarism, because the degree of similarity between two theories is too great to be coincidental. We are not aware of any evidence that Kuhn came up with his ideas earlier than he found them in Fleck’s work. It is probable that even Kuhn’s interest in Gestalt psychology was awakened by Fleck (D.G. Cedarbaum interviewed Kuhn in November 1979: “Kuhn allows that Fleck’s discussion of, and emphasis on, Gestalt psychology may have awakened his own interest in the field”). We know that Kuhn not only read the library copy of Fleck’s book, but bought his own copy in which highlighting and annotations testify to careful reading. Kuhn prided himself on his ability to understand other texts. Kuhn did not quote Fleck even once. After the publication of the American edition of Fleck’s book, the editors felt compelled to authoritatively claim that Thomas Kuhn had not committed plagiarism. There is little evidence of Kuhn’s self-ascribed drive to publish Fleck in English, and he was reluctant to agree to cooperate with the editors of the American edition of the “Entstehung...”

In this talk, we will investigate the nature of this relationship along with all the circumstantial evidence at hand, sketching lines of further research that will help to determine what ideas Kuhn took from Fleck, and how. From parallel discovery to plagiarism, all possibilities will be considered.
References


We still don’t understand the mind/brain relationship in mental disorders very well, and consequently, we often help ourselves to analogies in trying to make sense of it. The hardware/software distinction is one such prominently used analogy philosophers of psychiatry introduce to argue that physicalism does not entail that all mental disorders are brain disorders. However, its precise use in this argument is underdeveloped. In this short piece, I briefly outline the argument, then show that there are a number of ways in which the hardware/software analogy is uninformative or even misleading. Looking at ways in which the mind/brain distinction is similar to the software/hardware distinction can only be a first step in investigating the relationship between mental dysfunction and brain dysfunction. Nevertheless, looking at similarities and dissimilarities between mind/brain and software/hardware is instructive, because it provides reasons to rule out some theoretical options and lays the foundation for further research.

How does the analogy between the hardware/software distinction supposedly show that mental disorders are not (necessarily) brain disorders?

The master argument

Computers can have software problems (programming errors) without there being anything wrong with the hardware.

The distinction between problems in the brain and problems in the minds is analogous to that between hardware and software problems.

Therefore, there can be a problem at the level of the mind (software) without a corresponding problem at the level of the brain (hardware).

But in what ways exactly are mind/brain and software/hardware analogous; and how far does the analogy extend? If successful, does it show that mental disorders can’t be brain disorders, or does it merely show that there is conceptual space for mental disorders that are not brain disorders? Some (eg. Graham 2013) take the analogy to show that a condition can be either a mental disorder or a brain disorder, but not both. But the (putative) fact that there is a dichotomy between hardware problems and software problems in computers does not establish that the same is true of human mind/brains.

Two features philosophers which have stressed in making the analogy are multiple realisability and causal difference makers (i.e., points of intervention).

Multiple Realisability

One commonality between computers and humans frequently highlighted when pressing the distinction between mental dysfunction and brain dysfunction is multiple realisability. In computers, software problems can be realised on different kinds of hardware (Papineau 1994, Kendler 2015), so software problems are problems independently of the way they are instantiated in the hardware.
However, the claim that mental disorders are multiply realised in the brain is hostage to empirical findings. To the extent that they are not, the analogy between computers and humans falls down. There are reasons to believe that the case for multiple realisability has been overplayed (Polger & Shapiro 2016), and that kind splitting (i.e. individuating psychological disorders more finely in a way that aligns them with their realizers) will remove a lot of apparent multiple realization of mental disorder. In as far as the analogy to the hardware/software distinction relies on multiple realisation, we would lack an argument that mental dysfunction is not brain dysfunction should we find stable realizers or identities.

Difference Makers

A further way in which mind/brain is supposed to be analogous to software/hardware is in having separate difference makers: software and hardware problems arise in distinct ways and are fixed in distinct ways. If you have a hardware problem, you call the hardware engineer, if you have a software problem, the programmer needs to make some changes. This then shows that there are distinct problems at different levels, which have a different etiology and are fixed (treated) in different ways. This way of thinking is deeply rooted in people who posit a clear separation of mental dysfunction and brain dysfunction. (cf. eg Borsboom et al 2018). However, this supposed analogy does not hold. As Kendler (2012) points out, there are many difference makers at different levels in psychiatric illness, both when we look at the etiology of disorders (risk factors) and at treatment options. Thus, there are psychological, environmental and biological risk factors for depression, and treatment can also target psychology, physiology or environment. If we think that mental dysfunction is something fundamentally different from brain dysfunction, then the multifactorial account would seem to exclude the idea that mental disorders can be brain disorders. However, an alternative conclusion is available: mental disorders and brain disorders should not be considered as dichotomous.

We have seen that only one of the two supposed analogies, multiple realisability actually holds up under scrutiny, and even this is hostage to empirical findings. Furthermore, there is one important disanalogy.

Disanalogy

One important, but underappreciated, difference between mind/brains software/hardware is that a key criterion for whether something counts as a hardware problem is the design specifications for hardware. We therefore have clear criteria for what the hardware should be doing in the case of computers in a way that we do not for brains. Brains are not artefacts, therefore we cannot get the criteria for proper functioning from the designer’s intentions. Several accounts of dysfunction are available in philosophy of biology and psychiatry, and it is an open question whether dysfunction at the level of the brain and dysfunction at the level of the mental have separate conditions of function and dysfunction. In fact, one theoretical option is that dysfunction conditions should be the same if we endorse an evolutionary account of function and dysfunction, as the brain and the mind have evolved together.
Given this important difference, more weight will have to be put on the existing similarities between mind/brain and software/hardware. This means that multiple realisability will have to do all the heavy lifting. Given empirical uncertainty about the extent of multiple realisation, it follows that drawing an analogy to the hardware/software distinction does not show that mental disorders are not brain disorders. At most, it tells us that if mental disorders are massively multiply realised, there is no corresponding brain disorder.

References:


The claim that all models used in economics are false (but that some are useful) is an old chestnut. But in practice, economists constantly face model uncertainty – there is always some amount of misspecification, in that any particular model will fail to capture all the details of the economic system of interest. Some of this misspecification can be attributed to the inevitable and necessary use of idealization and abstraction. What to do about such uncertainty is a pressing issue that needs resolution in practice; policymakers rely on models to make decisions – important ones, like how to set interest rates. Yet, it is difficult in economics to establish reliance on more than very coarse models.

This paper considers how economists in practice navigate the construction of a model for policy purposes. Of particular concern is when multiple models must be used, as is practice at central banks, and the final product of the economist’s exercise is not itself just an augmented model or combination of models. We have two aims. The first is to provide a descriptive sketch of economics as it is practiced at institutions like central banks – what does go on, anyway, when economists try to model something, and want to grasp the causal structure of an economic system in order to assess or perform interventions on it? The second is to show that the popular assumption that the “model-target” relationship is a sensible one needs to be re-examined. There are epistemically interesting things include activities and infrastructure that don’t explicitly appear in the final model at all, but are worth paying attention to. In addition, overlaying the popular “model-target” scheme on top of economists’ modelling efforts would be misleading given the way they do produce models.

First, we orient the discussion in the philosophy of science literature. We will distinguish our approach from other approaches involving multiple models to ameliorate model uncertainty, such as robustness analysis. Furthermore, we are not dealing with cases that ostensibly “stitch” different models together by linking them together. Then, we examine a concrete example: the integration of the Bank of England’s COMPASS model with information from its accompanying suite of auxiliary models via a MATLAB modelling toolkit (MAPS) that allows for this to be done in a structured way. Specifically, we discuss coordinating the core COMPASS model with another one that differs by including the financial sector. This coordinative activity – a substantial amount which occurs off either model and requires a good deal of expert judgment – is itself epistemically productive in a rather interesting way. Notably, a good deal of this expert judgment takes place in the information technology infrastructure that surrounds the COMPASS model. Finally, we will give an account of representational accuracy in a setting where (1) information but multiple models, which may be very different from one another, are integrated, and (2) there is a main model, as there is at the Bank of England case, is involved. The latter is of particular interest because in our case study the main model, even after integrating information from without, may not explicitly represent important and causally relevant sectors that indeed exist, such as the financial sector.

Representative References:


The explanatory limits of predictive processing: self-deception and the personal/subpersonal distinction

The prediction error minimization framework (PEM) denotes a family of views that aim at providing a unified theory of perception, cognition, and action (Clark, 2016; Hohwy, 2013). According to PEM, the brain enables adaptive behavior by minimizing the precision weighted prediction error (i.e. a mismatch between its hierarchical generative model and the environmental stimuli). The minimization of prediction error is accomplished either by perceptual inference (i.e. model updating) or active inference (i.e. changing the environment to fit the model-based expectations). PEM has already been used to explain a variety of neurocognitive abilities and phenomena, ranging from attention, emotion, action, and consciousness to psychiatric disorders. In this regard, Jakob Hohwy emphasizes the broad ambitions of the PEM paradigm, asserting that it has the capacity to account for “perception and action and everything mental in between” (Hohwy, 2013, p. 1).

Given its theoretical ambitions, PEM should be able to account for phenomena such as self-deception. Here, self-deception is construed as a type of motivationally biased belief-forming processes (see, e.g. Mele, 2001). According to this view, the desire that p be the case explains typical instances of self-deception by playing a casual role in the formation of the belief that p is the case. However, PEM does not have a separate category for motivational states such as desires (see, e.g. Hohwy, 2013, p. 89). Therefore, it seems that PEM cannot explain typical cases of self-deception.

It should be noted that self-deception is typically construed as a personal level phenomenon, while PEM is typically construed as offering subpersonal explanations of the underlying neurocomputational processes (Clark, 2016; Hohwy, 2013). Thus, when discussing the theoretical limitations of the PEM paradigm we should to be careful not to confuse the levels of explanation or misconstrue PEM’s domain of application.

Having this distinction in mind, I argue that supporters of PEM face a dilemma with respect to its (in)ability to explain self-deception. Either it shows that the PEM paradigm is more limited in its ability to “illuminate all aspects of perception and action” than advertised by its principal proponents (Hohwy, 2013, p. 1; see, also Clark, 2016). Or it is simply false that it can shed light on all phenomena that are of interest to cognitive science and philosophy of mind.

The argument proceeds by distinguishing between three views on the relation between personal and subpersonal explanations: autonomism, functionalism, and the co-evolutionary model (Bermúdez, 2005; see, also Colombo & Fabry, forthcoming). According to autonomism, personal and subpersonal levels of explanation form autonomous domains and there are no explanatory relations that span across subpersonal and personal explanations. The functionalist view includes the top-down methodology and construes the relation between personal and subpersonal levels as a relation between role functions and their underlying implementations. Alternatively, the co-evolutionary model allows that bottum-up considerations stemming from subpersonal explanations can be used to constrain and revise how we think about phenomena at the personal level (Colombo & Fabry, forthcoming). I argue that none of these views can save PEM from the criticism that it cannot appropriately explain self-deception.

On the one hand, adopting the autonomist view would insulate PEM from the objeetion, but at the cost of limiting its capacity to shed light on personal level phenomena. In particular, the autonomists see personal and subpersonal levels as constituting categorically different explanatory domains.
(Bermúdez, 2005, p. 51). Thus, if self-deception is a personal level phenomenon, PEM could at most be seen as providing necessary preconditions for its occurrence, but without explanatory import.

On the other hand, adopting the functionalist view or the co-evolutionary model leaves PEM vulnerable to the motivation-based objections. I argue that PEM, without a functionally independent construct of desire, cannot provide a sufficient explanation of self-deception that could at the same time distinguish it from other forms of biased reasoning.

In the case of the functionalist view, the argument can be summarized as follows. Given that PEM eschews desires as functionally separable states, we can imagine a person who is prone to confirmation bias in the sense of updating beliefs with information that favors their prior beliefs regardless of the desirability of the content of those beliefs. However, a person who self-deceives will show similar confirmation bias only with respect to the content of beliefs she cares about, desires to be true, and so forth. Therefore, given that at the subpersonal level, PEM construes all processes in terms of prediction error minimization, it cannot capture what is distinctive about self-deception as opposed to other types of biased reasoning.

In the case of the co-evolutionary model, I argue as follows. Adopting this model requires that we reconceptualize desires in terms of higher-order beliefs about what lower-level prior beliefs an agent finds rewarding. Accordingly, self-deception is reconceptualized as a phenomenon where a higher-order prior belief via active inference leads to action and belief-formation that is biased towards maintaining the lower-level prior belief about the desirable state of affairs. This reconceptualization leads to a problem. At the most general level PEM leaves open how prediction error (PE) will be minimized. The functional roles of desires and beliefs are determined by their world to mind and mind to world directions of fit, respectively. Prior beliefs are supposed to play both roles. Consequently, when prior beliefs confront PEs, there is no a priori constraint determining when PEs will be minimized by prior beliefs leading to action or by their being updated. Thus, no matter how high up in the hierarchy of inferences we go, prior beliefs are not, even under ceteris paribus conditions, sufficient to produce a self-deceptive belief.

References


The argument from inductive risk [A-IR] aims to impugn the value-free ideal of science (qua regulative principle) and demonstrate that scientists qua scientists must appeal to non-epistemic value judgements when drawing inferences from evidence (Churchman 1948; Rudner 1953; Braithwaite 1953, Ch.7; Douglas 2000, 2008, & 2009). The rationale undergirding [A-IR] turns on a conceptualisation of scientific reasoning as a process of decision-making under uncertainty, wherein inferences are recast as decisions to accept (reject) an hypothesis. Given the inductive nature of scientific evidence, acceptance (rejection) decisions face unavoidable risks of error. Therefore, when deciding whether the available evidence is strong enough to justify accepting (rejecting) an hypothesis, scientists must first assess the magnitude of error: ceteris paribus, the more consequential the mistake, the stronger the evidence required. And when errors carry moral, political, or social costs, error assessments will require scientists to make non-epistemic value judgements.

[A-IR] has engendered a bevy of replies, notably Jeffrey (1956). Therein Jeffrey argues that scientific inferences are not acceptance (rejection) decisions, but rather evaluations of evidential strength. And within an inquiry the sole remit of scientists is the evaluation of evidential strength via the assignment of probabilities to hypotheses. Consumers of scientific research are then at liberty to accept (reject) an hypothesis by deciding to use (or not to use) research findings to guide rational decision-making. For Jeffrey, the consumption stage necessitates the exercise of non-epistemic value judgements; the assignment of probabilities stage doesn’t. Commentators have labelled Jeffrey’s defence of the value-free ideal the oddsmaker view of the science (cf. Leach 1968 and Martin 1973), and it has since been endorsed by Hempel (1965). [Levi (1960, pp. 352-353) attributes a similar view to Carnap (1950, pp. 205-207).]

The oddsmaker view of science garnered criticism from defenders of [A-IR], however, beginning with Churchman (1956) and Leach (1968), and culminating most prominently in Douglas (2000). In a nutshell, Douglas argues that, at various stages of an inquiry, scientists face numerous choices regarding the implementation of methodology, characterisation of data, and interpretation of results. Each decision point plays a crucial role in drawing inferences from data and carries a corresponding measure of inductive risk. Douglas concludes that no assignment of probabilities to hypotheses at the end of an inquiry can capture the compounding epistemic uncertainties associated with each decision. Thus, scientists qua scientists simply cannot avoid making non-epistemic value judgements, and, therefore, the value-free ideal (qua regulative principle) falters.

Betz (2013, 2017) defends the value-free ideal (qua regulative principle) from Douglas (2000), I contend, by adopting a variation on the oddsmaker theme, namely, his full uncertainty disclosure proposal. Betz argues that scientists may avoid non-epistemic value judgements in scientific inference by fully and transparently disclosing all substantial uncertainties encountered at each decision point, and advertising only those hedged hypotheses which are certain beyond any reasonable doubt. Betz’s logic here is straightforward: if inferences are certain beyond any reasonable doubt, scientists face no reasonable risk of error — in which case, there exists no need to make non-epistemic value judgements regarding evidential sufficiency.

Herein I defend Betz from Douglas (2017), in which she argues:
(i) Full uncertainty disclosure cannot encompass the multitudinous risky acceptance decisions scientists must make regarding the employment of theoretical framework and methodological approach.

(ii) Betz’s defence of the value-free ideal presumes the only risk communicating scientific findings faces is the risk of advertising a false claim and, therefore, that it improperly ignores the risk of failing to provide actionable, decision-relevant results in a sufficiently timely manner.

Regarding (i), I argue that the acceptance decisions to which Douglas refers are rather inferences, and as such admit of probabilistic treatment through the use of robustness or hierarchical analyses (e.g., sensitivity analyses) and imprecise probability theory. Concerns remain, of course, about the interpretation and decision-relevance of imprecise probabilities, but their resolution falls within the purview of science advisory, not scientific inquiry. Moreover, under Betz’s view, scientific inquiry ought to proceed under a democratic division of labour, according to which the scientist’s brief is the provision of evidential assessments. And the aim of evidential assessments is purely epistemic: communication of the relative strength of scientific evidence and the specification of any outstanding scientific uncertainties.

Per their epistemic mandate, I contend, scientists ought to construct inquiries which maximise the probability of eliminating false hypotheses, conditional upon the prevailing ethical, methodological, and budgetary constraints enacted prior to the commencement of the investigation. How consumers of scientific research, including other scientists, ultimately use evaluations of evidence to adjust beliefs, inform future inquiry, and make decisions is for consumers to decide. However, should they desire more or less risk averse scientific research, stakeholders and democratic decision-makers are certainly at liberty to incentivise research centres and institutions appropriately via negotiations and funding decisions.

Regarding (ii), I argue that, once an inquiry has kicked off, the sole aim of the scientist should be to avoid producing misleading evidence assessments. For those scientific contexts where results are especially time sensitive, however, scientists may deploy an adaptive design, an inquiry that adopts a stopping rule which defines the manner in which decisions to close an inquiry and draw an inference depend upon the accumulating evidence and pre-specified evidential thresholds (e.g., posteriori probabilities, Bayes factors, or likelihood ratios). Decisions to adopt an adaptive design, however, occur in the design of experiment phase and, therefore, prior to an inquiry and any inferences made therein. And neither Betz nor any other defender of the value-free ideal deny that non-epistemic value judgements have crucial roles to play at this planning stage of scientific inquiry.

Whilst Betz’s defence of the value-free ideal is far more robust than Douglas’s critique lets on, I conclude that it possibly fails on largely the same grounds Jeffrey’s oddsmaker defence might, namely, that probabilistic beliefs and preferences (values) are inseparable.

Drawing upon work by Karni and Safra (1995), Karni and Mongin (2000), and Nau (2001), I argue that the choice theoretic devices (i.e., hypothetical bets whose payoffs are consequences assumed to have state-independent utility) employed by the majority of theories of decision under uncertainty to measure agents’ degrees of pure belief fail. In which case the probabilities assigned by Jeffrey’s oddsmakers very likely contain non-epistemic value judgements — at least, it’s not clear we can know that they don’t. Whilst Betz does not embrace Jeffrey’s radical probabilism, insofar as we may interpret “certain beyond a reasonable doubt” as a graded degree of belief that needs to be elicited under choice theoretic conditions, his practically certain scientists suffer the same fate as Jeffrey’s oddsmakers. I end by expressing scepticism that the inability to disentangle pure degrees of belief
from values presents any particular problem for defenders of the value-free ideal of science (qua regulative principle).

It’s entirely plausible that the preferences inextricably enmeshed with elicited degrees of belief are simply the backdrop against which scientific activity takes place -- in which case, the value-free ideal may serve to regulate the role subsequent values judgements, those not already embedded in assignments of probabilities and practical certainties, play in scientific reasoning. But, of course, it’s entirely plausible the value-ladenness of probabilities poses a serious challenge to the value-free ideal. I remain agnostic.


Extrapolating causal effects is a widespread epistemic activity in biomedical sciences, social and behavioural sciences, and in Evidence-Based Policy and Practice. It involves measuring the causal effect of an intervention in some study population and predicting the effects of the same or a similar intervention in a distinct target.

Extrapolation is often difficult as study and target settings can differ in important respects and extensive theoretical and empirical resources are needed to clarify how similarities and differences bear on an effect of interest. Theory is needed to help us tell which causal assumptions are needed for certain kinds of conclusions about a target and for helping us foretell how differences between study and target systems bear on an effect. Moreover, empirical evidence is needed to clarify whether particular similarities and differences obtain between study and target settings.

The existing literature has made important progress on elucidating the nature of extrapolative inference and the challenges it presents, developing general strategies for extrapolation, and emphasising the importance of theoretical and empirical resources for underwriting inferences sanctioned by such strategies.

What has received rather little attention is the role of uncertainty. Extrapolation invariably involves uncertainties: in principle, because clarifying similarities and differences between populations cannot require perfect causal knowledge of a target (Steel 2009) as that would obviate the need to extrapolate at all. In practice, causal knowledge is also often scant and crucial assumptions require additional support. Since many high-stakes instances of decision-making depend on extrapolation, such as deploying policy interventions that can cause not just benefit but also harm, there is hence a pressing practical need to articulate and manage the uncertainties that remain.

In this paper, I aim to make progress on these important issues by focusing on two questions: first, how can we express the uncertainty/confidence concerning specific causal assumptions and the causal hypotheses they entail? Second, how does uncertainty/confidence concerning specific assumptions compound and propagate onto a conclusion?

The first question can be usefully approached with bayesian networks (see e.g. Landes et al. 2018): by encoding evidential relationships between causal hypotheses, their observable consequences, evidential reports about whether these consequences obtain, and information about the reliability and relevance of these reports, bayesian networks help us compute probabilities for specific hypotheses given available evidence.

However, a bayesian networks approach is also limited: it cannot tell us how to assess the weight of evidence for specific hypotheses (i.e. how confident we may be in first-order probabilities) nor can it elucidate how different assumptions work together. The relationships between these assumptions are not (merely) evidential: we need to consider which assumptions are necessary to make specific inferences; how relevant they are individually; and how they interact, e.g. whether they need to be jointly satisfied, or are causally, logically, or probabilistically related. Existing work on extrapolation does not elucidate these issues either. Cartwright and Stegenga (2011) argue that an extrapolative conclusion is only ever as strong as the ‘weakest link’, i.e. the assumption enjoying the least support. But this seems false: just because assumption A enjoys less support than B, doesn’t mean that our confidence in a conclusion shouldn’t increase if we increased the support for B. Nor is every assumption equally important for an inference. So, to make progress on the second question, we
must consider not just how evidence bears on the probability of specific assumptions and the weight in their favour, but also how relevant these assumptions are for a conclusion and how they work together in enabling it.

To make progress on these aims, I sketch a hybrid causal-graph based approach, called support graphs. Support graphs involve three layers. The first encodes causal knowledge and assumptions: in the tradition of graph-based causal inference (Pearl 2009), structural causal models and corresponding graphs capture causal knowledge about the populations of interest and the assumptions required for an inference. The second layer is a support layer. Drawing on a bayesian networks approach, it encodes how available evidence bears on the assumptions contained in the first layer. The third layer is a relevance layer. It encodes how relevant assumptions about certain causal relationships are for a specific conclusion. For instance, assumptions pertaining to relationships that are outside of a ‘conditioning blanket’ afforded by the d-separation criterion are irrelevant. Assumptions concerning other relationships will differ in relevance. Their relevance can be investigated by performing sensitivity analyses, i.e. comparative static changes to the parametric and non-parametric details of the structural causal model/graph at the first layer to learn how a conclusion changes under these manipulations. Together with information from the second layer, performing such analyses also allows us to map out which conclusions enjoy how much confidence (cf. Roussos et al. 2020): e.g. coarse-grained conclusions might be more robust to changes in the validity of assumptions, so they enjoy more confidence than finer-grained ones.

In sum, by integrating distinct layers of analysis, support-graphs can help investigators clarify several interconnected issues in an integrated way: 1) which causal assumptions are needed for an inference, 2) how relevant these assumptions are for a conclusion, 3) whether they enjoy sufficient support, and 4) how confident we may be in certain kinds of conclusions. In virtue of these promises, detailing a support-graph approach can mark important steps in facilitating our ability to articulate, manage, and ameliorate uncertainties in extrapolation.

Bibliography


Will eliminating ‘innate’ block essentialist interpretations?

The word ‘innate’ is used both in biological and psychological sciences. Some philosophers and scientists argue that ‘innate’ should be expelled from scientific explanations (see e.g., Griffiths 2002; Machery et al. 2019). Call this position innateness-eliminativism. I will argue that a common argument that innateness eliminativists present for their position provides poor support for innateness eliminativism. I will first summarize the relevant argument for innateness eliminativism. I will then outline what I think is wrong with this argument.

Here is how the innateness eliminativist argues in support of her position.

‘Innate’ is used both in lay and scientific discourse. In lay discourse, ‘innate’ is associated with a concept according to which being innate implies being developmentally determined, being species-typical (Linquist et al. 2011). And some theorists, including some innateness eliminativists, hypothesize that this folk concept is nothing but the folk-biological concept of inner nature (Griffiths 2002; Linquist et al. 2011; Knobe and Samuels 2013) which is widely agreed to be developmentally robust species-universal feature of human psychology (Gelman 2003). The consensus view is that this folk concept of innateness/inner essence captures no theoretically interesting category. Are things different with ‘innate’ as used in scientific contexts? Views vary. Some think that in at least some scientific contexts, ‘innate’ has a technical and explanatorily useful meaning. Others disagree; they think that scientific innateness notions and scientific explanations with reference to innateness are but a useless relic of folk biology (e.g., Machery et al. 2019). One reason why innateness eliminativists think that ‘innate’ has no legitimate place in scientific explanations is precisely the alleged absence of an explanatorily useful scientific innateness concept. But there is also a second reason. Some innateness eliminativists are on the view that ‘innate’ should be expelled from scientific explanations even if it has (or had) a stipulated, explanatorily useful meaning in a given explanatory context. The reason is the following. ‘Innate’ is likely to trigger the activation of the corrupt folk concept of inner nature and thus give rise to essentialist interpretations of nativist explanations in terms of this concept. Such interpretations are false of, and give rise to false inferences from, scientific innateness hypothesis. In order to prevent such false interpretations and false inferences, one should stop saying ‘innate’ and express the intended technical meaning of the word in a way that makes this technical meaning transparent. Using Griffiths’ words:

“Substituting what you actually mean whenever you feel tempted to use the word ‘innate’ is an excellent way to resist this slippage of meaning. If a trait is found in all healthy individuals or is pan-cultural, then say so. If it has an adaptive-historical explanation, then say that. If it is developmentally canalized with respect to some set of inputs or is generatively entrenched, then say that it is.” Etc. If the best explanation of a certain trait differences in a certain population is genetic, then call this a genetic difference. If you mean that the trait is present early in development, what could be simpler than to say so?” (Griffiths 2002, 82)

I argue that this language-reform – eliminating ‘innate’ and substituting it with some other phrase – is unlikely to block the undesired essentialist interpretations of scientific theories.

Here’s how I argue.

Let’s take for granted that (1) in (the English) lay discourse, ‘innate’ is associated with the concept of inner nature. Let’s also take for granted that (2) the concept of inner nature is a developmentally fixed and universal feature of human psychology. Given (2) and (3), the association between ‘innate’
and the inner nature concept is surely contingent in the following two senses. First, given that the concept of inner nature is a developmentally fixed human universal, possessing and deploying the concept does not presuppose having ‘innate’ – or any other word, as to that matter – in one’s vocabulary. Second, other words and expressions besides ‘innate’ are, have been, can be, and probably will be associated with the concept of inner nature. This much is evident if only from the observation that there are words in English (e.g. ‘in the DNA’) that are sometimes used synonymously with ‘innate’ (Linquist et al. 2011).

Now, suppose that we do what the eliminativist calls for: we swap ‘innate’ for some other, arguably more accurate expression in a scientific theory. The innateness eliminativist envisages that by so doing we block the theory from being falsely interpreted via the folk concept of inner nature. But given (2) and (3), this is not the most likely thing to happen. Any suitable alternative to ‘innate’ (as those proposed by Griffiths) would be a word or phrase such that already resides in the semantic vicinity to ‘innate’. As such, it will be in the “gravitational field” of the ingrained folk concept of inner nature and is therefore no less likely than ‘innate’ to trigger the activation of the folk concept of inner nature. So, the relevant nativist theory is likely to be interpreted in terms of the inner nature concept despite change in words.

References


Humans face various threats throughout their life. In order to deal with these threats, evolution has put in place a number of cognitive systems. Our understanding of these systems is incomplete. Consider the famous case of the patient SM. Since the destruction of her amygdala, SM has ceased to respond to obvious threats, showing no fear while being held at gunpoint or handling snakes (Feinstein et al., 2016). Yet, SM is not fearless. She suffered a panic attack upon the inhalation of CO2-rich air and reported a great fear of painful medical procedures (Feinstein et al., 2016). Thus, SM only reacts to threat cues conveyed by the internal but not by the external senses. Why?

In order to explain SM’s selective deficit, Feinstein and colleagues (2016) have argued that there are two distinct fear pathways: exteroceptive and interoceptive. In SM, the exteroceptive pathway is damaged, but the interoceptive pathway is preserved. What remains unclear, however, is the structure of interoceptive fear. I seek to fill this gap by (1) developing a cognitive architecture of interoceptive fear and (2) presenting independent evidence for its existence.

The proposal is as follows. There is a single fear system activated by threat detection (Fanselow & Pennington, 2018). Threat detection involves the appraisal of some objects / events as dangerous, i.e. associated with some ‘bad outcome’ and thus requiring elimination. There are two threat detection mechanisms. The exteroceptive mechanism analyses the exteroceptive (visual, audial) percepts and tags some of them as threats (Feinstein et al., 2016). The interoceptive mechanism (IM) works differently, I maintain: it analyses the interoceptive percepts (e.g. pain, air hunger) and extracts from them the information that the body is under threat. From this, it infers the presence of threats. For example, if I break a wrist, then moving it will become dangerous as it will disrupt the healing process. But the only way I may learn about this danger would be through the pain that accompanies movement. Thus, I suggest that IM has evolved to track threats that cannot be tracked via the exteroceptive mechanism. I further suggest that IM is able to learn the threat value of the objects / events frequently associated with pain and initiate a response to them (e.g. flinching) when pain is only anticipated.

I then provide an excellent reason to accept my proposal: it can explain pain asymbolia (PA). Those who suffer from this condition retain the ability to experience pain, yet make no attempts to avoid or terminate these experiences. Instead, they willingly offer themselves to be tortured. Crucially, asymbolics also have a non-pain deficit: they do not react to some threats, like a match struck next to their face (Klein, 2015). To explain both deficits, Klein (2015) and Bain (2014) have argued that PA arises following a breakdown of the capacity for bodily care. Presumably, asymbolics fail to be motivated by pain or threats to their bodies because they no longer care for the said bodies. This view, however, has been challenged on the grounds that it is not well supported by the clinical data (de Vignemont, 2015).

My proposal provides a good, simple alternative to the bodily care view. In SM, the exteroceptive mechanism is broken. I suggest that in PA, the interoceptive mechanism is broken. This means that the threat value of internal percepts cannot be computed. The consequences of this are twofold. First, the affective processing of pain is inhibited, so pain is not experienced as aversive. Second, the threat elimination procedures (withdrawal, flinching) are not initiated. Note that this theory of PA is testable: if asymbolics panic upon inhaling CO2-rich air, then I am wrong.

Overall, my paper makes a contribution to the study of human threat defences by developing (1) an architecture of interoceptive fear and (2) a new explanation of pain asymbolia.
References:


Larroulet Philippi, Cristian

Against Prohibition (Or, When Using Ordinal Scales to Compare Groups is OK)

Standard measurement methodology distinguishes between ordinal scales and quantitative scales (e.g., interval or ratio). In quantitative scales, the differences between subsequent levels of the scale (called ‘the intervals’) are equal in magnitude. Say, the difference in temperature between 2ºC and 3ºC equals the difference between 4ºC and 5ºC. In contrast, in the case of an ordinal scale such as one measuring an attitude, it is said that we don’t know whether the distance between “strongly agree” and “agree” equals that between “agree” and “neutral.” That is why ordinal scales only provide rank orders.

A widely held view on measurement inferences, that goes back to Steven’s (1946) theory of measurement scales and “permissible statistics”, defends the following prohibition: you should not make inferences from averages taken with ordinal scales (vs. quantitative scales). For example, if you measure happiness in an ordinal scale, and a group of people scores higher on average than another group, the prohibition forbids inferring that the first group is happier than the second. This prohibition is general—it applies to all ordinal scales—and it is sometimes endorsed without qualification. Now, few scales in the social and biomedical sciences are widely considered quantitative. So the prohibition bears tremendously on their research practice. If adhered to, not much causal research can be done. No wonder, there is frequent discussion about the quantitative status of specific measurements, and deep disagreement about the legitimacy of the prohibition, between practitioners and methodologists (Michell 1990, Sherry 2011).

In this paper I challenge the general prohibition when applied to real-world scales typically considered ordinal (vs. formally defined ordinal scales). I assess the issue drawing from the leading contender for a general framework of scientific inferences: Bayesian epistemology.

The argument for the prohibition turns on two premises (see Michell 1990, 40-46): the set of admissible transformations that define ordinal scales (namely, all order preserving transformations) and the following invariance principle. ‘When inferring claims from measurement results, only the conclusions that remain true under all admissible transformations are validly inferred.’ The idea behind this principle is that all admissible transformations of a scale represent the phenomenon equally well. So, to infer a conclusion when derived only under some specific transformation would be premature. We need to verify whether it is derivable under every admissible transformation. Otherwise the conclusion doesn’t follow.

Alas, the argument given for the prohibition cannot justify it in its (commonly endorsed) unqualified form—namely, ‘no inferences can validly be made with averages from ordinal measurements’ (e.g., Merbitz et al 1989, Michell 1990). The reason? The unqualified prohibition overlooks the argument’s invariance clause. That is, if the relative average comparisons between groups made with ordinal measurements remain invariant (e.g., group 1 always has a higher average than group 2, no matter which admissible transformation we use), then the argument allows for the inference. This is no hair-splitting—the invariance clause holds sometimes, leaving the relevant inference unchallenged by the argument. I illustrate with a real example (from Easterlin 1974), and provide a formal way of stating when the relative comparison will be invariant by using the mathematical notion of first order stochastic dominance. Thus, pace Michell, when the differences among groups are strong enough (so that stochastic dominance holds), the standard argument cannot block inferences about relative comparisons.
But I believe we can go beyond this qualification of the prohibition and arrive at a rejection of it as a applied to real-world scales, once we move to a subtler epistemic framework; one that allows for belief states between knowing that all intervals of a scale are equal (so that the scale is quantitative) and knowing nothing about how they compare (so that the scale is ordinal). Bayesian epistemology can be of help here.

I propose we can assess this issue as an inferential problem, by modelling the difficulty involved in working with non-quantitative scales as a case of “unreliable evidence” (cf. Howson and Franklin 1994). In our context, the source of unreliability comes from the potential heterogeneity of the intervals (i.e., differences between levels). We can picture a researcher measuring with a given scale, and represent her beliefs about the relative sizes of the intervals of this scale with a probability distribution. If she is certain that all the intervals are equal, she can compare groups with averages without uncertainty, since this amounts to having an interval scale. I show, however, that the rational beliefs (about intervals’ differences) needed to confirm relative average hypotheses are not as strong as the prohibition states. That is, they need not be beliefs adequate for a quantitative scale. Much less is needed. The pivotal issue for comparing groups ends up being how different in magnitude specific (vs. all) intervals of the scale are, and which intervals these are depends on the specific evidence (i.e., the measurement performed for a given comparison). More generally, I argue, the epistemology behind the prohibition is too blunt to be of use in all cases. It forbids sound inferences by over restricting both the possible credences of researchers about intervals’ differences and the possible kind of inferences (to only deductive ones). Indeed, I show that trying to adhere to the prohibition in this Bayesian framework can lead to incoherent credences.

Then I address objections. Most important among them is that researchers working with ordinal scales cannot rationally hold the relevant beliefs. My response highlights the narrow view of evidential context here presupposed, and the gap that exists between the standard definition of an ordinal scale and real-world scales considered ordinal. After all, most scales deemed ‘ordinal’ in actual research are so only because they fail to be quantitative, not because they have been shown to be strictly ordinal. To make my point, I elaborate on the paradigm example of an ordinal scale: Mohs’ scale of minerals’ hardness. I show that the literature has mischaracterized it.

This paper, thus, offers a principled response to the prohibition. Moreover, it offers a potential resolution to the tension between practice and methodology that besieges measurement scales.
Quantum entanglement is a key ingredient of quantum mechanics, and might even be central for the very distinction between the quantum and non-quantum theories. In a composite system made of several single sub-parts, this phenomenon refers to the situation for which it is impossible to attribute independent definite states to each of the sub-systems. Instead, the composite system has to be considered as a whole, and its sub-parts are said to be non-separable.

Philosophically, quantum entanglement has important implications. First of all, it is at the core of the main conceptual puzzle arising from the development of non-classical physics, known as the measurement problem. Secondly, entanglement allows for the observation of very peculiar, non-classical correlations between spacelike-separated events, which cannot be explained by any local causal model and are thus said to be nonlocal.

Overall, many interpretations of quantum mechanics have been developed in order to solve the measurement problem and provide an account for nonlocality (in terms of its nature, origin and properties). Each interpretation commits to a particular set of assumptions regarding the properties of reality and space-time, by specifying the ontology of the theory and its dynamics. Which of those interpretations is the most successful is still an ongoing debate.

Yet, entanglement is not an exclusive property of standard quantum mechanics. Instead, it is also found in further theoretical developments of quantum physics, e.g. quantum field theory, and is expected to remain an important feature of a future theory of quantum gravity. Investigating the philosophical implications of entanglement in this broader theoretical context could shed a new light on the conceptual problems in quantum theory. Indeed, since (i) the way entanglement and nonlocality are accounted for is partly conditioned by our conception of spacetime, and (ii) a radical shift in our conception of spacetime is expected to take place as quantum physics is developed into more general theories (where gravity would be ultimately taken into account), looking at entanglement in new theoretical frameworks generalizing standard quantum mechanics could help us understand the nature of entanglement and its connection to space and time in a radically new way. Identifying the relation between entanglement and spacetime is crucial for developing a consistent ontology of reality.

For these reasons, this work focuses on a particular theoretical development of quantum mechanics, called the process matrix formalism (Oreshkov et al. 2012). Within this more general theoretical context, we investigate the connection between quantum entanglement, spacetime, and new phenomena predicted by the formalism called causal non-separability and noncausal quantum probability distributions (Oreshkov & Giarmatzi 2016, Branciard et al. 2015).

The process matrix formalism is expressed within the operational framework for physical theories. This operational framework allows one to formulate different physical theories using the same conceptual tools, independently of any device or physical model (Brukner 2014). In essence, a physical experiment is encapsulated in a probability distribution correlating the initial states of the system with the possible outcomes of a measurement on this system. Hence, physics is described by operational notions such as preparation and measurement procedures in the context of an experimental setup. Such a framework is useful namely to reconstruct quantum mechanics from a set of basic principles, or axioms, that any quantum experiment needs to satisfy. These principles provide an interesting basis to study the physical foundations of quantum mechanics, as they are based on concrete operational processes and avoid the use of abstract, non-directly observable
concepts like the ones referring to some mathematical objects such as Hilbert spaces. Going further, the operational reconstruction of quantum mechanics can be generalized by modifying its basic set of principles from which it is derived, in order to broaden the range of physical experiments that are allowed by the theory. This results in a generalization of the theory, possibly predicting real natural phenomena not described by standard quantum mechanics.

The process matrix formalism is precisely such a generalization of quantum mechanics. It was obtained by dropping the basic principle according to which the causal structure of any physical experiment is always well defined prior to any measurement. That broader theoretical context predicts the existence of quantum processes (a generalization of the concept of quantum state allowing to represent quantum-like correlations between events over multiple parties without mentioning their spatio-temporal locations.) that are causally non-separable, for which one can see some analogy with the spatial non-separability involved in entangled systems. For a causally non-separable process, there is no definite causal order among its interacting parties. The corresponding causal structure is said to be indefinite. In operational terms, the probability distribution encapsulating the results of a measurement performed on a system can violate the causal equivalent of Bell inequalities, called causal inequalities. Such a distribution is said to be noncausal.

The objective of this study is, first, to discuss the connection between the notions of quantum and causal nonseparability. More precisely, since the notion of causal nonseparability is inspired by some analogy with quantum nonseparability, we will investigate the extent and the limits of this formal analogy. Secondly, in a realist framework, we will have a preliminary reflection regarding the potential implications of standard and causal nonseparability for the notion of spatio(temporal) relations. Regarding those questions, two corresponding statements will be defended. First, while quantum processes can indeed be seen as generalisations of quantum states, the two notions are mathematically and, most importantly, conceptually distinct concepts. Secondly, causal nonseparability implies some kind of indeterminacy of spatiotemporal relations. It is argued that, in spite of the disanalogies between standard and causal nonseparabilities, such implications for spatial relations can already be defended in the context of standard quantum mechanics.

References:


This paper contributes resources from the philosophy of science to identify differences in explanatory norms across disciplines and analyse how such differences affect interdisciplinary collaboration. Explanatory norms are the implicit and explicit rules that govern what scientists consider to be explanatory sound. Examples of explanatory norms are generality, tractability, precision, etc. Explanatory norms are domain-specific insofar as certain domains tend to praise some norms over others, for instance, a domain may favour tractability over generality; and even when different domains endorse the same explanatory norm, such as generality, they may have different standards for what counts as an adequate level of compliance with that norm (see, on this, Marchionni 2013; Woody 2003, 2015).

The body of literature on domain-specific explanatory norms is rapidly growing and providing increasing evidence on the impact of such norms on interdisciplinary collaboration (Knuuttila & García Deister 2019; Love 2015). What emerges from this literature is that the standards of validity of knowledge claims are often domain-specific and that differences in explanatory norms channel and constrain the kind of interdisciplinary interaction that takes place across disciplines.

In the literature, however, there is still no consensus on a theoretical framework that allows us to systematically compare norms across disciplines. The lack of a conceptual background that guides this analysis makes it difficult to proceed in a more structured manner. The extrapolation of explanatory norms is a rather unconstrained process; the literature is scattered around a variety of case studies, each of which is tight to the specificities of the domain under study. Ultimately, this hinders general philosophy of science work on the topic.

The aim of this paper is to provide a methodological framework that enables 1) to draw a comparative analysis of explanatory norms across fields; and 2) to use differences in explanatory norms as an indicator of distance between disciplines, which affects interdisciplinary work accordingly. To pursue this aim, the paper will build on philosophical work on explanation as a starting point to identify domain-specific explanatory norms.

A particularly promising candidate in this field is the work on dimensions of explanatory power by Ylikoski and Kuorikoski (2010). Dimensions of explanatory power refer to the features that qualify the goodness of an explanation over another, i.e., those features that make an explanation better or deeper than another, even when they are both explanations of the same phenomenon. Ylikoski and Kuorikoski identify five dimensions of explanatory power, i.e. non-sensitivity, precision, factual accuracy, degree of integration, and cognitive salience. To illustrate, non-sensitivity indicates the degree to which an explanation is insensitive to changes in background conditions. The less sensitive an explanation, the more powerful it is. Factual accuracy refers to the idealizations that are considered to be adequate for an explanation. The fewer the idealizations, the more powerful the explanation, etc.

Overall, the previous approach is particularly attractive in the context of a comparative analysis of explanatory norms and interdisciplinarity for various reasons, among which that it enables different disciplines to have their own disciplinary norms, which may pull towards different directions. But the account has been mainly developed in the context of explanation. It still needs to be shown: i) that specific (levels of) dimensions of explanatory power track disciplinary explanatory norms. This way, scientific domains can be represented as disciplinary profiles that include disciplinary norms; ii) how to link differences in disciplinary profiles to patterns of interdisciplinary collaboration.
To address the two issues above, the paper proceeds in the following way: it compares Ylikoski and Kuorikoski’s account to other accounts developed in the literature on dimensions of explanatory power and depth (e.g., Hitchcock & Woodward 2003); it connects it to the recent literature on explanatory norms in general philosophy of science (e.g., Woody 2003); and to the literature from the philosophy of science in practice that analyses interdisciplinary case studies (e.g., MacLeod & Nagatsu 2018).

In tandem with this, the methodological framework will be developed further on the basis of the results of a survey that identifies explanatory norms and their role in interdisciplinary collaboration. The survey collects data from scholars from different fields and experts on interdisciplinary science on items concerning explanatory norms and how their differences predict challenges in interdisciplinary work.

The paper concludes by showing the relevance of its result for the assessment of interdisciplinary research. The survey itself offers a concrete tool that can be used by science-policymakers to incorporate differences in explanatory norms as a factor that affects research outcomes. Intuitively, the closer two disciplines are with respect to their explanatory norms, the more suitable their disciplinary criteria for the evaluation of interdisciplinary projects, and vice versa. This part of the work will contribute to the “much awaited institutional turn in philosophy of science” (Mäki 2016, p.337), by offering an example of the way in which methodological and institutional factors are intertwined in shaping scientific research and, thus, of the importance to inform science-policy decisions on the basis of philosophy of science work.

References


Two birds with one stone? Not when arguing for Bayesianism and Credal Veritism

Theories of credences give rationality constraints for the degrees of belief of idealised agents. The dominant such theory is Bayesianism, formulating rationality constraints in terms of probability theory. It has a very impressive track record of fruitful applications in philosophy of science, epistemology, metaphysics, decision theory, economics, cognitive science, and more. Although dominant, Bayesianism is not without rivals. For example, ranking theory is a well worked out alternative to Bayesianism. For this and other reasons, it is worthwhile for Bayesians to argue for their theory.

One recently popular strategy to do so is set in the framework of accuracy-first epistemology: Bayesians have started to formally characterise measures of accuracy, i.e. closeness to truth, and give theorems to show that the Bayesian norms are in some sense accuracy-conducive. As a next step, accuracy-firsters consider the idea that accuracy is the sole fundamental source of epistemic value. Richard Pettigrew calls this claim Credal Veritism (2016b). This is not to say that there are no other epistemic values, but only that their value derives from the value of accuracy. Now, if Credal Veritism is true, it seems that the accuracy-conduciveness of the Bayesian norms speaks in favor of them. But why think that Credal Veritism is true? One popular line of reasoning considers other proposed fundamental sources of epistemic value and attempts to show that their value in fact derives from the value of accuracy. For example, Pettigrew explicitly subscribes to this strategy (2016b). This paper argues that Pettigrew’s way of providing this reasoning for Credal Veritism is question-begging against dialectical opponents like ranking theorists.

It is important to note that there are at least two views one can have on the accuracy-first programme. First, there is what I call the deductive argument view. According to this view, accuracy-firsters are supposed to provide a list of true premises from which the relevant Bayesian norm deductively follows. Importantly, for these arguments to be convincing, the premises have to be justified in a non-question-begging way. Second, there is what I call the coherent package view. According to this view, the primary goal of the research programme is to show that there is a coherent package of Bayesian norms (Probabilism and others) and epistemic axiology (Credal Veritism and measures for accuracy). A non-question-begging motivation for the epistemic axiology is not necessary, though maybe it should not be too harshly at odds with common sense. On the downside, however, the existence of such a coherent package by itself does not provide an argument for Bayesianism. In fact, ranking theorists have also proposed an axiology that fits their norms (Raidl and Spohn, 2020). To my knowledge, nobody has carefully compared available packages. Thus, given the coherent package view, accuracy-first epistemology has not yet delivered a complete argument for Bayesianism.

It seems to me that the coherent package view is more tenable. However, what is presented in the literature is the deductive argument view. For example, Pettigrew (2016a) speaks of delivering valid arguments for the Bayesian norms and at times explicitly tries to avoid question-begging justifications for the premises. Thus, it is fair and worthwhile to critically examine the success of accuracy-first epistemology given the deductive argument view. The present paper is supposed to contribute to this by showing that Pettigrew’s argument for Credal Veritism is question-begging.

Pettigrew (2016b) wishes to show that other proposed fundamental values in fact derive from the value of accuracy. More precisely, he considers the values of coherence (the credences in different propositions fit together) and evidential proportionality (the credence in a proposition matches the proposition’s degree of support given by the available total evidence). His defense of Credal Veritism
consists in two steps, here illustrated with coherence. First, he claims that coherence is precisely that which is required by Probabilism, i.e. a credence function is coherent iff it is probabilistic. Second, he points out that the accuracy-conduciveness of Probabilism (which is independently argued by accuracy-firsters) seems to show the following: if we suppose that accuracy is valuable, then it follows that coherence is valuable, because being probabilistic (thus, being coherent) promotes accuracy. That is, the accuracy-conduciveness of Probabilism shows that the value of coherence is derivable from the value of accuracy.

The paper argues that the first step of this reasoning is question-begging. It cannot be presupposed that coherence is precisely that which is required by Probabilism. For example, ranking theorists put forward their own account of coherence: a credence function is coherent iff it is a ranking function. The same holds for evidential proportionality. The paper uncovers that, even though accounts of coherence are not strictly implied by the respective norms, the relationship between the two is too close for the accuracy-firster to presuppose her account of coherence. Thus, this defense of Credal Veritism cannot be put forward as a justification for a premise in a convincing valid argument for a Bayesian norm. The accuracy-conduciveness of the Bayesian norms is supposed to play a double role in accuracy-first epistemology: on the one hand, it appears as a premise in accuracy-first arguments for Bayesian norms. On the other hand, it is used to defend another premise, namely Credal Veritism. If Bayesianism and its accounts of coherence and evidential proportionality could be presupposed, then Credal Veritism could be defended on grounds of Bayesian accuracy-conduciveness. And if Credal Veritism could be presupposed, then Bayesianism could be defended on grounds of Bayesian accuracy-conduciveness. But the accuracy-firster cannot kill both birds with the same stone.

References


Against Symmetry Fundamentalism, in Defense of Symmetry Deflationism

Symmetry fundamentalism holds that symmetries are features of physical reality and should hence be employed as guides to what’s fundamental. In this presentation, I call for philosophical caution when symmetry fundamentalism is employed for metaphysical research. First, I distinguish between two views on symmetries—by-stipulation and by-discovery. Then, I argue that the role that symmetries play in modern physics supports, to a great extent, the by-stipulation view, which conflicts symmetry fundamentalism. It follows from this that symmetry fundamentalism is not adequate to construe the role that symmetries play in current physics since it overloads their ontological import. I conclude that symmetry deflationism, instead of symmetry fundamentalism, should be recommended.
Wave function realism (WFR) stands out as a prominent interpretative option concerning the ontology of quantum mechanics (see Albert 2015, Ney 2012). Its central tenet is that the key ingredient of quantum mechanics, i.e. the wave function, is a physical and concrete object. In particular, it is a field, defined over the points of an incredibly high-dimensional space. But, even though the defenders of this view please this interpretation of the wave function as strikingly elegant and natural, WFR faces a straightforward crucial challenge: how can we account for our ordinary three-dimensional space and for spatially located objects, given that the fundamental ontology inhabits such an alien space? The mainstream solution, shared by most of the wave function realists, has been to adopt a functionalist approach to the problem, similar in spirit to the functionalist theory which is widespread in the philosophy of mind. In a slogan, functionalists say that ‘there are tables as long as there is something which plays the table-role’. As long as we can deliver a satisfactory story of how the wave function can play the causal and functional role that we associate with three-dimensional objects, we can make sense of the existence of tables. In this way, not only can we account for our appearances, but - in a certain way - it can be maintained that three-dimensional objects are still real, even in a world fundamentally built out of a wave function. In other words, they are real but they are also deemed as derivative or secondary entities, belonging to a derivative space, which is distinct from the high-dimensional space where the wave function is located.

However, in this talk, I want to highlight an aspect of WFR which has remained deeply unclear and which has been partially overlooked across the literature. In fact, even though the appeal to functionalism can account for the existence of three-dimensional objects, it does not settle the ontological question about the relation between the latter entities and the wave function. That is, what is the ontological relation between the fundamental ontology and the three-dimensional one? What is, precisely, the ontological status of the derivative ontology? Functionalism is compatible with different views about the nature of the functionalized entities. Should we endorse some form of reductionism? Do three-dimensional objects exist over and above the high-dimensional ontology, at a different level? Within the literature, the relation between the two levels is often sketched in various different ways, using notions that are usually considered to be non-equivalent, and sometimes also incompatible (e.g. reduction, emergence, realization, and so on). Thus, it is often unclear which is the exact ontological commitment of WFR. My main aim here is to develop a framework that can clarify this issue.

I will show that - once properly formulated - a clear ontological picture stems from the combination of functionalism and WFR. In particular, I shall defend a form of functional reductionism that was originally proposed by David Lewis (1970, 1972), and that has been recently defended by Jeremy Butterfield and Henrique Gomes (2020a, 2020b) within the philosophy of physics and by Jaegwon Kim (2005) in the philosophy of mind. This is a reductionist version of the functionalist account, which entails type-identity relations between the functionally realized entities and those entities belonging to the more fundamental ontology which play the role associated with them. Thus, in the end, what we really have is just one level, ontologically speaking. According to Lewis, this view is not only preferable, but it is mandatory, once we embrace a functionalist account similar to the endorsed by the defenders of WFR.

The framework is not without pitfalls, however. In fact, the intuition that the wave function determines the three-dimensional ontology – via functional realization – and is, therefore, more fundamental seems sensible, and it seems to be lost if we adopt an identity theory. It is the three-
dimensional ontology that emerges from the wave function, and not the other way round, of course. In other words, the wave function seems to be asymmetrically related to the three-dimensional ontology, but identity is not an asymmetrical relation. Thus, we should try to find a way to preserve the crucial intuition about asymmetry, making it consistent with the Lewisian reductionist view, in such a way that ontological identity can be combined with a 'multi-level' view of reality. I will sketch a strategy to satisfy this desideratum, based on Christian List's (2019) formal account of levels. He distinguishes between ontological levels and levels of description.

Ontologically speaking, the entities which are functionally reduced are straightforwardly identical, since they are co-extensive. In this respect, they belong to the same level. On the other hand, however, those entities are originally picked out by different theories, which introduce them via different languages, at different levels of description - and those levels are hierarchically ordered through supervenience mappings (between the theories). Thus, even within the Lewisian functional reduction model, there is still room for saying that reduction embeds asymmetrical relations: it is not ontological asymmetry, but asymmetry of description.

Summing up, my aim for this talk is twofold. Firstly, I will propose a Lewis-style functional reductionist account of wave function realism, that provides a clear picture of the ontology of the theory, especially concerning the relationship between the high-dimensional wave function and the three-dimensional ontology (i.e. identity). Secondly, I will propose a strategy to reconcile identity with the alleged asymmetry between the wave function and the derivative spatial ontology, by distinguishing between ontological levels and levels of description.

Finally, I want to stress that the scope of this discussion goes well beyond the debate on WFR. For example, a similar form of functionalism to the one proposed by Ney and Albert is currently at the center of the debate within the philosophy of quantum gravity. Indeed, some authors (see Lam and Wuthrich 2018) have proposed to frame the emergence of spacetime within quantum gravity theories in terms of functionalism. The account which I propose can be naturally extended also to that context.
I explore the multifaceted relationship between mathematical rigor in scientific representations and the ability of those representations to support scientific understanding. I consider two historical uses of unrigorous mathematics that have been argued to promote better understanding than their rigorous counterparts: "intuitive infinitesimals" in calculus and Heaviside's operational calculus in electrical engineering. Despite these arguments, each unrigorous technique nevertheless fell out of favor (the recent pedagogical resurgence of infinitesimals notwithstanding). I argue that this apparent tension can be explained by the sometimes conflicting contributions of mathematical representations to scientific understanding. Unrigorous techniques have both costs and benefits for scientific understanding and persist in the face of rigorous techniques when these benefits cannot be achieved rigorously.

To streamline the discussion, I frame these contributions to understanding in terms of Kuorikoski and Ylikoski's (2015) factive inferentialism about scientific understanding, though they could be reframed in terms of a number of other views. I think this lends itself most naturally to thinking about unrigorous mathematics, particularly because the cases in question seem to support an inferential account of mathematical representation as well (McCullough-Benner 2020). According to factive inferentialism, scientific understanding of a target phenomenon T consists in the ability (in practice, not in principle) to make appropriate counterfactual inferences about T. The view is factive in the sense that it requires that the features of T underwriting these inferences are accurately represented. This means a mathematical representation can contribute to scientific understanding either by making more (salient) inferences available or by making inferences more reliable.

An unrigorous mathematical representation might promote scientific understanding better than its rigorous counterpart either by making the relevant mathematical problems more tractable, thereby making more inferences practically available, or by reducing human error by simplifying calculations. Since the relevant mathematical objects are not clearly well-defined, this comes at the cost of making it less clear what inferences are licensed by the representation. This impedes understanding both by making the accuracy of the representation more difficult to evaluate and by introducing the possibility of human error in carrying out inferences apparently licensed by it. When these benefits exceed the costs, there is good reason to favor the use of unrigorous mathematics and vice versa. This pattern is exhibited by both examples I consider.

Before the calculus was put on a rigorous foundation in the 19th century, inconsistent infinitesimals made new counterfactual inferences accessible, particularly in classical mechanics. The switch to a calculus based on rigorous limits in the 19th century made little difference to the practice of physics, as expressions like "dy/dx" could still be manipulated as if they referred to ratios of infinitesimals, even while one understood them to really pick out the limit of a ratio of finite quantities. So there was little to lose in adopting the rigorous calculus. (This isn't always the case. For example, the Dirac delta function largely continues to be used instead of its rigorous distribution-theoretic counterpart because the delta function is much easier to work with.)

On the other hand, for pedagogical purposes, infinitesimal approaches seem to have been more intuitive than approaches based on rigorous constructions of the real numbers and epsilon-delta limits, allowing students to learn how to reason using the calculus without the significant overhead of first having to understand expressions like "dy/dx" in terms of limits and thereby promoting better understanding (through greater inferential competence) when the calculus is applied to simple physical problems. Indeed, this is the thrust of many contemporary arguments for using
infinitesimals to teach introductory calculus to non-mathematics students. Infinitesimal-based textbooks continued to appear until at least the publication in 1912 of the 12th edition of Kiepert’s Grundriss der Differential- und Integral-rechnung, one of the most popular calculus textbooks in Germany in the late 19th and early 20th centuries.

In contrast, the transition from Heaviside's operational calculus to more rigorous methods, particularly Bromwich's method of using the Laplace transform and its inverse (the Bromwich integral), was much more difficult, and so it remained in use for some time after the introduction of these more rigorous techniques. While Heaviside's methods were notoriously unrigorous, they consistently led to correct physical results, usually in a very small fraction of the space it would take to derive the same results rigorously. In practice, this made electromagnetic problems more tractable, thereby contributing to superior understanding of that domain. Bromwich himself even wrote to Heaviside that he was primarily using Heaviside's methods, relegating rigorous proofs using his own methods to footnotes. On the other hand, Heaviside’s techniques were notoriously difficult for students to master, not least because Heaviside had no interest in articulating in general terms the conditions under which they could safely be applied, insisting that it was better to learn when these techniques could be applied on the basis of experience (Heaviside 1899, §282, p. 128). Perhaps in the long run this could lead to greater inferential acuity than a more rigorous approach, and so have some payoff in terms of understanding; however, in addition to making his techniques difficult to learn, it was a significant source of human error that did not affect more rigorous approaches. As the rigorous techniques were further developed, it became common to find the Laplace transforms and inverse Laplace transforms using lookup tables for common functions rather than calculating them by hand (a tedious task in non-trivial cases). At this point, the operational calculus no longer made more inferences accessible than the more rigorous techniques, since it was as easy to derive the same results rigorously, and so it was ultimately replaced by them.

In each case, an analysis in terms of the contributions of unrigorous mathematical theories and their rigorous counterparts to scientific understanding helps us to explain both the positive epistemic status of the unrigorous techniques, as well as why they were eventually replaced.

References:


Pantazakos, Themistoklis

Perceptual Pluralism and Theory-Ladenness of Observation

The theory-ladenness of observation debate has taken on many forms since its advent, now more than fifty years ago. While all of these forms retain a central problematic, namely the potential undermining of scientific objectivity by acting subjective factors, I suggest that discussing all forms in bulk often serves to obfuscate rather than enlighten the issue. Instead, we should clarify what observation amounts to by analysing it into its constituent parts, specify which candidate factors we are considering each time and which parts of the observation process they purportedly act on, as well as how the various kinds of ladenness may feed off each other. For this reason, the first aim of this paper is to provide a model of theory-ladenness that brings the issue into systematic view.

By drawing on contemporary advances in neuroscience, I analyse observation as a five-stage path consisting of: sensation; perception; observation; data; phenomena. ‘Sensation’ refers to all processes that result in the production of the retinal image and regards light intensity computation. ‘Perception’ introduces structures with boundaries and regions, as well as shapes, spatial relations, contours, textures, and other general characteristics – Marr’s ‘2 1/2D’ sketch of the passing scene. ‘Observation’ regards objects proper and the ascription of meaning to the the scene. The systematization of successive observations comprises a set of data, while phenomena are abstractions from data, most commonly found in causal form.

These five stages are what subjective, ladenness-responsible input factors may act on. Regarding the factors themselves, the literature suggests the existence of several kinds, of which the most central are linguistic frameworks; conceptual schemes; prior beliefs; the architecture of perceptual systems; environmental cues. Each candidate factor acts typically on just one stage. This analysis suggests a two-variable, input-output model for the theory-ladenness of observation, whereby each thesis under its rubric becomes a token case of the following general statement:

Theory ladenness of <STAGE> by <INPUT>: the contents of <STAGE> are, among others, a function of <INPUT>.

This consists a model of bringing the theory-ladenness debate into general view. I put forward that the model is instrumental in evaluating the epistemological significance of X-ladenness, as it prompts one to specify at which stage down the path it happens, whether that ladenness is likely to be undone downstream and, if not, whether it is likely to affect a scientific statement’s truth-value.

The particular case of theory-ladenness that I interrogate within this paper is the cognitive penetrability of perception, and thus the contention that the contents of perception are, among others, a function of cognition. Claims against such penetrability were inaugurated by Jerry Fodor and are carried forward today by, among others, Athanasios Raftopoulos. Fodor’s and Raftopoulos’ arguments revolve around two axes, which I name isolation and commonality. Isolation holds that perception (here: early vision) is insulated from the higher cognitive centres (e.g. the prefrontal cortex). Simply put, the higher cognitive centres, activated only about 150 ms after stimulus onset, are not hooked up to perception via pathways fast enough to inform its contents. Commonality holds that, even though perception is amenable to moulding by perceptual learning and environmental cues, this does not amount to its cognitive penetrability. This is for two reasons. First, because perceptual learning can be controlled for by controlling for the available percepts. Second, because we all live in roughly the same world, having the same set of percepts when found in similar spatiotemporal regions and direct our attention similarly. In other words, the kind of theoretical principles integrated in the human perceptual modules are shared by all.
I suggest reasons why isolation is wrong and argue against commonality. Recent research in neuroscience adds to the wave of downgrading the spatial and temporal accuracy of tools called upon by Fodor and Raftopoulos to evince perception's isolation from higher cognitive centres. Simply put, the level of both spatial and temporal accuracy which these authors need to back up their theses of cognitive impenetrability is extraordinary and likely unavailable. Thus, precise calculations of when the effects of cognition on early vision begin to seem underfunded.

The bulk of my counter-argument to the cognitive impenetrability of perception though consists in attacking it from the angle of perceptual pluralism. I cite recent review articles that are illuminating regarding the variety of the perceptual modules output from human to human. Conditions such as metamorphopsias, gnosanopsias, agnosopsias, and being on the autism spectrum strongly support the existence of a variety of early vision within the human species. Even if we live in roughly the same world, and even if we control for percepts, the output of our perception/early vision is not the same for all. The principles transforming the scene into percepts encapsulated in the human perceptual modules vary importantly and, in so varying, undermine the cognitive impenetrability of perception.

Towards the conclusion of the paper, I address a possible counter-argument to my own. It could be supported that the modes of perception I am pointing to above are problematic, less successful, or erroneous, and it is no coincidence that they are often encountered in a clinical setting. In response to this, I provide a number of cases of varying perception that are not of clinical interest. Moreover, even for perceptions that are subsumed under clinically relevant conditions, I cite medical research that testifies to a large number of these perceptions and conditions not (necessarily) accompanied by a disease or pathology. Thus, the argument from success does not work in cancelling out these perceptual variations as less successful or fined-grained attempts at registering the scene, and cognitive penetrability is upheld.

References


Newton asserted in his 1672 draft of a “New Theory about Light and Colors” that his conclusions followed with a “most rigid consequence.” He further stated in his 1675 piece “An Hypothesis Explaining the Properties of Light” that his conclusions on light were “mathematically demonstrable.” Did Newton always think his conclusions were mathematically certain, or did he hold a more moderate position? Alan Shapiro argues that there were two shifts in Newton toward a moderate position, namely probabilism. Kirsten Walsh argues that no such shifts happened. The purpose of this essay is to clarify this debate and to defend a new stance: I argue that Newton shifted toward a moderate position only once and that this position is best understood as aiming for ‘experimental certainty’ in science. My contribution focuses on Newton’s training in the Aristotelian textbook tradition and on the aether as a recurring hypothesis in his thought.

In section 1, I discuss the term probabilism and then provide an outline of the paper. In the literature, mathematical certainty is contrasted with the moderate position of probabilism, yet some differences exist among the commentators. Alan Shapiro believes it is the thesis that “all explanations and theories [are] hypotheses with varying degrees of probability” (1993, 13). Kirsten Walsh suggests that probabilism is the hypothetico-deductive method (2017, 869). More problematic is the fact that Newton scarcely used the term “probabilism” apart from an early, 1678/9 letter to Boyle. I suggest that it is more accurate (by today’s lights) to call Newton a fallibilist, and it is misleading to label him a probabilist. However, for the sake of coming to terms with the commentators, I also use term probabilism.

Probabilism as it is used in this discussion, is an ordinal—not numerical—concept. Specifically, I define it as the belief that one is not justified in thinking that empirical statements can achieve demonstrable certainty and that one’s confidence in these statements are more or less probable depending on the evidence for them. Robert Boyle, Robert Hooke, Joseph Glanvill, and Issac Newton were all probabilists in this sense. Crucially, the former three as well as many others believed that natural science obtained only moral certainty—just one of many positions that fall under probabilism. The general early modern English understanding of ‘moral certainty’ was certainty beyond any reasonable doubt (Boyle 1966, 4: 42-5, 182, van Leeuwan 1970, Shapin and Schaffer 1985, Shapiro, 1983).

The aim of this paper is to clarify the recent literature on Newton’s position concerning certainty in science and to defend a new stance: I argue that Newton shifted toward a moderate position only once (after 1675) and that this position is best understood as aiming for ‘experimental certainty’ in science. So, the two guiding questions are these:

(1) When did Newton become a probabilist?

(2) What kind of probabilist was he and why did he think his scientific conclusions were more than morally certain?
Newton didn’t use the term moral certainty, however he did consistently claim that his conclusions were more certain than those of his contemporaries (Boyle, Hooke, and Huygens), who thought moral certainty was the highest kind of probability achieved in science.

In section 2, I introduce some conflicting passages, motivating these two questions. I then (section 3) show where two main positions go awry: Shapiro thinks Newton’s probabilism is less than mathematically certain, but more than morally certain, something I call ‘experimental certainty’ (1993, p.14, 19-21). I agree with Shapiro’s conclusion, although his account has two problems. First, Newton’s Lectures casts doubts on Shapiro’s reading that Newton moved toward probabilism in 1672. Second, Shapiro’s account suffers from internal inconsistencies and implies that scientific hypotheses are different in kind—not degree—from other scientific propositions. Walsh argues that Newton endorsed a notion of certainty as “compelled assent” (2017, 877). I attempt to differentiate compelled assent from moral certainty. I conclude that the criteria for compelled assent, while very convincing, need a bit more in order for moral certainty and compelled assent to be robustly distinct.

I then offer my own interpretation (section 4), defending a version of probabilism. I focus on Newton’s Aristotelian training and on the aether hypothesis. My answers to the two questions are these:

(1) There is strong textual evidence for thinking Newton shifted toward probabilism only once, which was after 1675/6.

(2) Newton’s adoption of the Aristotelian version of the dual methods of analysis and synthesis, coupled with why he remained committed to the aether hypothesis reveals parts of his method that licensed his belief that his conclusions were experimentally certain.

Newton’s “Hypothesis” was read to the Royal Society in 1675, which still included Newton’s ambiguous bids for demonstrable certainty. So, it was after this time that his position shifted. There was only one such shift in Newton’s thought—there was no second shift in 1710, contra Shapiro (2020). There is a tension between his accepting scientific induction in 1710 (i.e., the second shift) with three other beliefs he accepted before 1710: induction in mathematics, ampliative reasoning, and probabilism around 1675/6. So, given Newton’s earlier commitments, his explicit use of the term ‘induction’ is not significant enough to merit the reading of a second shift.

As for why Newton suggested that his scientific propositions were more than morally certain, I think there are two reasons. First, Newton adopted the Aristotelian version of dual methods because it tracked causal inferences. Second, he thought scientific propositions (including hypotheses) needed strong empirical support, which defenders of moral certainty never hashed out. What distinguished Newton from his contemporaries was that his propositions could have causal import and become deductions from phenomena; therefore, they were causally apt for the Aristotelian dual methods.
The strongest empirical support were things like deductions from phenomena—they had both a quantitative and qualitative relationship to the world. Newton suggested that hypotheses are permitted only as causal conjectures to be answered by experiments (Newton 2004, 165). So, hypotheses are probable propositions with only some empirical support and causal import. The aether had causal import and some empirical support, namely as the possible cause of gravity and the experiments by Hauksbee. Yet, it was conjectural in nature and fell short of being deduced from phenomena. If Newton had done more experiments, this hypothesis would have elevated to a proposition that was deduced from phenomena and generalized by induction, or a quam proxime true proposition that was experimentally certain.

References


Evidence-Based Medicine (EBM) is a medical methodology relying on evidence, where, traditionally, this evidence is a statistical correlation between a disease and its cause established through association studies like clinical or observational trials and confirmed through systematic review and meta-analysis (e.g., Howick, 2011).

However, more recently, evidence of a mechanism linking the disease and its cause has been added to the EBM approach (e.g., on EBM+, Parkkinen, et al., 2018).

How much useful is EBM+ with respect to the new SARS-CoV-2 pandemic? This question has attracted, and is still attracting, a lot of scholars and scientists from many diverse backgrounds. Aronson and colleagues (Aronson, et al., forth.) have recently argued that, when assessing causal claims in the case of an infectious disease like COVID-19, it is insufficient to rely merely on association studies; clinicians must also rely on mechanistic reasoning. In a nutshell, COVID-19 is here used as an example for the fruitfulness of an EBM+ approach, which combines association and mechanistic studies.

The authors’ main line of argumentation is as follows: (i) the reliance on association studies for assessing the effectiveness of causal interventions on the coronavirus disease is insufficient, for “[…] correlation is insufficient for causation: a correlation may be attributable to chance, bias, uncontrolled confounders, inappropriately controlled colliders, or relationships other than causation” (p. 2); (ii) coupling association studies with mechanistic reasoning is beneficial to the research on COVID-19, for, “[…] if an established mechanism links […] two variables, it may be possible to regard […] biases as less plausible explanations of the observed association on COVID-19 outcomes” (p. 4).

The present talk is divided into two parts. After having presented, in Part I, the above EBM+ approach with the coronavirus pandemic as an illustration, Part II raises two different but related general challenges to the EBM+ approach, and proposes some solutions to them.

First, by associating into a single methodological framework, namely EBM+, both association studies and mechanistic ones, it seems that EBM+ ultimately relies on two opposing – rather than complementary or merely compatible - theories of causation. Indeed, as a certain medical methodology, EBM+ is based on certain metaphysical presuppositions about causation: (i) the idea that we come to know a causal relationship in medicine through association studies very likely relies on an agency or counterfactual theory of causation, which analyzes causation in terms of manipulability (e.g. Pearl, 2009); (ii) the idea that we come to know a causal relationship through evidence of an underlying mechanism obviously relies on a metaphysical theory of causation as a mechanism relating a cause to its effect (e.g. Glennan, 1996). Can we coherently reconcile presuppositions (i) and (ii) into a single methodological framework? Or, do we have to go along a strong pluralist way by arguing that presuppositions (i) and (ii) cannot be unified (even minimally)? Indeed, agency or counterfactual theories of causation and mechanistic ones have been historically developed as two competing and (seemingly) irreconcilable theories of causation, where the former focuses on possible worlds, while the latter on the actual world.

A second related challenge on EBM+, in all its generality, is that we may wonder whether mechanistic studies are eventually not sufficient by themselves to assess causal claims in medicine. Indeed, cannot we consider association studies as means to know of a mechanism, that it to say to
get evidence about, actually, an underlying mechanism? In that sense, it is not the case that association studies are supplemented by mechanistic reasoning, but that mechanistic reasoning - under a very strong sense (Aronson et al. (forth.)), thus - is supplemented by association studies establishing through a statistical correlation between two variables, so to say, the boundaries of the mechanism in question. Furthermore, the suggestion that mechanistic reasoning is to be put primary would also solve the first challenge raised here, for EBM(+) would ultimately rely on a single metaphysical theory of causation.

References:


Much of the predictive-processing (PP) framework’s appeal seems to lie in its unificatory power. I take to be the core assumption of this approach its adherence to the claim that cognitive systems minimize precision-weighted prediction-error (i.e., the discrepancy between incoming signals and prediction signals) on average and in the long run. The framework is often said to encompass many, if not all, aspects of cognition, perception, and action.

The problem is that there is no broad consensus about how exactly PP manages to `unify’ these aspects. Traditional criteria of unification such as simplicity and generality seem to apply to some extent, but PP proponents' characterizations are diverse. For example, while Hohwy (2020) and Litwin and Miłkowski (2020) describe the framework as a `toolbox’—a common stock of algorithms that can be used to solve a variety of different cognitive tasks—, Clark (2013, p. 61) describes PP as a sort of ‘bridge’ across all of Marr’s three levels of analysis of complex information-processing systems, including the neural implementation. On the contrary, Prinz (2018, p. 235) considers PP to be aesthetically unifying in the sense that it "offers new ways of seeing" the mind that provides compelling narratives, rather than productive explanations about its neural base.

In this paper, I argue that PP unifies various aspects of cognition by providing a common schema, as opposed to providing a common cognitive process that drives all these different types of cognition. This argument builds on Danks’ (2014, ch. 8) distinction between schema- and process-centered unifications. Roughly, a schema-centered unification is an abstract structure that is shared by a variety of distinct cognitive models that are instantiating that structure. In contrast, a process-centered unification occurs when a system as a whole employs processes or representations that are shared across a wide variety of cognitive activities being modeled. I argue that, while PP does qualify as a schema-centered unification, it fails to qualify as a process-centered unification due to its lack of ascertaining a process common to the relevant cognitive capacities in question.

This novel characterization of PP’s unificatory virtues complements existing ones. It fits best with the toolbox and the narrative interpretations. Similarly to the toolbox interpretation, the same set of algorithms is used over and over again to solve different sets of cognitive tasks. The schema that makes these algorithms so widely applicable is that they all perform, in one sense or another, precision-weighted prediction-error minimization (PPEM). Nevertheless, the cognitive processes instantiating PPEM may do so in a plurality of different ways. Furthermore, the schema common to different PPEM algorithms may be the basis of PP’s narrative appeal, as it functions as a reference point in monitoring PP’s development over the past years. In contrast to Clark’s characterization, PP’s limitation to a schema-centered unification implies that it does not (yet) unify potential answers at Marr’s algorithmic and implementational levels. For instance, how PPEM is carried out remains underdetermined both by the fact that different PP models make different predictions about which algorithm correctly solves a given task, as well as the many ways it could be neurally implemented.

The remainder of the paper addresses two implications for PP’s explanatory status and its relationship to generic Bayesianism. Firstly, it is by now commonly understood in the philosophy of science that there is no necessary relationship between the unificatory power of a theory or model and its explanatory power (Morrison, 2000). However, some proponents of the PP framework seem to implicitly assume a necessary relationship between the two. Yet, if PP is a schema-centered unification, this assumption turns out to be unwarranted, since having a single PPEM schema implies no commitment to its truth or predictive power.
Secondly, if PP provides an improvement towards generic Bayesianism, this is not in terms of PP’s unificatory status but rather something else. While Bayesianism also qualifies as a schema-centered unification, it has been accused of being unfalsifiable and removed from empirical tests. It is commonly held that PP relies on principles of Bayesian inference. So, is PP susceptible to the same problems? Building on Rescorla’s (2015) application of Kuhn’s notion of a paradigm to generic Bayesianism, I suggest reinterpreting PP’s role as part of the normal science of Bayesianism. PP researchers make precise Bayesian rational analyses by offering a plurality of PPEM algorithms, and, in so doing, they facilitate specific applications that may increase the match between Bayesian predictions (e.g., the Bayesian Brain Hypothesis) and specific observations about cognitive systems (e.g., the brain).

References


Pulkkinen, Karoliina

*Could scientific representations help to cultivate epistemic humility?*

It has been suggested that experts (alongside everybody else) are vulnerable to the bias of overconfidence (Angner 2006). In scientific practice, one of the most pressing worries regarding overconfidence is that the provisional nature of scientific knowledge gets underplayed when the experts communicating the relevant knowledge overpromise what it can deliver. How can this be remedied? Recently, Erik Angner (2020) has argued that practicing epistemic humility could be one way to tackle experts’ overconfidence. Epistemic humility is an intellectual virtue that specifically characterises experts’ knowledge and the abilities relevant for their expertise. Some authors emphasise that such humility involves the acknowledgement of the limitations related to one’s knowledge and intellectual abilities (which includes taking responsibility of such limitations), while others prefer to analyse it in terms of setting one’s confidence at an appropriate level.

How can such humility be cultivated? In this paper, I suggest that we have good reasons to believe that scientific representations that possess specific valuable attributes (or values) help to cultivate epistemic humility in those examining and using the representations. I will evoke two examples of scientific representations – one from 19th century chemistry, and the other from modern climate science – to demonstrate how specific values associated with representations help to cultivate epistemic humility. By doing so, I bridge the gap between two vibrant areas of scholarship: accounts on values in science and intellectual virtues of inquirers. Considering that both deal with qualities that are used to characterise and appraise science and experts, surprisingly little has been written about the relationship between the two. Although some authors have explored whether the virtues of experts’ trickle to their work, it seems that there are compelling reasons to refrain from drawing connections between virtuous inquirer and the virtues of their work (Ivanova and French 2020, 5-6).

However, I suggest that instead of looking into how a virtuous expert might construct a scientific representation with valuable qualities, it is more fruitful to examine whether scientific representations that possess a certain set of values help to cultivate virtues of those examining and using the representations.


Erik Angner, “Epistemic Humility – Knowing Your Limits in a Pandemic.” Behavioral Scientist (2020)


Milena Ivanova and Steven French, eds. The Aesthetics of Science: Beauty, Imagination and Understanding. (Routledge, 2020)

Rathkopf, Charles

On the empirical content of deep neural networks

Deep neural networks (DNNs) predict activity in the human brain, despite not being trained to do so. Such unsolicited predictions appear to justify the claim that the DNN is “doing the same thing as the brain.” The fact that the network formalism applies to both biological and artificial systems seems significant. I articulate its significance by asking whether it can be viewed either as an instance of multiple realization, or as an instance of model transfer. I conclude that there is a spectrum of phenomena between these paradigm cases, and that predictive DNNs occupy a middle position on that spectrum.
Inference to the Best Explanation (IBE) recognizes a confirmational role to explanatory considerations. That is, the best explanation of the data is probably true or receives a higher probability. It has been argued that IBE and Bayesian confirmation theory are incompatible by highlighting, for example, that Bayesianism makes no reference to the concept of explanation (e.g. Douven, 2017). The most popular attempt to reconcile the two (e.g. Okasha, 2000; Lipton, 2001) shows that explanatory considerations help to determine the terms in Bayes’ theorem. Namely, IBE is framed in Bayesian terms.

The aim of the talk is twofold. Firstly, I show that framing IBE in Bayesian terms sometimes leads to the Problem of the Old Evidence (POE). According to this, Bayesian confirmation theory cannot explain why an evidence already known confirms a scientific theory. To do that, I point out that the same pattern that gives rise to POE can be found in some examples of IBE (e.g. Lipton, ibid.; Douven, 2017). Here, in fact, an anomalous phenomenon E, known before it is discovered that the theory T best explains it, confirms T. The latter point is due to IBE core idea that explanatory considerations have a confirmational import, and to its “self-evidencing” character (Lipton, ibid.), i.e. an evidence explained by a theory, in turn, confirms the theory precisely because it is explained by the theory.

As for the second aim, I investigate if the solutions proposed to the dynamic and static dimensions of POE work when they are expressed in best explanatory terms. In the dynamic POE in best explanatory terms, we want to explain how discovering that T best explains E raises our degree of belief in T. This is a sensible claim for IBE, given that explanatory considerations have a confirmational import. The standard approaches to the dynamic POE attempt to show that learning that T implies E, i.e. $X = T \vdash E$, confirms T, namely $P(T|X, E) > P(T|E)$. These approaches, however, do not work when the dynamic POE is expressed in best explanatory terms as they do not capture that what raises our degree of belief in T is learning that T best explains E. Moreover, the explanatory relationship between T and E is deductive, but it is problematic to plug a Deductive-Nomological account of explanation in IBE (Lipton, ibid.).

Two solutions proposed by Eva and Hartmann (2020) go in the right direction as they recognize that cases of confirmation by old evidence are instances of IBE. However, an explicit connection to IBE is never made. I claim that such a connection allows to gauge the weaknesses and strengths of the models. The first one aims to show what we are interested about: learning that T is the only available hypothesis that adequately explains the old evidence E raises our degree of belief in T. In fact, they prove $P(T|X \land \neg Y) > P(T)$, where $X \equiv T$ adequately explains E and $Y \equiv T^{*}$’s best competitor ($T^{*}$) adequately explains E. $X$ and $Y$ do not commit to any specific account of explanation capturing the fact that, although we have a semantic grip of an account of explanation to plug into IBE, we still lack a precise understanding of it (Lipton, ibid.). However, the condition HF4*: $P(T|X \land Y) = P(T|\neg X \land \neg Y)$ that carries the burden of the proof, is not an expression of IBE idea that explanatory considerations have a confirmational import. According to HF4*, when both $T$ and $T^{*}$ adequately explain $E$, $T$ receives a negligible confirmation. But, since for IBE, explanatory considerations have a confirmational role, $P(T|X \land Y)$ should be ranked strictly higher than $P(T|\neg X \land \neg Y)$. In fact, $X \land Y$ implies that $T$ adequately explains $E$, whereas $\neg X \land \neg Y$ does not.

The second model aims to show that scientists increase their degree of belief in $T$ because they become more confident of $A$, where $A \equiv T$ is the only available hypothesis that adequately explains $E$. However, they can never learn $A$ for certain as they cannot be sure that $T$ is absolutely the only available hypothesis that adequately explains $E$. The latter point is known in IBE literature as
“argument of the bad lot” (van Fraassen, 1989). To model this, one minimal constraint is used, i.e. A1: P(T|A) > P(T|¬A). Namely, the probability of T is higher supposing A than supposing ¬A. From A1, by using Jeffrey’s conditionalization, it follows that when the scientist becomes more confident in A, T is confirmed: P*(T) > P(T). In IBE terms, the scientist is using a definition of IBE (condition A1) which implicitly relies on the belief that T is absolutely the only available explanation that adequately explains E. However, it is an open question which definition of IBE scientists descriptively use (Douven, 2017).

In the static POE in IBE terms, we want to explain how E confirms the theory T that best explains it more than the other theories that do not explain the evidence as well as T does. This claim makes sense in IBE context, given the confirmational import of explanatory considerations, and IBE self-evidencing character. The standard approach to the static POE evaluates the confirmational relationship between T and E by using a counterfactual degree of belief function where E is not known. This allows a meaningful comparison between the likelihoods and the priors to establish confirmation judgements. Since the aforementioned conciliatory approach proves that better explanations have higher priors and likelihoods, E confirms T more than the competing theories. However, a classic problem of the counterfactual approach applies in this case: we would not have any particular degree of belief in T if E is not known since T would not even be formulated had E not be learned.

References


Neuroscientists are interested in understanding neural populations. This often involves comparing the effect of an intervention on population P1 with a controlled population P2. Since it is the only difference between populations, the intervention is taken to have causal powers (Woodward, 2003; Waters, 2007). Yet, neuroscientists are not interested in merely comparing manipulations or (statistical) differences. They instead want to understand the mechanisms that explain those differences, i.e., the ways in which causes influence their effects.

Broadly speaking, causal influence comes in two varieties. On the one hand, it can be ‘coarse-grained.’ For example, neurosurgeons may lesion a brain area in order to reduce neural hyperactivity, e.g., during epilepsy. Indeed, this was done to patient H.M. when his medial temporal lobes were bilaterally removed. On the other hand, causal influence can be ‘fine-tuned.’ Rather than lesioning hyperactive areas, the medical doctor may, for instance, prescribe medication that promotes the activity of neurons, e.g., GABAergic, that inhibit neural firing. Neural hyperactivity, then, can be reduced in both ways, yet it makes a great deal of difference whether this occurs in a coarse-grained or fine-tuned manner.

Philosophers of science have proposed ‘causal specificity’ as a way of understanding relative causal control (Weber, 2017a). According to this view, the causal set with the highest number of functional mappings onto the effect is most fine-tuned. What does this mean? Let A:{x, y, z} denote the set of sub-causes that influence population P. Here, A is the cause and x, y, and z are sub-causes. Let P:{f, g, h} denote the set of sub-effects influenced by A. Here, P is the effect and f, g, and h are sub-effects. Let us understand this via an example. The cause may be an anti-convulsant (A), which acts via psychoactive proteins (x, y, z) that bind to post-synaptic receptors (f, g, h), thereby reducing neural hyperactivity (P). So, according to causal specificity, the degree to which the sub-causes of A -- the values of the domain -- map onto those of P -- the values of the co-domain -- constitutes the degree of causal control A exercises over P (relative to P's other causes). Indeed, the A-to-P mapping must be functional: (1) All values of the domain must map onto at least one value of the co-domain, and (2) each value of the domain must map to no more than one value of the co-domain. The proponents of this view advance a numerical interpretation of ‘functional mapping,’ according to which “the number of values that the variables on both sides of the relation can take is vastly higher ... than that of any other causal variables that bear the [same] relation” (Weber, 2017b, p. 32).

My presentation will be a systematic critique of this view. First, the framework fails to capture the phenomena of dual effects, wherein a sub-cause influences two sub-effects. For instance, the protein x in anti-convulsant A may bind with receptors f and g in P. Once this is represented using functional mapping, x maps onto both f and g. Condition (2) above requires that each domain value must map to no more than one co-domain value. Accordingly, this mapping violates a central condition, and so the requirement that the mapping be functional is overly restrictive. Second, the framework fails to capture causal redundancy, wherein a sub-cause fails to map onto any sub-effect in P. Hypothetically, z may map onto a sub-effect in a non-P population, yet when we only consider the A-to-P mapping, z remains unmapped. This violates condition (1) of functional mapping, namely that all domain values must be mapped. Thus, the framework fails to capture causal redundancy, a phenomena that is ubiquitous in neural populations.

Third, the advocates of this view claim that causes which bijectively map onto their effects exercise the highest degree of causal control. Let me provide some background on functional properties before explaining what is meant by -- and is problematic with -- bijective functions. Surjective
functions are as those in which all values of the co-domain are mapped. Injective functions are those in which each mapped co-domain value is mapped onto by at most one domain value. Therefore, a functional mapping is bijective if and only if it is surjective and injective; this mapping exhibits one-to-one relations.

The supporters of causal specificity claim that bijective functions are “most specific.” I want to suggest that this is not the case given the numerical interpretation above. Instead, surjective non-injective functions can be more specific. Consider a cause F:{1, 2} and its effect Z:{a, b}. Suppose the F-to-Z mapping is bijective. Consider another cause G:{1, 2, 3} and its effect Z:{a, b}. Suppose the G-to-Z mapping is surjective non-injective. Let G(x)=F(x) but let G(3)=G(1)=F(1), meaning that G-values ‘1’ and ‘3’ and F-value ‘1’ map to Z-value ‘a.’ Calculating causal specificities using the numerical approach yields values 2 and 3 for the former and latter mappings, respectively. That is, two values of F and three of G map to Z, meaning that G-mapping (surjective non-injective) is more specific than F-mapping (bijective). Thus, the framework is incoherent when the numerical interpretation is advanced alongside the assertion that bijective functions are most specific.

Finally, I want to suggest that the framework includes incommensurability: it cannot compare the mappings of the anti-convulsant A onto P with those of lesion B onto P. This is because B does not have any identifiable sub-causes. Instead, it appears to act on the entirety of P. By contrast, A has sub-causes x, y, and z that specifically map onto sub-effects f, g, and h. So, this is an identifiable mapping. It follows, then, that we cannot even compare A-to-P with B-to-P, for the latter does not fall within the purview of the framework. Accordingly, it seems that the very problem that motivated the discussion, namely whether a schematic can be developed to capture relative causal control, cannot be accommodated. As a result, causal specificity as a view of causal control is misguided.

References


Stanley, Shaun

Cultural Evolutionary Theory and the Difficulties of Inferring Microevolutionary processes from Macroevolutionary Patterns

In the preamble to their discussion on point typologies, Bettinger and Eerkens tell us that “it can be argued that typologies themselves are of no intrinsic interest; they are merely intermediary constructions useful in investigating the “real” behaviors and processes we want to study” (Bettinger and Eerkens 1999, 231). In other words, a theoretical aim in archaeology is that inferences from the contents of the archaeological record to the processes of behavioral interactions in historical human societies be both possible and reliable. Recently Garvey (2018) has tried to make good on this sort of inference within the Darwinian framework of Cultural Evolutionary Theory (CET) and provides some arguments for the further inclusion of archaeology within CET. I present a two-fold argument. Firstly, I argue that these inferences are subject to two epistemic difficulties. Secondly, as against the remarks of Mesoudi et al (2006) and other enthusiasts for the view that the social sciences might be unified within CET, that these difficulties call into question the unity of CET itself and suggest a more disenchanted view regarding the Darwinization of social science.

Regarding the epistemic difficulties, the first difficulty I call (following Garvey) the Resolution Problem. The problem is that there is an epistemic gap between the long-term, large-scale patterns of the archaeological record and the short-term, small-scale, processes of biased transmission involved in cultural microevolution, and this epistemic gap is widened by preservation biases (factors affecting the likelihood that the artifacts and their spatial relationships have been preserved through time). The Resolution Problem, when it cannot be solved, weakens the warrant of such inferences. The second difficulty I call the Underdetermination Problem. The problem is that artifacts can come to be, and can be used to exhibit their function, in many different ways such that each of many distinct forms of behavior could, in principle, account for the artifacts form or function. One overcomes this form of underdetermination by amassing assumptions about standard forms of behavior in the relevant context given, say, by independent contextual data (for example that “…the objects that accompany people in death tend to be those that accompanied them in life (Haas et al 2020, 5)” allowing one to make inferences from burial remains to historical forms of behavior). However, insofar as these independent data are themselves called into question, so too are the inferences made about which behavioral processes were linked to the artifacts in question. So, when the Resolution and Underdetermination Problems cannot be adequately solved (which may be quite often) we must accept diminished warrant for inferences made about behavioral practices in historical human societies. Given preservation biases, sadly, we must also expect a limited range of material from which to make such inferences. Jointly, I believe this poses a threat to the epistemic unity of CET, and given all these limitations, contra the glowing expectations suggested to us by Mesoudi and others, I think this suggests a more disenchanted view as to the CET’s unifying power for the social sciences.

References:


This talk has three goals. First, I’ll argue that the popular distinction between explanatory, objectual and pragmatic understanding is currently ambiguous concerning whether it refers to a distinction between states, modes or objects of understanding. I argue that this distinction is most fruitfully characterized as one between epistemic states, rather than modes or objects. Second, I address the question of fundamentality. What do we mean when we say that one of these kinds of understanding is more fundamental than the others? I distinguish between several possibilities, and focus on “humble reductionism” – where one kind of understanding is necessary for another kind of understanding. Again, taking the three main kinds of understanding to refer to three different kinds of epistemic states, I’ll argue that pragmatic understanding is the most fundamental kind of understanding (in the above-mentioned sense). Third, because pragmatic understanding is the least discussed of the three main kinds of understanding, I’ll sketch a preliminary account of it, drawing what lessons I can from the existing literature and new ethnographic research performed on working scientists across several fields. Roughly, I claim that an agent S has pragmatic understanding of X when S has the ability to reliably and successfully manipulate X, where those abilities come with the right dispositions; exist in the majority of nearby possible worlds; and are robust. The more robust S’s ability, and the more relevant ways S can successfully manipulate X, and the more successful those manipulations are, the more pragmatic understanding S has of X. I close by considering how this account informs the debate about scientific progress. Specifically, it allows us to capture what seems right about both the so-called noetic (understanding-based) and functional (ability-based) accounts of scientific progress, combining them into a single account of scientific progress. I take this to be a point in favour of the characterization I give of pragmatic understanding.

References


Suárez, Mauricio

The Complex Nexus of Evolutionary Fitness

In evolutionary biology, fitness has long been appreciated by many to be a probabilistic disposition, or propensity, to reproduce successfully (see Brandon, 1978; Mills and Beatty, 1979). This propensity interpretation of fitness (PIF) is part of a larger tradition in evolutionary thinking that takes fitness or adaptiveness to be a causally explanatory concept (Sober, 1984, 2001). Yet, there has been little consensus as to the specific kind of propensity fitness is. On the contrary, there is much disagreement in the field as to how to formally represent fitness, how exactly it is an explanatory concept, and what exactly it explains. Critics have been quick to latch onto such disagreements in order to argue that fitness is not causally explanatory after all (Walsh, Ariew, Mahen, 2016), that it does not reflect causal relations (Walsh, 2010), and that there are no propensities underlying adaptation phenomena in evolutionary biology (Ariew and Ernst, 2009). The current impasse suggests that there are fundamental issues at stake regarding the nature of propensity and its explanatory power that are yet to be clarified. In this talk I offer a more complex and nuanced framework than is typically assumed for modelling chancy phenomena in general, the ‘complex nexus of chance’ (CNC). Contrary to what has been conventional in the philosophy of probability, this approach distinguishes clearly propensities from both probabilities and the finite frequency data that are used to test them. I argue that the application of this framework to fitness bears significantly on a few important problems currently discussed in relation with (PIF).

The first concerns the exact formal or mathematical representation of fitness as propensity. The relevant discussion here broaches two technical aspects of statistical modelling, informing what are sometimes known as the moments problem and the delayed selection problem (Sober, 1984; Beatty and Finsen, 1989; Pence and Ramsey, 2013). On the one hand there is the demonstrable empirical fact that fitness is often sensitive to higher moments of the statistical distribution for reproductive success. Hence identifying fitness with just the statistical mean average (the expected value, or expectation) of a probability distribution will often miss out critical differences down the lineage. The differences can be so critical as to entirely reverse judgements of relative fitness between individual organisms (or traits, or genes – more about this later on). But the idea that fitness, understood as a propensity, must necessarily be identified with some or other moment of a probability distribution presupposes that all propensities are statistical functions, or formal moments of the distributions, which is nowadays questionable in the philosophy of probability, and it is indeed rejected by the CNC. I consequently suggest that CNC accounts for the statistical modelling of fitness without such assumptions, and thus delivers us from the problem of moments.

The second narrow technical issue concerns whether fitness is short or long term; i.e. whether it involves reproductive success in the most proximate generations, or perhaps even just the next generation; or whether, by contrast, fitness refers meaningfully only to reproductive success down the generations – or perhaps even hypothetical success in some infinite reproductive limit. On a propensity analysis, the issue may at first sight seem merely a version of the debate regarding ‘single case’ versus ‘long run’ propensity interpretations of probability. If so, the delayed selection problem would boil down merely to a difference regarding the appropriate type of propensity involved, where those advocating long term fitness would be implicitly if not explicitly adopting a ‘long run’ propensity account. However, I argue that these distinctions are in fact tangential. Long term fitnesses, in particular, are perfectly compatible with ‘single case’ propensities, as advocated by the CNC. This has consequences for the precise mathematical definitions that are appropriate when modelling fitness in different contexts, and whether or not they issue in contradictions.
Then there is the second and more general issue, namely the explanatory role of fitness. Advocates of the PIF typically defend the view that fitness is a causally explanatory property of biological entities – and for this reason they are sometimes known as ‘causalists’ (Abrams, 2012). Critics of the PIF by contrast, tend to view fitness as not particularly an explanatory concept – certainly not a causally explanatory one –, but rather a descriptive or generalising concept. In the last part of the talk, I argue that CNF shows both ‘causalists’ and ‘statisticalists’ to be in part right. Propensities are indeed explanatory entities, but in accordance to CNC they typically explain not frequencies in data, but the single case chances that they give rise to within particular chance set ups. Thus ‘fitness’ is indeed often a name for an explanatory propensity, but not merely that: it is also a name used for the distinct probability distribution within a statistical model that is adequate for the purpose of representing the single case chances manifested. And in turn these chances are used to account for the actual data recorded in observational studies of reproductive success, where ‘fitness’ is also sometimes confusingly used to refer to the finite frequencies in the data for reproductive success. The disambiguation of these three distinct but mutually related uses of fitness is essential for a better understanding of its explanatory power.

Selected References


Sober, E. (2013), “Trait Fitness is not a Propensity, but Fitness Variation is”, Studies in History and Philosophy of Biological and Biomedical Sciences, 44, pp. 336-341.

Behavioural economics has taught us that human agents don't always display consistent, stable, and context-independent preferences in their choice behaviours. Can we nevertheless do welfare economics in a way that lives up to the anti-paternalist ideal most economists subscribe to? Against recent pessimism, this paper argues that, on the right view of what constitutes subjective interest, the choice-theoretic framework developed by Douglas Bernheim and Antonio Rangel (2007, 2009) lives up to this challenge.
Biologists and psychologists have argued social learning is important for the intergenerational transmission of cultural information (Boyd & Richerson, 2005; Henrich, 2015), including technological knowledge as well as social norms and traditions. Some have argued further that cumulative cultural accumulation, roughly understood as cultural information that could never be re-innovated by an individual (Mesoudi & Thornton, 2018; Miton & Charbonneau, 2018), is a uniquely human trait as a consequence of specific forms of social learning we alone rely on. The homo imitans hypothesis (henceforth HIH) suggests the copying of conspecifics’ behavior, or imitation, is the uniquely human social learning process that accounts for cumulative cultural accumulation (Call, 2009; Tennie et al., 2009; Tomasello et al., 1993).

HIH would be disconfirmed if nonhuman animals imitate. Variability in intraspecific traits that are not the result of ecological or genetic differences has led some comparative researchers to argue that other species have cultures in the relevant sense (Aplin, 2019; Danchin et al., 2018; Garland et al., 2017; Jesmer et al., 2018; Schuppli & Schaik, 2019; Whiten et al., 1999). Furthermore some nonhuman species arguably imitate others in the process of learning communicative signals and causally opaque behaviors (Klein & Zentall, 2003; Kuczaj & Yeater, 2006; Whiten et al., 1996, 2009; Zhao et al., 2019). Taken together, one might speculate that nonhuman animals sometimes imitate in order propagate and maintain culture-specific information.

Others have criticized this line of reasoning (Claidière & Sperber, 2010; Galef, 1988; Schlingloff & Moore, 2018), but the ‘Zone of Latent Solutions’ hypothesis (henceforth ZLSH) stands as the most developed critique of the idea that nonhuman animal cultures are the result of imitation (Bandini et al., 2020; Reindl et al., 2018; Tennie et al., 2020; Tennie & Hedwig, 2009). ZLSH suggests that apparent cases of imitative learning in animals are actually the consequences of socially-mediated individual learning: instead of copying the behavior of a conspecific, the presence of a conspecific aids in an animal performing similar behaviors in similar contexts. Tennie and colleagues (Bandini & Tennie, 2020; Neadle et al., 2020) have argued that since naïve individuals can arrive at similar solutions to complex problems without demonstration, imitation is not necessary for great apes to maintain the relevant traits. Instead, individuals simply re-innovate complex behaviors, which puts pressure on the idea that imitation is the cause of intraspecific group differences. If true, ZLSH would support HIH since no other animal imitates, its reasonable to suggest that imitation is what makes human culture unique.

While individual and social learning are clearly contrastive in terms of the environmental circumstances in which they occur, the psychological mechanisms that afford both individual and social learning likely overlap (Heyes, 2012; Heyes & Pearce, 2015). This fact, I argue, makes drawing conclusions about how social or solitary experiences can lead to changes in behavior more complicated that ZLSH portends. This paper will develop two objections to ZLSH that center on this confusion about the relationship between individual and social learning.

Firstly, all socially learned behaviors permit of (re-)innovation by a solitary individual in principle. When researchers claim that some trait or skill could never be re-innovated by an individual, they often underestimate solitary learning processes (assuming, e.g. individuals are limited to trial-and-error learning) and underspecify the scope of the counterfactual. In practice, those social learning techniques that most likely permit of naïve innovations are also responsible for stereotypically culturally mediated traits in humans (Caldwell & Millen, 2009). Therefore, imitation is not necessary
for trait acquisition, though its likely sufficient (Legare & Nielsen, 2015). Since the same psychological mechanisms underwrite individual and social learning, ZLSH does not offer a productive competing hypothesis in light of the evidence that nonhuman animals imitate.

Secondly, different human cultures utilize divergent social learning techniques (Legare, 2017; Lew-Levy et al., 2019; Mesoudi et al., 2016), most of which ZLSH could be readily applied to show the strategies are not necessary for the acquisition of traits. In fact, those skills and norms that do not practically permit re-innovation are rare and late developing in human history (Coolidge & Wynn, 2016; Stout, 2011). If ZLSH holds for early and modern human social groups, then it cannot secure HIH. But the variability in these human social groups are what cultural evolutionists are trying to explain, regardless whether the explanation relies on claims about human uniqueness. ZLSH relies on a problematic contrast between individual and social learning, and as a result it fails to explain the roots of human and animal culture.
Pure Shape Dynamics (PSD) is a new framework for constructing relational theories of motion, both classical and quantum. It differs from standard Shape Dynamics (SD) in that it supplies a description of the evolution of a physical system solely in terms of the geometric properties of an *unparametrized* curve defined over the relational configuration space of the theory (called shape space), without invoking either best-matching, Jacobi's principle, or monotonically increasing parametrizations of the dynamical curves over shape space (see Mercati, 2018, for a comprehensive survey of standard SD).

In this talk, we will first provide a self-contained presentation of the technical machinery of PSD and its physical significance. We will then consider the metaphysical implications that the adoption of unparametrized dynamical curves have on the notions of time and change. It is usually claimed that SD nicely accommodates a notion of B-series-type change, i.e., the physical evolution being given by a sequence of configurations ordered by an "earlier than" relation (cf. Gryb and Thébault, 2016, for a full articulation of this argument). PSD, on the other hand, seems to strip this ordering to its topological bare bones, so to speak: The dynamics is still given by a sequence of configurations ordered by a "nearness" relation but, without any monotonically increasing parametrization at hand, there seems to be no more room in the formalism to distinguish whether a configuration is "earlier than" or "subsequent to" another one. The discussion will show that this is in fact not the case: To the contrary, PSD is capable of providing an account of time and change centered around the notion of intrinsic complexity of a system, which is much more faithful to relationalism than the monotonically increasing ordered structure of SD is.

Bibliography:


In this paper I will criticize current concepts of the immune system (IS) for focusing entirely on mammals. I will continue the groundwork laid by early feminist scholarship regarding bias in immunology to provide an alternative position resisting taxonomic chauvinism. Ideally, a framework that resists biases is also a framework that describes the processes studied more accurately. Examining an invertebrate IS, I will show that: 1) there are invertebrate adaptive IS, based on nucleic acids and thus transgenerational; 2) the distinction between biotic and abiotic stressors breaks down. Reworking our concept of the IS to address the vast heterogeneity of multiple realized IS will be helpful both to the philosophical debate whether the self/nonself-, danger-, or continuity theory best predicts the outcomes of IS interactions, as well as epistemically helpful to scientists, who study organisms that do not match the mammalian prerogatives of IS.
I. Single models alone?

Much has been learnt about scientific models since philosophers turned their attention to them. While there is still disagreement about their epistemic import, their ontology, or the representational relation in which they stand with their targets, these issues are nowadays debated precisely because models do no longer stand in the shadow of theories. The significance of their role in scientific practice has been acknowledged; it is not up for debate. This widespread understanding of their role has been possible, at least partly, due to the capacity of philosophers across the sciences to communicate with each other. Issues such as the requirements of a model-based theory of scientific representation (Frigg & Nguyen, 2017), the role of idealisations in modelling (Cartwright & Jones, 2005), or whether models are explanatory (Reiss, 2012) have been answered with considerable level of generality by dint of the analysis of relatively simple models that are easily explained and understood by non-experts. In other words, philosophers of economics and biology alike have been able to learn from simple models like Schelling’s model of segregation (Ylikoski & Aydinonat, 2014).

In this paper we shall argue, however, that analysis of individual models is not enough. Models are not islands; they don’t stand on their own. Thorough philosophical understanding of models and, particularly, of their epistemic import requires analysis of families of models and of the relations between individual models. We shall further suggest that the extant philosophical literature is not well equipped to embark in the analysis of families of models.

II. The CAPM: a family of models.

In order to make our point, we shall rely on the Capital Asset Pricing Model (CAPM; pronounced “cap-em”); the cornerstone of modern portfolio theory, which granted Harry Markowitz and William Sharpe the Nobel Prize in Economics in 1990. This is a family of models with a number of features that make it particularly susceptible to philosophical scrutiny. The CAPM is a) popular and important in practice. Besides having been accredited with the Nobel stamp, it is used by both academics and practitioners. But, b) its empirical support is contested; there is evidence in its favour and evidence against it. More damning, perhaps, is that c) there are principled concerns over its testability.

What’s the use of a family of models that is, on the one hand, at worst untestable and at best contested and, on the other hand, favoured by academics and practitioners? This is a pressing philosophical question. We interpret it as being confronted with a diverse family of models whose scientific aims are, at best, unclear. How to address it with accounts of models that have largely focussed on the individual model? We suggest that we should take a step back and do descriptive work first: how are the different models in the CAPM related to one another? In how far do they cohere and differ in the questions they ask and the answers they give? What types of contributions might we ascribe to different single models or clusters within the CAPM family?
III. A descriptive framework for model diversity.

In order to address these questions, we put forward a descriptive account of the CAPM family in a way in which we can be faithful to its inherent diversity with unclear, differing and complex scientific aims. Our framework is simple, and focuses on the relations between models within a diverse family of models. We ask: in how far are the models complementary to each other? We define two relations, vertical and horizontal complementarity. With vertical complementarity we refer to the fact that a model contributes new aspects to the development of the common theoretical framework. By horizontal complementarity we mean that a model reveals new aspects of the phenomenon in question. We analyse three cases of complementarity in the CAPM family—one with a predominantly theory-focused complementarity (vertical), one with a predominantly phenomena-related complementarity (horizontal), and one with both.

We draw three main lessons from this analysis. First, with respect to the epistemic import of individual models, the variety of relations reveals that not every cluster within a family of models contributes in the same way. Indeed, contra (Aydinonat, 2018) we argue that it is far from clear how and whether we may systematically connect positive epistemic import with bringing about additional complementary models. There are cases within the CAPM family which seem to bring about befuddlement rather than better explanations. Second, there might be instances in which, if judged by the understanding a model offers about a phenomenon, its epistemic import may be considered to amount to little, or nothing. And yet, the impact it has in the discipline can be considerable, because of the theoretical (vertical complementarity) insight it provides. To be able to recognise such contributions philosophically is valuable because it sheds light on the specific research area. Finally, the CAPM is a case with important general import, as it simply combines two known features of scientific inquiry: (a) that there are families of models, and (b) that it can be sometimes hard to pinpoint what the exact aim and goal of their scientific aim is. The CAPM is simply a hard case of (a)-(b) entanglement. It is thus not only interesting in and of itself—we will also learn something for any other part of science in which there is such diversity.

References


Vos, Bobby

Integrated HPS? Formal versus Historical Approaches to Philosophy of Science

*Main thesis*

The separation between formal and historical approaches is ill-founded. Formal approaches not only can—but *have*—incorporated nuanced treatments of the historical dimension of science into their frameworks. Conversely, historical approaches can benefit from the conceptual rigour offered by formal analyses, although this possibility remains unactualized in practice. In particular, some of the major criticisms faced by the various historicist frameworks of the 1960s and '70s, most notably those of Kuhn and Laudan, could have in part been avoided by closer engagement with the developments in formal philosophy of science at the time. The charge that formal approaches have ‘nothing to offer’ to historical analyses is therefore misguided. Moreover, the fact that formal approaches continue to play no role of significance in contemporary reflections on ‘Integrated HPS’ is primarily due to sociological, rather than conceptual reasons.

*Outline*

While recent decades have seen a thorough-going synthesis of history and philosophy of science—most notably in the form of the ‘Integrated History and Philosophy of Science’ movement—the gap between history and formal philosophy of science remains as wide as ever. In this talk, I argue that the divide between historical and formal philosophies of science is ill-founded.

I start by considering the origin of the supposed dichotomy between history and formalism in philosophy of science, which dates back to the emergence of post-positivist, historical philosophy of science. A survey of the philosophical literature of the time shows a pervasive presupposition, especially among formal philosophers, that there is an intrinsic opposition between historical and formal philosophy of science—the ‘H/F dichotomy’.

Having established the perception of a dichotomy, I argue it is ill-founded. To this end, I first argue that it is not clear what the H/F dichotomy itself is supposed to consist in. After considering several possible construals of ‘historical philosophy of science’ and settling on the most plausible construal in this context, I argue that the resulting version of the H/F dichotomy is straightforwardly refuted by the history of the philosophy of science. More specifically, I shall draw attention to an often-overlooked, yet extensively developed research tradition in the philosophy of science, viz. formal accounts of theory-dynamics as developed within the structuralist approach to philosophy of science. Damböck (2014) has recently drawn attention to the interaction between Kuhn and the Munich structuralists in particular, as exemplified by Stegmüller (1976) and Balzer, Moulines and Sneed (1987, ch. 5). I show, however, that the interaction between historical and formal philosophies extends beyond Kuhn and the Munich structuralist and includes also the attempts by Pearce and Rantala (Pearce & Rantala 1983, Pearce 1987) to formalize certain aspects of Laudan’s problem-solving model of scientific progress. This establishes, I argue, the long-standing interest from formal philosophers in engaging with the historical dimension of science.

Following this, I consider—and set aside—the charge that these formal accounts have ‘nothing to offer’ to the historical philosopher of science. In response to this charge, I show that some of the
major criticisms faced by the various historicist frameworks of the 1960s and ‘70s, most notably those of Kuhn and Laudan, could have in part been avoided by closer engagement with the aforementioned formal frameworks. I conclude, therefore, that the charge that these formal frameworks have ‘nothing to offer’ is misguided, and subsequently follow Damböck (2014) in concluding that the H/F dichotomy in the post-positivist philosophy of science is grounded in sociological, rather than conceptual, factors.

Finally, I extrapolate these observation to modern reflections on ‘Integrated HPS’. More specially, I argue that, just as with post-positivist philosophy of science, the lack of engagement with formal methodology can be attributed to primarily sociological factors. This sociological divide, I argue, is rooted in the contingent, sociological phenomenon that most contemporary interest in Integrated HPS comes from philosophers who are concurrently operating within the context of the practice turn. While innocuous in principle, I argue that this sociological phenomenon can nevertheless provide a plausible explanation for the continued H/F dichotomy.

References


The notion of similarity has been discussed at length in the context of scientific representation. Using qualitative methods such as ethnographic fieldwork and interviews and drawing on a case study concerning the use of mouse models in cancer research, this paper aims to revisit the debate on similarity by investigating different epistemic roles that similarity plays in scientific practice. To this end, the paper first provides some scientific detail and then, for the purposes of philosophical analysis, it proceeds to disentangle three notions or research modes that often appear intertwined in practice but are conceptually distinct: model selection, model extrapolation and model creation. Importantly, the concept of similarity plays different epistemic roles in each of these research modes. Finally, the analysis provides grounds for more general conclusions concerning the concept of scientific representation.

Much of our current knowledge of cancer biology and cancer therapies owe to research conducted on laboratory mouse models which allow for in vivo exploratory as well as hypothesis-driven experimenting. Immunocompetent transplantable models are readily available and are most commonly injected with cancer cell lines to induce tumorigenesis. They allow for studying the tumor-immune system interactions. However, some of their advantages also reveal substantial limitations having mostly to do with the nature of cancer cell lines. Human tumors can be studied in immunodeficient transplantable models but because they do not possess a fully functional immune system, their usefulness in studying tumor-immune interactions and testing immunotherapies is severely restricted. Genetically engineered models rely on the transgenic expression of oncogenes or the inactivation of tumor suppressor genes and provide more physiologically relevant tumor-microenvironment. However, there are confounding factors arising from the contemporary techniques of genetic manipulation (Zitvogel et al. 2016). To overcome the translational gap, many researchers see a promise in further refining humanized mouse models of cancer which can be defined as immunodeficient mice engrafted with haematopoietic cells or tissues, or mice that transgenically express human genes (Shultz et al. 2007). There is a variety of ways in mice can be humanized, each with its own set of problems.

It should quickly become clear that no single model is optimal for addressing all the numerous questions that might be considered (Shultz et al. 2012). Instead, by capturing only some of the relevant aspects of the phenomenon, each type of model proves useful for answering some questions but not others.

Published scientific work often contains passages in which scientists explicitly allude to some kind of similarity between the model and its target (see, e.g. Morton et al. 2016). This is further confirmed by interviewing researchers. However, closer investigation shows that in different contexts the reference to similarities play different epistemic roles. To show this in detail, the paper distinguishes and discusses three research modes: model selection, model extrapolation and model creation.

When selecting a mouse model of cancer, researchers have to first establish what the research question is (see also Huber and Keuck 2013 for a similar point), on the basis of which they may choose between available models. Although the selection is guided by similarity consideration, i.e., what the relevant similarities (and dissimilarities) between a model and its target are given the question at hand, it is balanced by many other factors (see Dietrich et al. 2020 for an extensive overview). For instance, immunocompetent transplantable models permit experiments to be done in a timely manner, cancer cell lines are easily modifiable, and studies of immune cell infiltrates and rapid screening of new drugs can be performed. However, these models do not recapitulate the
tumor microenvironment or the multistep processes of tumorigenesis characteristic of spontaneous
tumor development.

Model extrapolation has received considerable attention in philosophy. While philosophers appear
to be in agreement with respect to the general claim that the success of extrapolation depends on
some sort of similarity between the surrogate and its target, they disagree on the particular ways in
which the justification proceeds (see, e.g. Parkkinen and Williamson 2020). Estimates differ, but it is
generally established that using findings obtained from pre-clinical studies to project efficacy and
safety of potential anticancer drugs in human trials suffers from high rates of failure. More
importantly, there are well documented cases in which experiments conducted on animal models
showed promise but had tragic consequences in first-in-human trials (Lemoine 2017; see also
Parkkinen and Williamson 2020). Humanized mice are said to help close the translational gap (see
also Piotrowska 2013).

Model creation concerns an active process of introducing targeted changes that give rise to new
animal variants. All mouse models currently in use have been modified to some extent, some more
than others. Contemporary cutting-edge research concerns creating humanized mice in hope to
overcome constrains posed by the mismatch between mouse and human physiology.

Similarity considerations play different epistemic roles in each of the research modes. There is a
difference between the initial act of selecting a model for the purposes of studying it, and the
consequent justification of the extrapolative claims based on obtained results. In other words, the
fact that a model is chosen for its purported similarity to its target does not by itself justify the
extrapolation. While in model selection and model extrapolation the considerations concern already
existing, pre-established similarities, the process of creating new mouse models differs in that
scientists have to deliberate which similarities to introduce into the model.

Appreciating the complexities involved in the process of creating new models such as humanized
mouse models of cancer, also proves illuminating for the debate on the similarity account of
scientific representation. According to one objection, the similarity account fails to appreciate the
distinction between representation and accurate representation: similarity is inadequate to account
for the former but can be a criterion of the latter (Frigg and Nguyen 2017, p. 54; see, e.g., Suárez
2010, p. 93). However, the deliberate introduction of similarities exhibited by humanized mice
(model-creation-mode) guides the representational use and influences the subsequent maintenance
of the representational relation. Thus, under these specific conditions, similarity considerations
ground the establishment of representational relation.

research organism. Studies in History and Philosophy of Science Part C: Studies in History and
Philosophy of Biological and Biomedical Sciences, 80, 101227. doi:10.1016/J.SHPSC.2019.101227

Springer Handbook of Model-Based Science (pp. 49–102). Dordrecht: Springer. doi:10.1007/978-3-
319-30526-4_3

Huber, L., & Keuck, L. K. (2013). Mutant mice: Experimental organisms as materialised models in
biomedicine. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy
of Biological and Biomedical Sciences, 44(3), 385–391. doi:10.1016/J.SHPSC.2013.03.001


The Aesthetics of Scientific Experiments

Milena Ivanova, Alice Murphy, Sophie Ritson, Helen De Cruz

There is growing interest in the role of aesthetics in science. However, the focus has primarily been on the “beauty” of scientific theories, and other aspects of scientific practice have been overlooked. The aim of our symposium is to expand the discussion of aesthetics beyond theories to consider another, often-neglected feature of science: the experiment. We will offer a novel exploration of the aesthetic considerations that go into the design of an experiment, and investigate how the reception of an experiment is impacted by its aesthetic value. In particular, we will discuss the following: Is the beauty of an experiment reducible to other properties such as simplicity or elegance? Are experiments across different time periods and across the different sciences valued for the same aesthetic properties, or are these subject to change? What is the relation between beauty and the attainment of understanding via an experiment? And finally, what is the significance of feelings of wonder and awe in the scientific domain? In addressing these, we also hope to demonstrate how attending to the aesthetics of experiments can indicate further fruitful avenues of research concerning the relations between philosophy of science and philosophy of art.

The relationship between aesthetics and science is beginning to attract substantial interest in the philosophy of science. In addition to a number of collections exploring connections between philosophy of science and philosophy of art, (Bueno et. al 2018; Frigg and Hunter 2010) and fiction and imagination in science (Levy and Godfrey-Smith 2020), there is now a volume dedicated specifically to aesthetic evaluations in science (Ivanova and French 2020). The latter offers a number of new directions on how beauty effects the evaluation of scientific theories and can guide theory choice in cases of underdetermination, the relationship between beauty on one hand, and truth and/or the acquisition of understanding on the other, and experiences of the sublime in science.

This has been an invaluable contribution to the debate on how aesthetics can feature in science. However, its primary focus is on scientific theories. The aim of this symposium is to expand the discussion of aesthetics in science beyond theories to consider another, often neglected feature of scientific practice: the experiment. We will offer a novel and comprehensive exploration of the aesthetic considerations that go into the design of a scientific experiment, and investigate how the reception of an experiment is impacted by its aesthetic value, or lack thereof. Our main aims are the following:

The first aim is to analyse the concept of beauty as employed by practicing experimenters. What properties do these experimenters value? Is beauty reducible to other properties such as simplicity or elegance? Are experiments across the sciences and across different time periods valued for the same aesthetic properties or are these subject to change? The practice of constructing an experiment has changed. It once involved one scientist, a small room, and a few tools. Today, thousands of scientists are involved in the construction of an experiment, and its location can surpass the boundaries of a country. Further, the information that an experiment produces can often surpass what can be published in a single monograph. If the practice of experiments can change so drastically, then should we assume the ways that experiments are evaluated aesthetically can also change?
The second aim is to address the role that aesthetic factors play in the construction and reception of an experiment. Can the beauty of an experiment affect how successful it is considered to be, and whether it should be pursued in the first place? Some argue that an experiment is beautiful only when it shows something new and expands our knowledge of the world, for instance, by making a new discovery. But this conception seems limiting, given that experiments are performed with many different aims in mind: to test theories, to study phenomena, to replicate and corroborate results. Is there, then, a more fundamental goal that seems to be accomplished by a beautiful experiment? Furthermore, the aesthetic language used in other domains, such as in the appreciation of art or of nature, is wide ranging, extending far beyond the application of “beauty”. Given this, perhaps it is too restrictive to focus solely on beauty when analysing aesthetic judgements in science. Is there, for example, an important function for “ugly” experiments?

The third aim is to explore how the discussion of the aesthetics of experiments offers a productive way to consider connections between (philosophy of) science and art. The history of art tells us of the common origins of artistic performance and experimentation but today, experiments are clearly delineated from artworks. Can experiments, like artworks, elicit genuine aesthetic responses and bring about emotions such as wonder and awe? If so, what is the function of such responses in science? Furthermore, we often take it that great works of art can leave us with a better understanding of the world, but it is a matter of debate as to whether this affects the value of a work of art qua work of art. Are there parallels, then, between the beauty of an experiment and its role in providing knowledge and understanding, and discussions concerning the relations between the aesthetic and cognitive value of art?

This symposium offers the first detailed discussion of the aesthetics of scientific experiments, and therefore takes a step towards filling a gap in the current aesthetics of science literature. In doing so, our session will also illustrate the productivity of potential connections between scientific and artistic practices via a consideration of the aesthetics of experiments.

Bibliography


Analogical reasoning is central to many historical and current scientific debates, and is closely related to the question of unity: unity is the endpoint of any claim that the relation between two notions is ‘more than an analogy’. This symposium explores analogy and unity in the context of the ontology of chemistry. It will bring together early career philosophers of chemistry with different approaches to chemical ontology. The symposium will show how the ontology of chemical entities — with a focus on chemical elements, molecules and molecular structure — is influenced by the use of analogical reasoning and by the unificatory connections that are posited with more fundamental theories.

Chemical entities take many forms — from the atom and its subatomic structure, through the molecule to chemical substances. One can think of chemistry as the study of these entities, their inner structure, and transformations. Philosophy of chemistry explores them by examining the concepts and chemical theories describing them, and studying the relations that exist between chemistry and other sciences. This symposium brings together early career researchers whose contributions to the philosophy of chemistry involve questions about chemical entities. For example, what was the role of analogical reasoning in the discovery of chemical elements, or in the formulation of the periodic table? To what extent are different theories of such entities unified?

The symposium will investigate chemical ontology by employing two key concepts: analogy and unity. The British chemist Joseph Priestley argued that “analogy is our best guide in all philosophical investigations.” According to the SEP, “an analogy is a comparison between two objects, or systems of objects, that highlights respects in which they are thought to be similar.” While the role of analogies and analogical reasoning as a methodological tool in physics have been thoroughly explored (building on the foundational work by Mary Hesse), a similar investigation with respect to chemistry is relatively absent.

The second concept is unity. We examine unity both as a hypothesis but also as a tool to understand chemical ontology. In the philosophy of chemistry, ideas of unity have been associated with reductionism and taken to undermine any genuine talk of chemical ontology. In this symposium we show that this need not be the case: there are forms of unity through which we can draw valuable conclusions about chemical ontology.

We take our symposium to be a positive contribution to discussions about how chemistry is related to physics. The philosophy of chemistry, in its attempt to defend its autonomy as a separate and autonomous branch of philosophy, has often investigated the relationship between physics and chemistry from a negative perspective. For example, arguing against the reduction of chemistry to physics, or against a quantum mechanical understanding of the periodic law, have been standard ways of defending the autonomy of the philosophy of chemistry. In opposition to this, our symposium is a novel contribution because it examines the relationships between physics and chemistry from a positive perspective. Both fields are highly interconnected and have mutually influenced and inspired each other. Highlighting these features of chemistry’s relation to physics does not undermine the validity of philosophising about chemistry, rather the opposite.

Lastly, this symposium contributes not only to specific themes within the philosophy of chemistry, but also to the philosophy of science as a whole. It will analyse the concepts of unity and analogy within the context of chemical (theoretical and experimental) practice, as well as propose new reflections and concepts that arise from thinking about chemistry. While the field of philosophy of
chemistry is still largely underrepresented within philosophy of science, the study of chemical practice and ontology cannot only help rethink some generally admitted ideas about science: it also leads to its own set of philosophical ideas that can enrich philosophy of science in general.

The talk “Chemical Analogy, Classification and the Nature of the Halogens” aims to contribute to the debate on the concept of chemical element, which is a current issue in the philosophy of chemistry, by studying the attribution of elementary status in the early nineteenth century. It shows a strong link between chemistry’s relational ontology and the discovery of new chemical elements: the nature of chemical bodies could only be considered within the context of classification, and never completely individually. Analogy in chemical properties was thought to indicate analogy in composition, and therefore a new elementary substance could only be established as part of a class of substances which all had similar chemical behaviour and fit into a coherent theoretical framework.

The talk “Group-Theoretical Analogy” explores two familiar themes in the philosophy of chemistry: the nature of the chemical elements and the periodic law. The talk offers a critical examination of recent attempts to explain the overall structure of the periodic system from a group-theoretical point of view, inspired by the success and in close analogy with the group-theoretical approach in elementary particle physics. It argues that the success of this symmetry program hinges on an important shift in the ontology of the chemical elements, and retraces the origin of this shift back to Heisenberg’s work on the isospin of nucleonic particles in elementary particle physics.

The talk “The Problem of Molecular Structure Just is the Measurement Problem” investigates the implications that different interpretations of quantum mechanics have on the metaphysics of molecular structure. While these problems are often recognised as formally analogous, since both involve the breaking of underlying symmetries on measurement, we argue that their relation is more than an analogy: that they are one and the same problem. In general, this talk aims to open the way for a mutually informative interchange between the foundations of quantum mechanics and the foundations of chemistry communities. While work in the former has clear implications for the latter discipline, insights from chemistry and its foundations could also lead to a deeper understanding of the nature of quantum mechanics.

The talk “The Unity of Chemistry with Quantum Physics” examines one of the most widely-discussed questions in the philosophy of chemistry; namely, how is chemistry related to quantum physics, and what does their purported relation tell us about chemical ontology? The talk presents a novel model of non-reductive unity which is based on how the discovery of chemical entities and their properties have been largely influenced by the explanatory, confirmatory and predictive interconnections between the two sciences. This model not only correctly characterises how the two sciences are related, but also allows one to be a non-eliminativist about chemical entities, without having to hold a pluralist or strongly emergentist position.
Though there is renewed critical interest in the field of mental health, psychotherapy has been relatively neglected, not only by philosophers but by the humanities and social sciences more broadly. Yet, psychotherapy is not only a staple of mental healthcare but contemporary life. Talking therapies are often presented as safe and versatile interventions, an alternative to pharmaceuticals with their potential side effects. In addition to treating mental illness, individuals increasingly face an expectation to seek professional therapeutic services when experiencing challenging circumstances more generally. In this symposium, we turn our attention to conceptual and evidential questions arising in relation to contemporary psychotherapy.

Psychotherapy is a staple of contemporary mental healthcare. It comprises a diverse landscape of practices that see talking as therapeutic, and perhaps curative. Such talking therapies are often viewed as safe and versatile interventions, effective alternatives to pharmaceuticals with their potential side effects. Cognitive Behavioral Therapy (CBT), for example, is the treatment of choice for depression and anxiety under the NHS. Psychotherapy is also ubiquitous in contemporary life. Even absent a formal psychiatric diagnosis, individuals frequently face an expectation to seek professional therapeutic advice in response to challenging life events—ranging from grief to marital strife to midlife crisis. Despite being such a widespread part of contemporary mental healthcare and life, several conceptual and evidential questions remain: What makes contemporary psychotherapy therapeutic? What is it about these talking therapies that makes them effective? Is it about the specific techniques that psychotherapists employ? Or is it something more general and diffuse—interpersonal qualities like empathy or the therapeutic relationship—that underlies their effectiveness? How should we assess their effectiveness?

Insofar as philosophers—and scholars in the humanities and social sciences more broadly—have surveyed the field of mental health, their focus has largely been on the dangers of psychopharmaceuticals or issues of psychiatric diagnosis. Less attention has been given to conceptual and evidential issues arising in psychotherapy—especially non-psychoanalytic varieties of psychotherapy that are tightly integrated with contemporary healthcare systems. These psychotherapies typically present themselves as scientifically grounded, their effectiveness established through the same methodologies in use across medicine more widely—namely, Randomized Control Trials (RCTs). For example, since the seminal 1977 study where CBT outperformed an antidepressant in a head-to-head RCT, CBT has gone on to become a prime form of treatment in various healthcare systems due to its impressive evidence base (Rush et al., 1977; NICE, 2009). But is the medical model generally—and Evidence-Based Medicine (EBM) specifically—an appropriate one for psychotherapy? Is this the right way to assess psychotherapy? And what exactly are we assessing?

In this symposium, we interrogate these questions about the conceptual and evidential basis of psychotherapy. The symposium is comprised of three talks addressing questions about (1) what it is that we are assessing when we assess the effectiveness of psychotherapy and (2) what kinds of evidence are most apt for assessing efficacy when faced with a varied and contextually-dependent practice. Sahanika Ratnayake argues that neglecting qualitative data—especially that gained through introspection—in assessments of psychotherapeutic efficacy is misguided and hypocritical. Riana Betzler uses the work of humanist psychotherapist Carl Rogers to argue for a more expansive approach to evidence for psychotherapeutic efficacy—specifically one that considers the tacit dimension of psychotherapeutic practice. Hannah Blythe points to conceptual problems with the extant division between the technical elements of psychotherapy and the contextual and
interpersonal elements. She argues that the lack of attention to context complicates assessments of efficacy and threatens their generalizability.

These three talks raise questions that intersect closely with issues of interest in many other areas of the philosophy of science—in particular, philosophy of medicine and the social sciences. Several philosophers have critiqued EBM and RCTs in recent years (e.g., Cartwright, 2007, 2010; Howick, 2011; Stegenga, 2014; Worrall, 2002). This symposium engages with and extends these critiques by applying them to the case of psychotherapy. Psychotherapy is an especially interesting case for considering the limitations of EBM and RCTs because of its multi-component structure; it may be the case that some elements of psychotherapies (for example, technical components) may be amenable to the use of RCTs while others (for example, interpersonal components) are not. But can these components really be separated from one another? The case of psychotherapy also raises important questions about generalizability and external validity analogous to those already addressed in Cartwright’s work on the use of RCTs for establishing the effectiveness of policy interventions.

Although psychotherapy, on the surface, may seem to fit within the structures of traditional medical models, it may actually be much more similar to policy interventions in that at least certain elements seem to be highly context dependent. Furthermore, context dependency in psychotherapy may occur at an even finer grain than that at work in policy, given that it sometimes seems to hinge on the individual relationship between therapist and client.

Beyond the clear intersection between debates about the limitations of the medical model and the evidential basis of psychotherapy, this symposium also engages with other themes that have been prominent in recent work in the philosophy of science—on the aims of therapy, whether care or cure (e.g., Stegenga, 2018); on the relationship between mental health and well-being (e.g., Wren-Lewis & Alexandrova, forthcoming); on the dichotomy between the art and science of medicine and whether such a dichotomy makes sense (e.g., Solomon, 2015); on medicalization and the dangers of pathologizing ‘problems in living’ (e.g., Sholl, 2017); and finally, on the role of the therapeutic alliance, doctor-patient relationship, and skills like empathy in healthcare (Ferry-Danini, 2018; Howick et al., 2018; Halpern, 2011).

By providing a critical examination of psychotherapy, a largely neglected area of practice thus far, this symposium extends current debates within philosophy of science and draws connections between diverse areas. Untangling the conceptual and evidential issues in psychotherapy is furthermore pressing given its ubiquity in contemporary medical practice and everyday life. This symposium provides a unique opportunity to begin to grapple with these central questions and take steps towards a more philosophically informed approach to psychotherapeutic practice.

References


Dependence in Physics

Alastair Wilson, Katie Robertson, Michael Townsen Hicks

The proposed symposium consists of three 20-minute talks, with 5 minutes Q&A following each talk and a 15-minute panel discussion to conclude. This symposium explores the general notion of dependence as it applies within physics, and contrasts various different relations of dependence which appear to be needed to understand the practice and content of physics. In particular, the symposium talks focus on the dependence relations of reduction (both ‘horizontal’ and ‘vertical’ varieties), grounding and causation. Reduction and grounding are typically conceived as (in different ways) relating multiple levels of physical reality, while causation is supposed to relate events at a single level. However, every case of reduction is highly contested, there are well-known difficulties in making sense of causation in fundamental physics, and grounding presents additional conceptual problems which have as yet received little attention. The aim of this symposium is to classify these different notions of dependence, to explore their similarities and differences, and to illustrate their application to physical contexts including the mechanics of planetary orbits, the relationship between statistical mechanics and thermodynamics, and the relationship between symmetries and conservation principles.

Dependence in physics is a topic of special contemporary interest, as it serves to connect up active literatures from distinct subfields of philosophy that have tended to proceed in isolation from each other. Causation and grounding have been studied by metaphysicians largely in isolation from actual scientific practice; philosophers of science have generally focused on causal explanation, intertheoretic reduction and (more recently) on non-causal explanation; philosophers of physics have tended to think about dependence in a case-by-case way, responding to specific interpretive difficulties. This symposium aims to connect up these debates and to thereby to move them forward, enhancing our overall understanding of dependence in physics.

The most familiar form of dependence in science is causal dependence, but its application to physics remains highly contested. Since Russell argued that fundamental physics has no room for a temporally asymmetric notion of causation, it has been widely assumed that causal notions are inapplicable there; but physicists continue to use causal language, and even setting aside causation we may still need to make sense of a notion of nomological dependence which has much in common with causation. The talks by Wilson and Hicks engage with these puzzles; Wilson also argues that making sense of some cases of causation in physics requires interventions which violate fundamental laws of nature and may even violate metaphysical constitutive principles.

Causation doesn’t seem to be applicable to the explanatory relations between earlier and later theories, or to the dependence of higher-level goings-on on lower levels, as when thermodynamics is explained using statistical mechanics or when stellar spectra are explained in terms of nuclear energy gaps. The classic theoretical tool in philosophy of science for handling such explanation is reduction, but this notion has fallen out of favour in recent times. Most interesting reductions remain deeply contested, and over the last few decades numerous challenges have been raised to a reductionist picture – challenges deriving from abstractions, approximations, idealizations and other practical features of a modelling methodology. Commonly, two forms of inter-theoretic reduction are distinguished: vertical reduction is a dependence relation between theories describing different scales, or subject matters, whereas horizontal reduction is a relation between an old theory and its successor. But a principled distinction between these forms of reduction has not yet been provided. In Robertson’s talk, she argues that whilst the practice of approximation is central to horizontal reduction, abstraction is central to vertical reduction. Distinguishing abstraction from approximation is a difficult task, but Robertson argues that the notion of a subject-matter can help. Thus, the notion
of reduction is defended from these attacks and emerges reinforced, as a strong potential contender for understanding the relations between levels in physics. While we may still opt to bring further non-causal dependence relations into the picture, this move should not be motivated by a perceived failure of reduction to account for the classic cases of interlevel dependence in physics.

Even if reduction can be rehabilitated, there are important cases of non-causal dependence in physics which appear to need a different treatment. The relationship between symmetries and conservation laws which is captured by Noether’s theorems is an example of particular interest, which plays a central role in Hicks’ paper. While the prospects for reducing symmetries to conservation laws (or vice versa) seem dim, and it is difficult to understand the relationship in causal terms, the theorems still seem to reveal a deep connection between them. Hicks argues that this connection is best understood as an instance of metaphysical grounding – of conservation laws in symmetries. This runs contrary to the influential proposals of Marc Lange (who appeals to meta-laws constraining nearby nomic possibilities) and of Harvey Brown (who explains both symmetries and conservation laws in terms of the underlying dynamics). Simultaneously with recent developments on the topic of scientific explanation, but often addressed in isolation from them, contemporary metaphysics has come to focus on a distinctive metaphysical type of explanation: grounding. When A grounds B, B is the case because A is the case—where ‘because’ is intended to be understood in a special metaphysical sense. Grounding, which is usually introduced by explicit contrast with causal explanation, plays an increasingly prominent role in contemporary metaphysics. Over the last decade, the notion has come to be the tool of choice for many philosophers in thinking about fundamentality and emergence. As physicists look to probe the fundamental laws of nature in particle accelerators and in cosmological observations, metaphysicians such as Bennett, Fine, Lewis, Schaffer, and Sider have seen their task as being to answer certain general questions about the fundamental nature of reality and about how non-fundamental reality is grounded in, or metaphysically explained by, the fundamental. Since we have clear logical frameworks for understanding grounding, developed by Fine, Correia, Litland, and others, grounding would seem like a natural tool with which to understand disputed cases of non-causal explanation within physics. But in practice, progress in metaphysics in developing the concept of grounding has largely been neglected by philosophers of physics, while metaphysicians working on grounding have rarely considered the complexities of applying the notion to real-world physics examples. The papers by Hicks and Wilson address this gap. Hicks explores the relationship between symmetries and conservation laws captured in Noether’s theorems, and defends the applicability of grounding in this case. In contrast, Wilson raises novel difficulties for grounding in physics which arise from the way that grounding relations generically entail counterfactuals with metaphysically impossible antecedents.

Overall, this symposium will disentangle different notions of dependence in physics and investigate their application to a range of pressing questions in philosophy of physics: concerning symmetries and conservation laws, causal dependencies in classical mechanics, and the relations between theories at different levels including statistical mechanics and thermodynamics. The upshot will be an enhanced understanding of how the varieties of dependence relate to each other and of how best to model them.
Epistemic Diversity and Argumentative Practices in Science

Silvia Ivani, Caglar Dede, Merel Talbi, Roosmarijn van Woerden, Elias Anttila

Argumentation seems to play an important role in science. Epistemic diversity is often seen as a factor that facilitates the epistemic success of argumentative practices in science. However, recent studies have raised questions on the actual epistemic merits of argumentative practices such as peer review in science. Understanding whether and how diversity can play a beneficial epistemic role in argumentative practices is a particularly urgent issue, given the increasing importance of interdisciplinary research and participation of different stakeholders in scientific practices. This symposium examines the role of epistemic diversity in scientific argumentative practices by focusing on four different contexts: the collaboration between experts and non-experts in designing scientific research projects; the joint evaluation of evidence-based policy decisions involving policy-makers, scientists, and laypeople; the interdisciplinary exchange within science; and the role of expertise and epistemic privilege in the ethics of argumentation in science.

Argumentation can be defined as the communicative activity of producing and exchanging reasons in order to support or challenge claims and positions. Argumentation is a pivotal (but demanding) human practice, especially in situations of doubt or disagreement. Across different societies and cultures, argumentation plays a prominent role in scientific inquiry, legal procedures, education, and political institutions. However, there are many instances in which argumentation does not achieve its presumed goals: instead of consensus, it leads to polarization; instead of circulation of reliable information, it leads to the propagation of falsehoods or dogmatism; instead of fostering sound decision-making, it leads to suboptimal choices. Recent developments such as the Brexit campaign, climate science denial, anti-vaccination movements, among others, seem to be examples of argumentation failing to fulfill its epistemically beneficial roles. These and similar occurrences put a strain on the idea that rational argumentation can be successful. Hence, it is critical to understand the conditions under which argumentation among epistemically diverse groups can achieve its presumed goals. This is a timely and important issue in contemporary social epistemology (Goldman 1994, 1997; Dutilh Novaes, 2020), and more recently, in scientific approaches to argumentation (e.g. Mercier 2020).

This symposium focuses on epistemic diversity and argumentation in science. Since science is a disciplined collective epistemic activity, it constitutes one of the best places to test, inform, and generate philosophical insights about argumentation. One prominent understanding of science presupposes that scientists seek to persuade peers to accept their findings and that scientific knowledge acquires its objectivity thanks to the peer review process among epistemically diverse agents holding different theoretical and methodological background assumptions (see, for example, Bonilla, 2006, Longino 1990). Epistemic diversity within science is hence understood to be crucial for scientific progress. After all, scientific evidence is often produced by diverse scientific methods whose outcomes can be verified, corroborated, or replicated so as to count as convincing arguments supporting the theories in question (e.g., Longino 1990, Goldman 1997).

However, recently, some philosophers have challenged this optimism regarding the virtues of diversity by demonstrating its limited epistemic merit (Weisberg and Muldoon 2009, Zollman, 2007; 2010, Heesen and Bright 2020). Moreover, a number of studies have highlighted how argumentation in scientific communities is not immune to the propagation of falsehoods and polarization (e.g., O’Connor and Weatherall 2019). For instance, Heesen and Bright (2020) directly challenge the desirability of peer review in science.
The symposium seeks to contribute to these recent debates concerning the nature and desirability of epistemic diversity in science by bringing together researchers with diverse backgrounds working on scientific argumentative practices and epistemic diversity in the ERC-funded project The Social Epistemology of Argumentation led by Catarina Dutilh Novaes at VU Amsterdam. The contributed papers will focus on the argumentative practices in science from a philosophy of science perspective. Specifically, each author critically analyzes a prominent instance or context in which scientists engage in argumentative practices. These include scientists’ engagement with citizens in designing research projects (Ivani) and in evaluating evidence-based policy (Dede), interdisciplinary exchange within science (Van Woerden and Talbi), and the role of expertise in the ethics of argumentation (Anttila).

In “Epistemic Diversity and Citizen Engagement in Science,” Ivani explores whether and to what extent the diversity injected in science through the practices of citizen engagement can be epistemically and socially beneficial. Dede contrasts the deliberative polling and the collaborative research practices as two rival approaches to identify citizens’ values for the complementing evidence-based policy evaluations. Van Woerden and Talbi analyze the tensions that emerge when interdisciplinary research aims to be epistemically plural and integrationist at the same time. Finally, Anttila argues that experts can accumulate further credibility as experts by committing deliberate or unintentional argumentative injustices.

References


Evidential pluralism and Causality in the Special Sciences

Jon Williamson, Michael Wilde, William Levack-Payne, Yafeng Shan, Samuel D. Taylor

Evidential Pluralism maintains that in order to establish a causal claim one normally needs to establish the existence of an appropriate correlation and the existence of an appropriate mechanism complex, so when assessing a causal claim one ought to consider both association studies and mechanistic studies. This thesis has led to fruitful philosophical work on the role of mechanisms in the biomedical sciences and to suggestions for improvements to evidence-based medicine (‘EBM+’). The question arises as to whether it can also be applied in other contexts. This symposium will examine Evidential Pluralism and causality, and their use in the social sciences, sports science and cognitive science.


Evidential pluralism is the thesis that establishing a causal claim in medicine typically requires both establishing a correlation and establishing a mechanism to explain that correlation (Russo and Williamson 2007). Jeremy Howick (2011) and Alexander Broadbent (2011) have put forward a number of potential counterexamples to this thesis. In particular, there seem to be cases where tightly controlled comparative clinical studies alone suffice to establish a causal claim, in the absence of evidence of a mechanism. Unfortunately, in its original formulation, evidential pluralism is ambiguous in a number of crucial respects (Illari 2011). This makes it difficult to determine whether the putative counterexamples to the thesis are successful. In this talk, I aim to make clear the commitments of evidential pluralism by focusing on the theoretical motivation behind the thesis. In particular, I suggest that the natural way to think about evidential pluralism has been as a response to both the problem of confounding and the problem of masking (Clarke et al 2014). However, I will argue that the proposed counterexamples show that this natural way of thinking about evidential pluralism cannot be correct. I then propose an alternative way of thinking about evidential pluralism that is not susceptible to these counterexamples. Hopefully, this talk will serve as a brief introduction to the main issues concerning evidential pluralism in the philosophy of science.


As many will know, the practices of Evidence Based Medicine (EBM) are being adopted outside of medicine. One area where this is taking place is in the sports sciences. There is a push to move into an Evidence Based Practice (EBP) framework, which involves the adoption of EBM evidence hierarchies. Just like EBM, EBP emphasises the idea that Randomized Controlled Trials (RCTs) produce strong evidence, which can regularly be relied upon to establish causation, and inform practice. In this paper, I outline the criteria commonly regarded as necessary for a RCTs to provide strong evidence. I, then, briefly, argue that RCTs in the sports sciences very often do not meet these criteria. I argue that, due to the nature of the research conducted in sports science, this is often unavoidable. This means that in many areas of sports science, RCTs and systematic reviews and meta-analyses of them are simply not sufficient to establish causal relationships.

This argument sets up the main bulk of the paper. This involves a discussion of the Russo-Williamson Thesis (RWT) and its applicability to the sports sciences. The RWT is the claim that we normally need evidence of a correlation and a mechanism in order to establish a causal claim. I discuss the historical case of creatine as a sports performance supplement, and argue that it took both evidence of correlation and mechanism to establish its effectiveness as an ergogenic aid. This provides motivation for the claim that the RWT, or something like it, is applicable in the sports sciences. I,
then, contend with a common objection to the RWT. This is the objection that, historically, effectiveness has been established on the basis of evidence from RCTs, without the use of evidence from mechanistic studies. As such, the objection continues, we do not need evidence of a mechanism to establish causal claims. I argue, informed by the work of Jon Williamson, that this objection does not counter the RWT, as evidence from association studies, such as RCTs can, if certain criteria are met, be sufficient to establish the existence of a mechanism. This allows that evidence from RCTs can, in some instances, suffice to establish causation. I finish the paper by using the case of caffeine research in sports science to give an illustrative case where evidence from association studies was sufficient to establish both a correlation, and a mechanism, in the sports sciences.

Talk 3: “Applying Evidential Pluralism to the Social Sciences” by Yafeng Shan and Jon Williamson (Kent).

Since around the year 2000, philosophers of science have produced a great deal of interesting research on the role of mechanisms in science. One strand of this research concerns the role of mechanistic evidence in establishing causal claims. Russo and Williamson (2007) argued that in the biomedical sciences, a causal claim is established by establishing (i) that the putative cause and effect are correlated, and (ii) that there exists a mechanism linking the two which can account for this correlation. This thesis has the following important consequence: while quantitative studies (in particular, randomised controlled studies) provide excellent evidence of correlation and, in the right circumstances, can provide evidence of the existence of a mechanism, it is important to also consider other evidence of mechanisms when assessing a causal claim. This motivates a kind of evidential pluralism.

In medicine, this form of evidential pluralism has led to a proposed modification to evidence-based medicine, called EBM+. Parkkinen et al. (2018), for instance, developed procedures for evaluating mechanistic studies alongside clinical and epidemiological studies, when assessing the effectiveness of an intervention or when ascertaining the effects of exposure to an agent.

This paper argues that evidential pluralism applies equally to the social sciences, where it leads to new foundations for mixed methods research. In the social sciences, as in the biomedical sciences, establishing causation requires establishing both correlation and mechanism—social mechanisms, in this case. While quantitative association studies can provide some evidence of mechanisms, in addition to good evidence of correlation, other sorts of study also provide good evidence of social mechanisms—notably, certain qualitative studies.

We argue that there is scope to apply evidential pluralism to the social sciences. First we show that the lessons from evidence-based medicine can be carried over to evidence-based policy, and that evidential pluralism can provide an account of the assessment of evidence in evidence-based policy. We compare this account to that provided by realist evaluation, which also has a central role for mechanisms. Second, we use case studies to argue that evidential pluralism additionally applies to more theoretical social sciences research, and can be used to elucidate the confirmation relations in basic social sciences research. Third, we show that evidential pluralism can provide new foundations for mixed methods research, because it offers a precise account of the need for mixed methods when establishing causation in the social sciences.

We then respond to two objections to the claim that evidential pluralism can be applied to the social sciences: one due to Julian Reiss and a second due to Francois Claveau. We conclude that evidential pluralism has much wider scope than originally envisaged, and sheds new light on the use of evidence in the social sciences.

According to some philosophers of cognitive science, only mechanistic explanations of cognition count as genuine, causal explanations of cognition, because only evidence of mechanisms reveals the causal structure of cognition (cf. Kaplan and Craver, 2011; Piccinini and Craver, 2011). On this view, a causal explanation of cognition will carry explanatory force only to the extent that it identifies, through analysis, the component parts of the mind/brain (e.g. neurons, modules, etc.) and their principles of interaction, before showing how these component parts causally interact to generate some phenomena (cf. Machamer, Darden, and Craver, 2000).

But this view is not without its critics. In fact, some philosophers now argue that we can have genuine causal explanations of cognition that abstract away from mechanistic detail to characterise the causal structure of cognitive systems in nonmechanistic terms (cf. Chemero, 2009; Chemero and Silberstein, 2008; Van Gelder, 1995; Van Gelder, 1998). On this competing view, we cannot rule out the possibility of developing genuine, causal explanations of cognition that make no reference to the mechanistic parts and activities of the mind/brain. The upshot is that we cannot jump to the conclusion that causal-nonmechanistic explanations of cognition are ipso facto spurious or defective.

My claim here is that the tension surrounding the possibility of causal-nonmechanistic explanations is grounded in conflicting intuitions about the mechanistic account of causality, which holds that C is a cause of E iff there is the appropriate sort of mechanism (or chain of mechanisms) linking C and E. I argue that critics of causal-nonmechanistic explanations are invariably committed to the mechanistic account of causality, but defenders of causal-nonmechanistic explanations cannot endorse the mechanistic account of causality without being committed to two inconsistent claims: that we can have causal-nonmechanistic explanations of cognition and that causal structure is entirely mechanistic.

The problem, however, is that we cannot jettison the mechanistic account of causality without introducing another account of causality from which to derive the satisfaction conditions for genuine, causal explanations of cognition. To fill the gap, I defend the epistemic theory of causality (ETC), which holds that “causality is a feature of the way we represent reality rather than a feature of agent-independent reality itself” (Williamson, 2013, p. 268). The idea of ETC is that we are justified in taking there to be a causal relation between X and Y iff the causal relation between X and Y is part of a causal structure that yields some set of successful prediction, explanation, and control (PEC) inferences.

I argue that ETC supports an epistemic conception of causal explanation in cognitive science, which gives rise to a new satisfaction condition for genuine, causal explanations of cognition: that for an explanation e to be causal, e must have appropriate epistemic utility by yielding successful PEC-inferences. I conclude that when we endorse ETC, we can reject the claim that only causal-mechanistic explanations of cognition are genuinely explanatory, while leaving open the possibility of both genuine causal-mechanistic and genuine causal-nonmechanistic explanations in cognitive science.
The proposed symposium consists of three 20-minute talks, with 5 minutes Q&A following each talk and a 15-minute panel discussion to conclude. This symposium explores the contemporary role of fictions in science and metaphysics beyond the recent developments in the literature on scientific models. We do so by addressing three questions — (1) What are fictional objects? (2) What can we know about fictional objects? (3) What is the relation between fictional objects and the world? — in three distinct but interrelated topics: (i) scientific thought experiments, (ii) counternomics, and (iii) computer programs. In doing so, we demonstrate the fruitfulness of the concept of fiction throughout philosophy of science and metaphysics, and stimulate further research in the world of philosophical fictions.

The concept of fiction is increasingly gaining attention and application throughout contemporary philosophy. Philosophy of science is no exception: demons, omniscient beings, empty universes, and superluminal travelers are indeed ubiquitous in science. Nonetheless, philosophical accounts of fiction have successfully been extended to branches of philosophy of science where forms of fiction do not at first sight appear to be present. Such is the case most notably in the study of scientific models, where the concept of fiction has given rise to various fiction views of scientific models (e.g. Frigg and Nguyen 2020). Besides fictionalist theories of models, the question now is where else in philosophy of science we might fruitfully apply the concept of fiction.

When we apply philosophical accounts of fiction in science and metaphysics, we must confront three types of questions concerning fictions: (1) What are fictional objects? (2) What can we know about fictional objects? (3) What is the relation between fictional objects and the world? This panel aims to increase understanding of fictions in science and metaphysics beyond scientific models by addressing questions (1)–(3) in the context of three interrelated topics: (i) scientific thought experiments, (ii) counternomics, and (iii) computer programs.

Question (1), the ontological problem, addresses the following questions: in which ontic category do fictional entities fit? Do we actually have good reasons to place fictional entities in an ontological category? Those who answer the second question in the affirmative belong to the fictional realist camp. They stand in contrast to fictional antirealists, who deny the existence of fictional entities. According to the antirealist, there are no such things as fictional entities. For instance, in pretense theory (Walton 1990), fiction is understood merely as make-believe. Proponents of fiction views of models realized that, in scientific inquiry too, models are often referred to as if they were concrete objects that exist in space and time; scientists often describe models of, e.g., springs, pendula, and predator-prey scenarios as if they exist and as if there exist facts about these fictions that can be discovered and studied. This latter statement concerns the aforementioned question (2): the epistemological problem of what, if anything, can be known about fictional objects.

In his talk A Fiction View of Scientific Thought Experiments, Rijken argues that there is a subject closely related to scientific models where questions (1) and (2) are surprisingly independent from question (3): scientific thought experiments. Rijken proposes an account of scientific thought experiments based explicitly on the fiction view of models, using the following characterization: to perform a scientific thought experiment is to reason about a model with the aim of solving a problem about that model — where models are construed as acts of pretense according to a specific fiction view (Salis 2019). A central problem concerning scientific thought experiments is often called the paradox of thought experiments: how can we learn about the world by performing thought experiments, that is, by merely using the imagination? (Kuhn 1964) Using his fiction view of thought
experiments, Rijken argues that we are able to dissolve the aforementioned paradox in the following way: by performing a scientific thought experiment, we learn directly about a mode
not about the world. We can thus exhaustively analyze the epistemic role of scientific thought experiments by answering only questions (1) and (2).

While the idea that, when modelling, scientists engage in pretense has at least a heuristic appeal, it leaves answered unsatisfactorily the ontological question of what scientific models actually are (question (1)); are they merely fleeting figments of the imagination or are there reasons to suppose they exist substantially? Recently, Thomasson (2020) and Thomson-Jones (2020) have suggested an artificialist view: pretense only plays a role ‘within’ the model, externally scientific models are abstract artifacts created by scientists when producing texts and descriptions. It is just the internal content of model descriptions that should be taken to involve pretense, while externally the models are abstract artifacts. In his talk A Fiction View of Computer Programs, Wiggershaus extends this artificialist view to computer programs. However, whereas models are usually held to merely explain or predict the phenomena of their respective target system, programs seem to prescribe the behavior of computational systems. This raises the question how programs qua abstract artifacts can be executed and bring about effects by guiding real-world computers (question (3)). For that matter, a comparison to works of music is drawn, which can be executed in performances or by special kinds of automata like player pianos.

Question (3) — how fictions relate to the world — also features prominently in Iranzo-Ribera’s talk A Guide to Fiction in Counternomic Reasoning. Counternomics are counterfactuals with nomologically impossible antecedents, which feature prominently in, e.g., scientific explanation, model-based reasoning, and (more relevantly for this proposal) thought experiments. Controversy surrounds counternomics, not least because on the view that laws of nature are metaphysically necessary, these counterfactuals come out vacuous under the application of the dominant Lewis-Stalnaker semantics over possible worlds. This contrasts starkly with scientific practice, which takes these counterfactuals to be substantive: some of them are deemed true (or at least ascertainable), others false. Iranzo-Ribera argues that one fruitful way to think about these counternomics’ antecedents is to regard them as fictional claims, whose truth (or at the very least, assertability) is grounded on facts about the world.

If facts about the world are what ultimately grounds fictional claims, one might worry why fictionalism is needed in the first place. Another useful distinction that has been drawn in the study of fictional entities is that between their ontology and the activity of engaging in scientific games of make-believe (Suárez 2009). In line with this distinction, Iranzo-Ribera claims that fictionalism about counternomics highlights the cognitive value of fictions: reasoning about nomologically impossible states of affairs is more akin to the practice of fictionalizing — engaging in games of make-believe — in science than to that of reasoning about (im)possible worlds equipped with some semantic theory.
Model Transfer in Science

Catherine Herfeld, Wybo Houkes, Tarja Knuuttila, Andrea Loettgers, Dunja Seselja

The transfer of scientific models within and across different domains is a conspicuous phenomenon. In almost all scientific fields, it is a common practice that models developed to advance knowledge in one domain are applied in another to study sometimes similar but often also fundamentally different problems. This transdisciplinary aspect of scientific modeling has been the center of recent literature about knowledge transfer more generally. At the same time, the phenomenon of model transfer is connected to issues arising in various debates in philosophy of science, from discussions in the philosophy of modeling to the debates about interdisciplinarity, application-driven research, and even about the public perception of science. Yet, the philosophical discussion on model transfer is still at an early stage and the phenomenon often not directly targeted in such discussions. This symposium aims to fill this gap. Contributions not only address questions around the forms that model transfers can take, the conditions that have to be in place in order to ensure their success, and the way in which success needs to be understood in this context. They also study the challenges that the transfer of models confronts. The aim is to take the debate about model transfer one step further and thereby link key questions and results to related discussions in philosophy of science.

Recent discussions in philosophy of science have emphasized the importance of a cross-domain transfer of scientific models as one of the central catalysts for scientific innovation and progress in a field. Such model transfer can take different forms. It can lead to an adaption to, or even integration of the transferred model with already established knowledge in the target domain. In such cases, model transfer is *prima facie* smooth and, ideally, advantageous in that it ultimately leads to new insights and thereby to progress in the target domain.

Examples in the recent literature of cases for such model transfers are numerous. Psychological models have been introduced into economics to acknowledge various motives behind human decision-making and explain phenomena such as pro-social, habitual, or addictive behavior. The underlying idea is that such models could help economic theories to better explain the effect of tax policies, trading behavior, or exchange of goods beyond the market place (Falk et al. 2008, Lisciandra 2018, Rabin 2002). Other examples are the transfer of models from engineering into synthetic biology, of game theoretic models from mathematics into evolutionary biology, or of models from physics and complexity theory into mathematical finance. In the latter example, known under the label of econophysics, the hope is to better predict the dynamics of financial markets by means of a different set of models than those based on standard economic assumptions (Knuuttila and García-Deister 2019, Thébault et al. 2018, Jhun et al. 2018).

While the literature on case studies has been flourishing, the philosophical discussion is still at an early stage. There are a number of crucial blind spots that require identification and systematic analysis in order to move the debate one step forward.

One set of open issues relates to the actual unit of transfer. When studying the phenomenon of model transfer itself, it becomes clear that what is being transferred sometimes involves specific models, such as, e.g., the Ising model, but sometimes also highly generic models, such as, e.g., coupled harmonic oscillators, Lotka-Volterra-equations, or a game theoretic framework. As a consequence, some philosophers take the unit to be the concrete model itself, while others follow Paul Humphreys (2002, 2004) and argue that what is been transferred is what Humphreys calls a formal template (e.g., Houkes and Zwart 2019, Knuuttila and Loettgers 2014). Such a template originates in some domain and becomes further specified in order to make it applicable in some target domain (Humphreys 2019). Such diverging views about the unit of analysis but also the
different implications of those views for the nature of the transfer process raise questions such as what formal templates exactly are, how they relate to scientific models, and what makes for their capacity to be transferred between sometimes very different domains and thereby applied to fundamentally distinct problems.

A second set of issues concerns the transfer process and its characteristics. For some time, philosophical analyses rested on the implicit assumption that model transfer occurs without substantial changes of the model in the transfer process. Recently, a number of case studies suggested that modifying the model and/or the subject matter to which it applies in the target domain is indispensable to a successful cross-domain transfer (e.g., Herfeld and Lisciandra 2019). For example, models from engineering have been increasingly used in synthetic biology only after extensive modification, where they have called into question a variety of basic principles of existing theoretical frameworks (e.g., Knuuttila and Loettgers 2013, 2014). Questions that arise in this regard are whether such processes have characteristic stages that can be identified in all transfer processes and that can be generalized across different domains (see, e.g. Herfeld and Doehne 2019). Furthermore, to better understand such transfer processes, both the conservation and adaptation aspects need to be studied. Finally, given that model transfer is a phenomenon that potentially influences the possibility for progress in a field, what needs to studied is whether there are different mechanisms at play, depending on source and target domain.

A third set of issues concerns the conditions that ground model transfers and enable their success. A first question that arises is why scientists engage in modeling transfer in the first place, rather than constructing a new model that is aligned with their accepted theoretical frameworks, their methodological and epistemological commitments, as well as their target problems. On the one hand, different social and psychological factors might be involved: opportunism, attempts for imitating success, the lack of a comparable alternative, or ‘imperialist’ tendencies have certainly initiated model transfer processes. On the other hand, epistemic and methodological factors such as the structural similarity of a problem, a shared set of methodological commitments, or the goal of theoretical unification equally play a role. While economists, for example are more open to model transfers if the object transferred can accommodate their main methodological, ontological, and epistemological commitments (Basso et al. 2017) they are hesitant to undertake them when a transfer would lead to fundamental conceptual changes (e.g., Sent 2004, Thébault et al. 2018, Bradley and Thébault 2019, Gingras and Schinckus 2011). However, those very same factors might equally prevent model transfer from being successful. Such cases provoke questions such as when model transfer is considered to be successful and whether we can identify some context-invariant success conditions.

Finally, while model transfer is frequently viewed as having occurred more or less successfully, such success may be only apparent. Particularly in instances in which models would have to undergo a modification or even transformation process in order for the model to become applicable, such a transfer can confront a number of challenges. Economics as a discipline is a prime example in which attempts for model transfers from physics, sociology, and psychology have been manifold but where such transfers often have to overcome substantial barriers to entry or sometimes even fail completely. Furthermore, even if a model is transferred and adopted in a specific domain, its transfer may not always have beneficial epistemic effects.

The contributions to this symposium will take one or more of those four issues and address them in a unique way. They thereby contribute to a clear agenda for philosophical work on model transfer in science, which includes methodological reflections on, and a set of normative guidelines for how such transfer processes should be studied in philosophy of science.
Models in Particle Physics: Three Challenges to Realism

Florian J. Boge, Martin L. King, Niels C. M. Martens

Particle physics aims at uncovering the fundamental interactions and building blocks of physical reality, and is therefore ripe for investigation by philosophers of science. Issues of realism can be fruitfully approached by examining the modelling practices involved in theory development and experimentation. We will address three challenges to scientific realism coming directly from modelling practices in particle physics:

1. How much reality can – and must – we attribute to the phenomenological simulation models used in collider experiments, given their various free parameters?

2. How can one be a realist in contemporary particle physics if physicists are increasingly turning to bottom-up approaches that do not represent anything that could be a candidate for reality?

3. In the specific case of models of the elusive dark matter particle – seemingly required by astrophysical observations but barely constrained by them – what would it even mean to be a realist about this myriad of conflicting models?

The debate over scientific realism is one of the long-standing topics in philosophy of science. With the continued attention it has received in recent years, as seen for instance in the lively debates on selective realism, perspectival realism, or realism in the quantum domain, its relevance needs no specific justification. Similarly, since the heyday of the semantic view of theories, scientific models have come into closer focus as theoretical entities in their own right, as reflected e.g. in recent monographs by Khalifa, Bail-Jones, Gelfert, or the classic collection by Morgan and Morrison.

Particle Physics is special as a case study for investigating the intersection of these two topics for a number of reasons. With its central entity – the infamous Standard Model (SM) – enjoying a high degree of experimental confirmation, particle physics offers a model of fundamental interactions between quantised fields as the primary candidate representation of reality. As is well known, however, there are several empirical and explanatory gaps that demand going beyond the SM (BSM) – one of them being the need for an additional, BSM dark matter particle, as the SM dark matter particle, the neutrino, only accounts for a fraction of the mass-energy that dark matter is expected to contribute to the universe. As is equally well known, none of the favoured scenarios for physics beyond the Standard Model have received any confirmation from the Large Hadron Collider. For that reason, the field is in a kind of crisis, and modelling-strategies have become more creative. An important recent trend is to move away from top-down models and towards model-independence, by focusing on simplified models (McCoy & Massimi, 2018) and bottom-up approaches like the SM effective field theory. Furthermore, as has been recognised by philosophers, such as Margaret Morrison, Wendy Parker, and Kent Staley, a number of distinct epistemological questions arise when models, especially simulation models, are integrated directly into the experimental practice of a field. Due to the vastness of data and the difficulties in approaching radiating and hadronising quarks from either theory or experiment directly, particle physics is one of the few fields where this is truly the case.

These facts, that theoretically favoured options in going beyond the Standard Model have failed whereas new alternatives need to be generated in ever more creative ways, and that many experimental details can only be tackled by equally creative modelling, raise a number of interesting
challenges for realism that have so far not been covered in the literature. Our symposium hence complements existing approaches to realism in the context of particle physics, such as those by Dawid, Williams, or D. Fraser.

In particular, we identify three challenges for realism about several entities and events that appear in three distinct stages of the modelling practice in experimental as well as theoretical particle physics. We discuss, in each case, what it would take to meet these challenges. The first challenge concerns the use of simulations in particle physics experiments, and arises generally for SM and BSM modelling. The other two challenges arise specifically in the context of the modelling associated with the theoretical response to the growing sense of crisis arising from a lack of BSM signs from particle physics experiments.

**CHALLENGE 1:** Given the presence of various free parameters in simulation models used to tackle the goings on in the ‘beam pipes’ of large colliders, is it possible to claim that these models represent those goings on?

**CHALLENGE 2:** Given the open-ended nature of the models formulated in the most general conservative approach to generating models beyond the Standard Model (the Standard Model effective field theory-approach), what conditions would have to be satisfied to even obtain a *candidate* representation of reality?

**CHALLENGE 3:** Given the diversity of modelling-approaches to dark matter and the lack of empirical constraints to remove the underdetermination between these different models, is it, at present, (already) possible to be a dark matter realist? If so, what would it even mean to be a realist about an entity that we barely know anything about? If not, what would it take to reach the point where we would be justified in being scientific realists about dark matter?

In three talks, we will offer precise formulations of the problems implicated in these challenges, as well as outlooks on how they could possibly be overcome. We are convinced that this can stimulate a fruitful discussion within philosophy of science, with implications exceeding the narrow context in which these challenges are developed.

For instance, the need to analyse the impact of parameters on simulation-intensive sciences has been recognised for disciplines like thermodynamics (Hasse and Lenhard, 2017), climate science (Parker, 2017), and astrophysics (Guegen, 2020). With particle physics containing a properly experimental (as opposed to observational) division, insights that can be gathered about the impact of parameters under such epistemically favourable conditions will likely have implications also for other branches.

Moreover, the importance of model independent methods and the implications for realism posed by effective field theories have recently begun to be explored by McCoy & Massimi (2018) and Williams (2018), but the present discussion focuses on bottom-up rather than top-down EFTs, which are more model independent than the simplified models examined by McCoy and Massimi, and as such present different challenges.

Finally, the discussion of the third challenge provides novel arguments that bear upon existing debates about realism in astronomy and cosmology (Hacking, 1989; Shapere, 1993; Anderl, 2016) and in biology – the dark matter case study will be illustrated by the very similar case study of genes.
Scientific Experts and the Pressures of Pandemic Policy Advice

Mathias Frisch, Katherine Furman, Axel Gelfert, Stephen John, Marion Vorms

This symposium aims to examine how scientific expertise and trust can be generated under conditions of deep uncertainty and urgency, as illustrated by the current phase of the Covid-19 pandemic. Drawing on case-studies from across Europe, the papers in the symposium focus on three inter-locking issues: first, the ways in which experts investigate, justify and communicate claims under conditions of deep uncertainty; second, the ways in which scientific testimony is conveyed and shaped by other social actors; third, the ways in which public reception of experts’ claims can be assessed and serve as epistemic resources. A distinctive feature of the symposium is the ways in which the contributors bring together a range of disciplinary perspectives, from social epistemology to philosophy of modelling to philosophy of public health, allowing for comparisons with other cases in medical science, public health and climate science. As a result, the symposium aims to draw on the current crisis to challenge and develop philosophical understandings of how the science/public-policy interface does and should affect both science and policy.

The Covid-19 Crisis required far-reaching decisions under extreme scientific uncertainty. Policymakers who wish their decisions to be informed by the best available science had to acknowledge that much of the science concerning SARS-CoV-2 was unsettled and is, even now, still changing: what is the virus’s infection mortality rate and how easily is it transmitted? What factors account for large differences in mortality rates among different countries?

This symposium investigates what, in such situations, should count as appropriate scientific advice, how expert advice should be properly credentialed and communicated, and how public debate about scientifically informed policy should be conducted. The interplay between science and policy is governed by a fundamental tension: On the one hand, scientists should provide policy advice that reflects economic, social, and political concerns. On the other hand, scientists should not import their own value judgments about what ought to be done into their expert assessments. At the same time, policymakers and the general public must decide which experts to trust, and how to properly evaluate and respond to competing scientific claims.

When there is broad consensus on what goal we want to accomplish in addition to a high degree of scientific consensus on how to accomplish it, the division of labor between scientists and policymakers seems straightforward: policy-makers set the goals and all scientists must do is to tell policymakers the best means to this end.

What makes the Covid-19 crisis a fascinating contrast case is that it involves both deep epistemic uncertainties and complex and contested values. The combination of high stakes (decisions of life and death) and fast (yet incomplete) science, makes for uncomfortable trade-offs. As the science becomes more uncertain, there are more – and more different –things scientists can reasonably say. For example, there are many different ways of modelling a disease’s progression. As such, an expert offering just one model might be accused of a dereliction of scientific duty, yet offering multiple models creates a problem for the policy-maker: which one should she use? And how should she, as non-expert, decide among them? At the same time, as values become more complex and contested, it becomes harder to determine for experts how best to provide policy-relevant information.

When the need for actionable results requires departing from established scientific procedures, we need to ask which departures are compatible with a well-ordered science that can fulfil its role of providing credible and trustworthy policy advice. How can scientific experts discharge their duty of providing policymakers with scientific information, when there is no established consensus? How do
we distinguish legitimate scientific disagreement from scientifically marginal views that may be discounted? These questions take on additional urgency in a context where researchers themselves take to, or are recruited by, the media, either by choice or because policymakers request their presence at press conferences. By contrast, scientific input regarding the climate crisis, for example, has taken the far more impersonal form of the IPCC collectively issuing reports. This new ‘personalized’ approach to policy advice, which contrasts with the IPCC’s institutionalized yet slow process of aggregating individual findings into consensus reports, places additional epistemic and political burdens on scientific experts. Predictably, this has led to accusations of partisanship.

In both climate science and pandemics, we are highly certain about some things—for example, that, in the absence of mitigation measures, climate change or Covid-19 will have largescale effects—and uncertain about many others—such as the precise effects of particular interventions. In both cases there are challenges to communicating uncertainty: first, related to how it can or should be quantified; second, related to how probability claims may be (mis-)understood by audiences; third, stemming from the risk that acknowledgment of uncertainty may be exploited by other interested actors . To some extent these challenges arise from a fundamental misunderstanding: a proposition’s “being scientifically established” is often treated as a synonym for “being certain”. However, from a philosophical perspective, the public ideal of “best science” is problematic for at least three reasons: first, philosophy of science has stressed that there is no such thing as a single, unitary “scientific method”, and that much scientific knowledge is provisional, uncertain, and contested; second, work on epistemology and communication has shown the difficulties of establishing who counts as an expert, and how failures to defer may be epistemically rational; third, political philosophers have long noted that, ultimately, all policy decisions must be based on values, rather than facts alone, and value pluralism may be irreconcilable. With the proposed symposium, we hope both to analyze and understand the source of such misunderstandings, and also to contribute to their future resolution.

The contributions to this symposium examine these topics from a variety of perspectives. John’s and Frisch’s contributions focus on the role of scientific experts in assessing and communicating their findings. Frisch asks what epistemic credentials uncertain evidence needs to possess in situations of extreme urgency, while John argues that scientist need to properly ‘curate’ their findings when communicating uncertain evidence. Gelfert draws attention to an often-overlooked participant in the communication of scientific opinions: science journalism and popular science, which can play a mediating role between accounts of scientific findings produced for other experts and the public. Vorms and Furman, finally, focus on policy-makers and the wider public, both as recipients of findings communicated by experts, who need to be able to distinguish reasonable doubt from undue scepticism, and as potential source of local knowledge. As such, the contributions draw on and contribute to debates over a wide range of topics in philosophy of science and allied disciplines, including philosophy of climate science, philosophy of modelling, social epistemology, and, more broadly, debates over the feasibility and desirability of “value free” science, all with the aim of furthering and improving real-world debates and practice.
SU(n) Surplusage: Symmetry, Gauge, and Equivalence

Neil Dewar, John Dougherty, Henrique Gomes, Tushar Menon, Sarita Rosenstock

It is widely held that there is an intimate relationship between the notions of symmetry, surplus structure, theoretical equivalence, and gauge freedom. This symposium brings together some of the most exciting recent work on this relationship, to explore what these interconnections can tell us about the nature of physical representation. Dewar considers the relationship between symmetry and equivalence in the context of quantum statistical mechanics, and their implications for how to represent phase transitions. Dougherty argues that the standard understanding of the relationship between symmetry and surplus structure is untenable, by showing how it leads to an inconsistency in the interpretation of quantum field theory. Gomes analyses how the intertwining of “gauge” and physical data on boundaries manifests a subtle and characteristic form of non-locality. Menon shows how the debate between different approaches to the interpretation of symmetries is illuminated by application to supersymmetric theories. Rosenstock exhibits a way of using category theory to capture the representational content of a theory, without falling into the trap of identifying a theory with a category.

A central topic in philosophy of physics (and in philosophy of science more generally) is the distinction between representational and surplus structure in our theories: how are we to determine the parts of our theories that do representational work, and those which are mere mathematical artefacts? And how do we use this identification to (in Twain’s phrase) “eschew surplusage”?

One popular way of identifying surplus structure has been the so-called method of symmetry (Saunders 2003; Dasgupta, 2016): theoretical structure can only be taken to be physically significant if it is symmetry-invariant. The application of this method has led to an explosion of interest in the metaphysics of symmetries and has helped shed a great deal of light on the structure of our best theories. Increasingly, however, there is an awareness that the application of this method is not as straightforward as one might have thought; and moreover, that merely identifying surplus structure may not guarantee that one knows how to get rid of it.

In this symposium, we explore the subtleties and limitations of this method. On the one hand, we will consider the subtleties that arise in applying it to cutting-edge physical theories: gauge theory (Gomes), supersymmetry (Menon), quantum field theory (Dougherty) and quantum statistical mechanics (Dewar). On the other, we will consider how the method relates to broader projects in philosophy of science, such as the detection of surplus structure (Dougherty and Gomes), category-theoretic representations of theories (Rosenstock), intertheoretic translation (Dewar), or the metaphysics of physics (Menon). We therefore anticipate that this symposium will demonstrate both the relevance of philosophical debates to scientific practice, and the importance of case studies for general philosophical questions.

Much of what emerges from the work in this symposium is a sense that many understandings of symmetry and surplus are not appropriate. Dougherty’s contribution, for example, argues that the currently extant understanding of this method—i.e., the identification of surplus structure with symmetry-variant structure—is not fit for purpose when applied to case studies in quantum field theory. In a similar vein, Menon suggests that when we determine that a theory has symmetries, requiring a “metaphysically perspicuous” reformulation of it (in such a way that the symmetry-variant structure is expunged) is too strong a requirement: in supersymmetric theories, he argues, we can have no guarantee that such a reformulation is to be found; and Gomes argues that keeping a gauge-variant part of the theory is a reasonable price to pay for locality. Rosenstock, too, argues
against supposing that some particular formulation of a theory—such as its formulation as a category of models—will capture everything that is physically relevant about a theory.

However, this symposium also shows how the analysis of symmetries can be very insightful for foundational work. Notwithstanding the above, Rosenstock also shows that a certain kind of category can play a privileged role in determining a theory’s representational content: in part, because of how it helps to facilitate comparisons of structural content within and between theories, such as those revealed by symmetries and equivalences. Gomes shows how the familiar trade-off between nonlocality and redundancy in gauge theories arises from the intermingling of physical and nonphysical data on the boundaries of regions in spacetime. Dewar considers the implications of broken symmetry for quantum statistical mechanics, and for our understanding of the nature of phase transitions.

Overall, then, this symposium will shed light on an area of high philosophical interest, through a variety of different perspectives.
The suggested symposium focuses on symmetry principles: conceptual or formal postulates that constrain the content or the interpretation of physical theories or guide their formulation. We will present three new perspectives on symmetry principles, bringing together considerations from metaphysics, philosophy of science and the philosophy of physics. All three papers support the view that a unified account of spacetime symmetries and more abstract kinds of symmetries is both possible and desirable. The first paper shows that Earman’s principles for spacetime symmetries can be naturally extended to internal symmetries. This extension provides internal symmetries with a parallel role to external symmetries in the determination of theoretical structure. The second paper critically examines the claim by Belot that the notion of equivalence cannot be used to provide a unified understanding of different symmetries. The third paper provides a new account for the role of the principle of general covariance to the understanding of gravity and spacetime structure, and contrasts the principle with the approach of gauge theories of gravity. The symposium aims to promote further consideration of conceptual issues that hold a promise for better interpretations of current physical theories as well as for the development of future physics.

Symmetry considerations are fundamental to foundational questions in modern theoretical physics, both providing the basis for the formulation of the laws in central contemporary theories, and functioning as a major tool for the interpretation of these theories. Both of these roles were first manifested in the context of spacetime symmetries (see, for instance, Friedman (1983); Earman (1989)). These discussions later served as a model for the extension of symmetry considerations to other symmetry transformations, including ‘internal’ degrees of freedom, such as the gauge freedom of electrodynamics. On the one hand, the methodological role of symmetry arguments in the formulation of the laws of physics is rooted in the role these considerations played in relation to spacetime symmetries in the development of the general theory of relativity. These considerations gave rise to Noether’s theorems and to the first formulation of the gauge principle by Herman Weyl, both of which became essential to future development of physics, providing, inter alia, the theoretical framework for the standard model of particle physics. On the other hand, the practice of interpreting the ontology of physical theories through their symmetries goes back to the debates over the nature of space and time in the context of Newtonian mechanics, and is to a great extent based on the notion that states that are connected by a symmetry transformation are physically equivalent. Symmetry considerations are commonly formulated in terms of symmetry principles: conceptual or formal postulates that constrain the content or the interpretation of physical theories and/or guide their formulation. For instance, Ismael and Van Fraassen (2003) describe symmetries as ‘a guide towards superfluous theoretical structure’. But despite the historical role of the analogy between spacetime symmetries and other types of symmetries, it is commonly believed that spatiotemporal symmetries stand out as special, for physical (Wigner 1964) as well as philosophical (Belot 2013) reasons. The suggested symposium will present three new perspectives on symmetry principles, bringing together considerations from metaphysics, philosophy of science and the philosophy of physics. Despite the differences in our approaches, all three papers support the view that a unified account of different kinds of symmetries is both possible and desirable.

In the context of field theories, spacetime symmetries are divorced from other types of symmetries via the external/internal distinction (i.e. the theory’s independent vs. its dependent variables). In the context of the former, Earman (1989) has famously shown how various symmetry arguments manifest two symmetry principles. Together, these principles aim to ensure that a theory has neither too much nor too little spacetime structure. The first paper of this symposium shows that these principles can be naturally extended to internal symmetries, for instance gauge transformations. This
extension provides internal symmetries with a parallel role to external symmetries in the determination of theoretical structure. In a slogan, symmetries act as a guide towards the structure of physical quantities.

Motivated by Earman’s (1989) discussion of spacetime symmetries, Belot (2013) introduced what later (Wallace 2019a) became to be known as the representational thesis (RT), a general symmetry principle that is meant to apply to all kinds of symmetries (i.e. not only spatiotemporal ones): Two solutions of a classical theory's equation of motion are related by a symmetry if and only if they are physically equivalent, in the sense that they are equally well- or ill-suited to represent any particular physical situation. Although this principle (or some version of it) has been widely adopted in the philosophy of symmetries, as Belot points out, he presents several examples where it seems to fail. In particular, in the general context of the symmetries of differential equations, Belot shows that in cases such as the harmonic oscillator or the heat equation, there are symmetries of the differential equations characterizing these systems that relate solutions which are representing physically inequivalent states of the system. The second talk critically examines Belot’s analysis of these examples and shows that they do not succeed once we are explicit about how the solutions of the differential equations actually model the dynamics of physical systems. More generally, the talk shows that when asking questions about the interpretation of symmetries, including principles such as (RT), we need to be particularly sensitive to questions about how we characterize the subsystem that we are attempting to model, what kind of environment we consider, and what kind of boundary conditions we impose.

The last paper concerns two approaches to the question of the relationship between spacetime symmetries and the gravitational force. The first is based on the principle of general covariance that guided Einstein in the development of the general theory of relativity, and is the basis for Earman’s second symmetry principle. The principle states that the dynamics should be invariant under general arbitrary coordinate transformation. Despite its role in the development of the theory, it was criticized immediately after the introduction of the theory for being devoid of physical content, and it remained the subject of ongoing controversy (Norton 1993, Pooley 2010). The second approach is that of gauge theories of gravity (Hehl 1995). This approach aims at a unification of gravitation with the other forces by treating internal and external symmetries on a par using the concept of localization of global symmetries that commonly describes internal symmetries. The latter approach leads to rich possibilities for spatiotemporal structure and is considered as a promising path towards quantum gravity. The paper shows that the principle of general covariance, understood in light of recent developments in the philosophy of symmetries, can lead to a clear and minimalist derivation of the exact same possibilities of spacetime structure as the gauge approach.

We hope that the symposium will promote further consideration of the use of these and other symmetry principles in physics. We aim to show that despite recent advances in the debate around symmetries, there is still much room to increase their scope beyond known applications. Although our case studies focus on current theories, their application is not limited to any particular theory. Thus, symmetry principles of the sort we discuss hold a promise for both the development and the interpretation of future physics.
What We Have Learned about Memory (and What Remains to be Learned)

David Colaco, Ali Boyle, Michael Levin, Sarah Robins

Learning and memory have long been the research core of cellular and molecular biology. This symposium will address challenges and opportunities that new research presents for theoretical conceptions of memory, its function, and its biological basis. In the past ten years, novel technologies have led to discoveries regarding the memory engram. Concurrent with these discoveries about the engram, researchers have discovered mechanisms that share properties with memory mechanisms, where these novel mechanisms often underwrite diverse neuronal (and even non-neuronal) phenomena. Further research challenges not only the mechanisms that underwrite memory, but our understanding of which organisms have memory, and in what form. Together, these discoveries suggest that memory (if this is the right word to use) operates in a manner that may depart from our intuitive, scientific, and philosophical conceptions.

This symposium will address challenges and opportunities that new research presents for theoretical conceptions of memory, its function, and its biological basis. In the past ten years, novel technologies have led to provocative (and sometimes inconsistent) discoveries regarding the memory engram (Asok et al. 2019; Josselyn and Tonegawa 2020; Langille and Gallistel 2020). Concurrent with these discoveries about the engram, researchers have purported to discover mechanisms that share properties with paradigmatic memory mechanisms in humans and other organisms, where these novel mechanisms often underwrite diverse neuronal (and even nonneuronal) phenomena (Shomrat and Levin 2013; Bédécarrats et al. 2018; Posner et al. 2019). Further research challenges not only the mechanisms that underwrite memory, but our understanding of which organisms have memory, and in what form. Together, these controversial discoveries suggest that memory (if this is the right word to use) may operate in a manner that departs significantly from our intuitive, scientific, and philosophical conceptions.

While these discoveries are no doubt attention-grabbing, their implications for our understanding of memory remain unclear. Though the biological bases of several varieties of memory have been topics of inquiry for decades, these research advances call into question whether memory substrates and mechanisms are compatible with scientific and philosophical conceptions of memory and their traces in the brain. At the same time, we should not assume that our concepts must be outright reengineered in light of these discoveries. Philosophical analysis of novel research is needed to determine how (or perhaps even whether) these discoveries ought to have an effect on how we think about memory, our functional analyses of it, and our means of explaining it.

The challenges and opportunities that result from these implications are well exemplified by growing concerns surrounding our conception of memory traces. Supported by the advent of optogenetics (Robins 2018), studies of the engram have produced results that call into question the stability and function of synaptic processes thought to underwrite memory consolidation. With these insights, there is increasing debate of the role of molecular details in underwriting this consolidation (Langille and Gallistel 2020). This research contributes to a continued debate about whether memories are maintained over time (De Brigard 2014), but it also calls into question what the target should be for a substrate of memory in the brain and other biological systems. Together, these questions tie fundamentally to how we ought to conceive of what is needed to underwrite memory processes.

Likewise, the debate over potential synaptic and molecular substrates of memory has affected our conceptions of what systems exhibit properties that warrant being called “memory.” New research suggests that signaling phenomena bare more than a passing resemblance to memory functions and mechanisms. However, what should we infer from these similarities? Pushes from philosophers
(Colaço forthcoming) and scientists (Baluška and Levin 2016) to compare paradigmatic memory phenomena with phenomena that occur in biological systems that do not even have a brain suggest that the traditional categories of memory, supported by both scientific and philosophical conceptions, must be reevaluated in light of novel discoveries. This reevaluation is controversial, as it not only calls into question scientific and philosophical accounts; it tests the limits of our folk conception of what counts as a memory as well.

While this symposium focuses on the philosophical and scientific upshots of the novel discoveries regarding memory, its function, and its biological basis, the topics we address have implications for debates about categorization, conceptualization, animal models, and cognitive ontology as well. The study of memory provides fertile grounds for analyzing how our cognitive kind concepts ought to inform and be informed by ongoing and disputed scientific investigations of these kinds. Importantly, the philosophical dimension of this debate is not passive. Rather, the scientific advancements can only be supported with rigorous analysis and engineering of cognitive concepts like memory, which is precisely the aim of this symposium.

References:


