Workshop

Explanation and Understanding

Book of abstracts

Date: 23-25 May 2018

Venue: Royal Academy of Dutch Language and Literature (KANTL)
Koningstraat 18 Ghent, Belgium

Organization:

Centre for Logic and Philosophy of Science

(Ghent University)

the Research Group Philosophy of Science and Technology (VU University Amsterdam)

Chairs

Henk de Regt (Amsterdam), Dingmar van Eck (Ghent), Erik Weber (Ghent).

Local Organizing Committee

Roxan Degeyter, Stef Frijters, Julie Mennes, Dingmar van Eck, Erik Weber, Dietlinde Wouters, Pawel Pawlowski.

About the workshop

This workshop is organized within the series "Logic, Reasoning, and Rationality" :

`http://www.lrr.ugent.be/`

The series is supported by the Research Foundation Flanders (`http://www.fwo.be/`) through the scientific research network on Logical and Methodological Analysis of Scientific Reasoning Processes (LMASRP):

`http://www.lmasrp.ugent.be/`

fwo Research Foundation Flanders Opening new horizons

Abstracts

Invited speakers

Matteo Colombo

I know that I know nothing. Explanation, Prejudice, and Intellectual Humility

People are prejudiced towards members of groups with a worldview they perceive to be dissimilar from their own. The relationship between perceived dissimilarity and prejudice is so stable that it has been referred to as a psychological law. In this talk, I shall bring together ideas and methods from existing literatures on explanation and intellectual humility, and discuss how people's intellectually humble explanatory reasoning might impact the relationship between dissimilarity and prejudice.

Caterina Marchionni

Explanatory norms as frictions to integration: the case of economics and its neighbours

By looking at the relationship between economics and neighbouring fields, I examine the way in which field-specific norms about explanation hinder the integration of mechanistic models across fields. I conclude by arguing that the mechanism-based unity of science championed by Craver and other mechanistic philosophers is better captured by the image of a cubist painting than that of a mosaic

Alexander Reutlinger

Understanding and Non-Causal Explanation

According to several accounts of understanding, understanding a phenomenon requires having an explanation of that phenomenon. This requirement faces a challenge from the existence of causal and non-causal explanations: whether there is a unified account of understanding seems to depend on whether there is a unified account of causal and non-causal explanations. If one adopts the standard causal account of explanation, then, clearly, there is no unified account of explanation - and, hence, of understanding - because causal accounts do not capture non-causal ways of explaining. I will argue that there is a unified (or, as I call it, monist) account of causal and non-causal explanations: the counterfactual theory of explanation. I will suggest that this theory of explanation could be a fruitful building block for a unified (or monist) view of causal and non-causal modes of understanding.

Contributed Talks

Roxan Degeyter and Erik Weber

Explanation and understanding in population genetics

Population genetics is described as "[...] the study of the patterns of genetic variation at the population level and how changes in these patterns that result from evolutionary forces bring about evolution over time. (Erlod and Stansfield 2010, p. 248). Population genetics is thus about causal explanations. Our aim is to explicate the structure of explanations in population genetics and propose several standard formats in which the causally relevant elements are clearly indicated and structured. Such standard formats may facilitate communication and discussion between geneticists.

We observe two kinds of explanations in population genetics. The first are explanations regarding an evolution over time. Often-cited examples are the evolution of moths from grey and speckled to dark melanic and the evolution of the fur colour of beach mice from brown to white. These evolutions can be explained by an earlier change in environment, which is causally relevant for the explanandum change, provided a precondition for the causal relationship is fulfilled, namely the presence of visually oriented predators. This precondition is a form of natural selection. Natural selection is thus a precondition for the existence of a causal relationship between a change in environment and a change in visual characteristics of a population.

The same thing can be observed in the second kind of explanations. These deal with synchronic differences, such as the brown and pink versus yellow colours of land snails or the striped versus unstriped patterns of walking sticks. These synchronic differences can be explained by reference to a difference in environment, which is causally relevant for the difference in colour, provided that a precondition for the causal relationship is fulfilled, namely the presence of visually oriented predators. Again, this precondition is a form of natural selection, although of a more hypothetical nature than in the evolutionary examples. In these explanations of synchronic differences, natural selection is again a precondition for a causal relationship.

In our talk we will present two of these examples. The example of the evolution of fur colour in beach mice will clarify the structure of explanations of evolutions over time and the example of different patterns in walking sticks will clarify the structure of explanations of synchronic differences between subpopulations. In both cases we will show that understanding is generated by means of an explanation that cites causes and, on top of that, tells us something about the underlying mechanism that produces the relevant causal relations that are used in the explanation.

Sven Delarivière

Explicating understanding: a conceptual framework and functionalist approach

My aim is to present a conceptual framework to explore the notion of understanding and strengthen a functional approach to it in particular. Considering the scarce and disjointed nature of the present literature, it will be fruitful to string together claims of epistemology (and other relevant fields) in a clear and systematic framework where we can keep track of where particular claims fit in and what they have bearing on. To begin, I present a conceptual framework intended to cover the epistemologically relevant aspects of the concept "understanding". That concept concerns a subject (S) possessing a particular property (P) concerning an object (X). Each requires the same philosophical steps of investigation: (i) characterisation, which requires us to define, demarcate or otherwise characterise the aspect in question, (ii) an analysis of that interpretation and its entailments, and (iii) extracting the parameters, which together track the dimensions of quality in understanding. To ensure that an account of understanding is consistent, each branch also requires the same philosophical approach to each of its steps. The functionalist approach, which I shall defend, consists of the following: First, the (P) p roperty of understanding is

(i) characterised as abilities, (ii) analysed in terms of (counter)factual performance, and (iii) qualitatively expressed through their stability (involving the performance's range within, and robustness through, counterfactuals). Second, the (X) o bject of understanding is (i) characterised through the several appropriate usages, (ii) analysed as what is valued and deemed appropriate by a practice, and (iii) qualitatively expressed through its scope (involving the width, sensitivity and accuracy of performances). And third, the (S) s ubject with understanding is (i) characterised by the structural integrity of the system implementing the performances, (ii) analysed as either a system or an epistemic agent (two related concepts which shouldn't be conflated), and (iii) qualitatively expressed through its system efficacy (involving resource economy and growing potential). The strongest danger for a functionalist approach would be cases of abilities w ithout understanding, of which the literature has proposed several examples. But with the benefit of the framework, each can be classified as either a case of (a) misestimated abilities or (b) subject-misattribution. The functionalist approach to the object allows us to carve up different kinds and degrees of understanding that object. So some examples (e.g. algorithmic understanding, verbatim memorisation, idealisation and luck) should be considered as misleading in that they make us (a) misestimate the abilities present. Under the framework, this can be expressed through a lack in scope-parameters. These parameters are a stronger and more flexible ally to conceptualise understanding than necessary or sufficient conditions (which make the problem of luck especially difficult for epistemology). Furthermore, the functionalist approach to the subject opens up the possibility of extended, collective or artificial understanders. So the problem in other examples (e.g. The Chinese Room, the barometer-using weatherman or a collaborating community) is (b) subject-misattribution, which can be dissolved by reconsidering the system or agent attributed with understanding.

Leen De Vreese

Risk factors, explanation and scientific understanding

The notion risk factor is omnipresent in contemporary medical research, medical practice (e.g. prevention campaigns) and layman understanding of health and disease. This is a recent phenomenon, in the sense that it started in the 1950s. During the second half of the 20th century and the first decade of the 21st century, talk in terms of risk and risk factors has become ever more pervasive. Nevertheless, the work of medical scientists and sociologists of medicine shows that there is no consensus about how the term is best used. In general, four different meanings of the notion "risk factor" can be discerned in the literature:

1. Risk factor_0 = any factor associated with the development of a given disease.
2. Risk factor_1 = risk factor_0 considered to be a cause of the disease.
3. Risk factor_2 = risk factor_0 of which it is not known whether they are a cause of the disease or not.
4. Risk factor_3 = risk factors_0 thought not to be a cause of the disease.

In my talk I will use this distinction as a basis for my analysis of whether and how risk factors can explain a disease and whether and how they provide scientific understanding. Getting clear on this is important to evaluate the importance of risk factor knowledge. Given that causal factors are generally taken to have explanatory power, it seems uncontroversial to claim that type 1 risk factors explain and thereby provide scientific understanding. The interesting question is whether and how this extends to type 2 and 3 risk factors. Do they explain? If so, in what sense? And what do they explain? Additionally, do they provide scientific understanding (with or without explanation)? And again: if so, in what sense?

As a starting point of my analysis, I will take the possibility that non-causal risk factors somehow explain seriously. I will do that by shifting my attention away from (causally) explaining the onset of a disease to explaining differences in chances. Understanding why person a has a higher chance of getting breast cancer than person b may require that we have knowledge about probabilistic dependency relations in the world, without these relations being causal. In my talk, I will explore whether taking this route helps us further in getting a grasp on how non-causal risk factors can explain and/or provide scientific understanding.

Armond Duwell

Modal Understanding: The Real Deal

This paper has two aims. One aim is to argue against the view that explanations of phenomena are essential to understanding them, as suggested by the received view associated with Strevens (2013), Grimm (2014), or Khalifa (2017), among others. The second is to argue that the modal view of understanding phenomena developed in Le Bihan (2017) and Duwell (forthcoming), a non-explationist view of understanding, doesn't succumb to objections that can be leveled at it.

While it is widely recognized that there are objects of understanding, e.g. propositions, that can be understood without explanations, it is not widely recognized that one can understand phenomena without explanations. Why? I suggest that it is because we conflate two distinct concepts of phenomena: surface and hidden phenomena, as distinguished by Feest (2011). Surface phenomena are patterns in data; hidden phenomena are taken to be the cause or ground of those patterns. If we distinguish surface and hidden phenomena, we are not tempted to associate understanding phenomena with their cause or ground, and hence explanations. In Section 3, I argue that explanationist understanding of hidden quantum phenomena is impossible to come by if that theory is fundamental, yet interpretations of quantum mechanics indicate otherwise. I also discuss recent work on quantum surface correlation phenomena where physicists are explicitly looking for understanding of surface phenomena independently of hidden phenomena. These serve as counterexamples to the received view. In Section 4, I describe the modal view of understanding and show how it accommodates these cases. In Section 5, I will discuss and apply Khalifa's (2013) strategy for turning purported counterexamples to explanationism regarding understanding into positive cases and show that it fails in the case of surface and hidden quantum phenomena. In Section 6, I discuss the completeness objection to non-explanationist views that claims that understanding of phenomena is complete when one grasps their explanations. That is at odds with scientific practice. In Section 7, I discuss an objection to the modal view which claims that understanding is too ubiquitous on the view, and argue it is not. I conclude that non-explanationist views of understanding, and the modal view in particular, deserve serious consideration.

Johannes Findl

The relationship between scientific understanding and explanation

The nature of understanding has attracted growing interest in both epistemology and philosophy of science. While in the latter, the focal point of the debate is the relationship between scientific understanding and explanation, in the former, it is the relationship between understanding and knowledge. In this talk, I will be primarily concerned with the debate in philosophy of science, and in particular my goal will be to provide some reasons to reject a number of influential views according to which scientific understanding has always to pass by the route of scientific explanations. In other words, I will claim that scientific understanding may sometimes come apart from scientific explanations. This conclusion, however, will also essentially depend on premises which gain support from a number of plausible considerations that have originated in the epistemological debate on understanding. I will proceed as follows: After disambiguating the term understanding and restricting myself to its use in relation to a why-locution in scientific contexts (henceforth: understanding-why), I will briefly argue that the intuitively plausible attempt to define understanding-why via identifying it with having an explanation is flawed because of unduly oversimplification. Moreover, I will also talk about the Hempelian view on understanding as rational expectation. I will argue that the Hempelian view is ambiguous between two natural yet problematic readings, which gives us a strong reason to reject it. I will then describe Michael Strevens' much improved and influential "the simple view", according to which understanding-why P is defined in terms of grasping a correct explanation of P, where an explanation is an explicitly expressed set of propositions. Importantly, Strevens thinks that "there is no understanding without explanation", and I take his view to be the strongest contender in the current discussion. Before I will present my own objections against the simple view, I will scrutinize Peter Lipton's and Daniel Wilkenfeld's and Jennifer Hellmann's objections to it. Lipton rejects Strevens' view since he thinks that to understand-why is to have cognitive benefits such as knowledge of causes, of necessity, of unification and of possibility. While Lipton agrees that explanations often do provide these benefits, he argues that in modelling, scientists obtain tacit causal knowledge of a phenomenon in the absence of its scientific explanation. Likewise, he says that we obtain knowledge of necessity in Galileo's famous thought experiment (i.e. the independence of acceleration from gravity) when there is no explanation available. Against Lipton, I will claim that having tacit knowledge of a phenomenon is not sufficient for understanding why this phenomenon was the case, and that the Galilean thought experiment only works if we accept some empirical assumptions which themselves form part of correct explanations. By contrast, Wilkenfeld and Hellmann object to Strevens that understanding-why is to a large extent non-propositional in nature, and therefore cannot be the mere product of propositional explanations. While I think that their objection is ingenious since it rightly emphasizes that understanding is an ability that cannot be reduced to a set of propositional instructions, I will argue that it is ultimately not successful. However, this has less to do with their general claims about understanding as an ability than with the very examples they use to illustrate them. Finally, in the last part of my talk, I will argue that the historical case of the water channel developed by the German Ludwig Prandtl demonstrates the possibility that a scientist may gain genuine understanding why of a phenomenon before any correct scientific explanation of it is available. The water test channel is a physical model which Prandtl once developed in order to visualize basic flow experiments, and, running these experiments, he was later able to find missing equations for Fluid Mechanics and Aerodynamics. I will claim pace Wilkenfeld that this example shows both that understanding is ability that is irreducible to propositional instructions, and that Prandtl's underlying thought processes are best accounted for in terms of beliefs with less than full credence. In a nutshell, my goal is to demonstrate that the water channel experiment is a case where understanding why has actually come apart from explanation.

Joachim Frans

The ontic-epistemic distinction and the study of explanatory proofs

Philosophy has a long and lively tradition of analysing the nature of explanation. The majority of this attention is devoted to explicating the notion of scientific explanation. When Hempel and Oppenheim (1948) set the stage for current debates concerning scientific explanation, the notion of mathematical explanation was initially left aside. This has changed in last decades, and mathematical explanation is now considered as a legitimate topic in analytic philosophy. The study of mathematical explanation entails two different ways in which mathematics might provide an explanation. The first sense is extra-mathematical explanation, where mathematics plays an essential role in the explanations of natural or social sciences. The second sense pertains to explanations within mathematics itself. Within the latter sense, there is a modest tradition of investigating differences between explanatory and non-explanatory proofs. An influential model to draw this distinction is due to Steiner (1978). After Steiner, several authors have proposed revised versions of his model (e.g. Weber and Verhoeven 2002; Salverda 2017). Others proposed alternative models (e.g. Lange 2014, Pincock 2015). My talk will concern the specific notion of an explanatory proof, but will also look at the literature on scientific explanation. Given the richness of the literature concerning scientific explanations, it is possible to see what insights can be gained for the analysis of mathematical explanation. A specific aspect of debates concerning the nature of scientific explanations can be fleshed out in terms of a tension between ontic and epistemic conceptions of explanation. This terminology was introduced by Salmon (1984), but the terms can evoke different things in specific debates. One way of making sense of this tension is to state that ontic conceptions of explanations define explanation independently from categories (such as understanding) linked with a subject, while epistemic conceptions define explanation through categories (such as understanding) that are linked with a subject. In this talk, I look at the ontic-epistemic distinction with respect to the notion of explanatory proofs. Most literature does not look at mathematical explanation in these terms (Pincock 2014 and Delariviere, Frans and Van Kerkhove 2017 are exceptions), as the relation between explanation and understanding is often not addressed explicitly. My aim for this talk is threefold. I will argue that: 1. It can be meaningful to address the notion of explanatory proofs in terms of an ontic-epistemic conception.This calls for a weak reading of the ontic conception, as Kaiser (2015) proposes in her work on explanations in biology. 2. Models of mathematical explanation describe both ontic and epistemic aspects of explanation. This perspective is influenced by Illari's (2013) work on integrating ontic and epistemic constraints on mechanistic explanations. 3. Introducing the views from (1) and (2) is fruitful. It leads to an improved setting, where new insights concerning the literature on explanatory proofs can be presented

Daniel Kostic

Minimal structure explanations, scientific understanding and explanatory depth

I outline a heuristic for thinking about the relation between explanation and understanding that can be used to capture various levels of "intimacy" between them, i.e. by using this heuristic we will be able to explain away some of the seemingly paradoxical cases in which there could be the understanding without explanation, as well as cases where there can't be understanding without explanation. The idea is that the level of complexity in the structure of explanation is inversely proportional to the level of intimacy between explanation and understanding, i.e. the more complexity the less intimacy, and vice versa. The structure of explanation should be understood as a description of the exact relation between the explanans and explanandum, and the complexity in this context should be understood as the number of components that are required to describe this relation. In this sense, it means the more components the more complex the structure of explanation, and vice versa. I further argue that the level of complexity in the structure of explanation also affects the explanatory depth in a similar way to intimacy between explanation and understanding, i.e. the less complexity the greater explanatory depth and vice versa. To avoid circularity when using the terms "grasping" and "understanding" in referring to the structure of explanation, following Strevens (2008, 2013) and Khalifa (2017) I distinguish between "understanding-that" and "understanding-why". Understanding-that refers to some basic cognitive abilities such as being a competent speaker of a language, knowing what certain mathematical relations mean, grasping the mathematical axioms and knowing what it means to say that they are logically primitive, or knowing that something is a fact. For example, in the D-N model there is also the understanding-that of the rules of inference, order of derivation, validity and soundness. The understanding-why refers to knowledge of why something is the case, which is based on the knowledge of counterfactuals. Another way to put it is that the understanding-why comes from the structure of explanation, and it has a form of counterfactual information about the dependency relations between the explanans and explanandum. The minimal structure explanations also support an account of explanatory depth, that can be applied to both causal and non-causal explanations. The explanatory depth should be understood in terms of richness of counterfactual explanatory relations that the explanation provides, so in this sense, the explanations which provide fewer counterfactual explanatory relations are less deep than the ones that provide more counterfactual relations. Depending on the complexity of the structure of explanation, the relation between explanation and understanding can be more intimate or less intimate, the more complex the structure of explanation the less intimate the relation between the explanation and understating, and vice versa. Because of the minimal structure and more direct relation between explanation and understanding, these explanations will be deeper, and more universal, because they will provide more counterfactual dependency relations for our grasping.

Insa Lawler

Understanding in the case of holistically distorted models and how-possibly models

There is a variety of models in science that feature idealizations that are utterly false, such as the assumption that the number of particles in a fluid approaches infinity. One popular divide is between models that have real-world target objects and models that have hypothetical target objects (short: how-possibly models). For instance, phase-transition models are considered to model realworld phase transitions. But the famous Schelling model of segregation is not considered to model real-world segregation; it models segregation in a hypothetical town. The different kinds of models are claimed to provide us with different (possible) results: how-actually explanations in the first case, and how-possibly explanations in the second case (e.g., Grune-Yanoff 2013; Bokulich 2014; Reutlinger et al. 2017). In my talk, I argue that this divide blurs the commonalities of such models when it comes to understanding. My focus lies on a comparison of what Rice calls 'holistically distorted models' (2017) and how-possibly models. Holistically distorted models are models that involve idealizations that are ineliminable, i.e., they involve idealizations that are essential to the model's mathematical representation of the target phenomenon (Rohwer and Rice 2016, p. 1134). I argue for a unified account of understanding in the case of holistically distorted models and how-possibly models. By means of examples, I argue that the following holds for both kinds of models: The aim of working with such models is to gain understanding pertinent to a real-word phenomenon. If there were so-called how-possibly understanding, it would be gained employing both kinds of models. Efficient constructions of such models are based on understanding what features of the involved entities might make a difference to the real-world phenomenon or which do not. This understanding is the basis for educated guesses that are crucial to construing the models. Only in the case of successful models understanding can be gained. If understanding pertinent to the realworld phenomenon is gained, the propositional content of the accounts that comprise the understanding are not identical to (a part of) the propositional content of the respective models that involves the utterly false idealizations. The models are not the explanations. Based on work by Alexandrova (2008), Pincock (2014), and Rice (2017), I argue that it is the information we extract from such models that figures into our understanding pertinent to the real-word phenomenon. Yet, understanding can only be gained if the extracted information is considered to be tenable in light of further evidence. Grasping true information is not sufficient for understanding. Such further evidence might be independent reasons or the justification of the idealizations involved in the respective models. In a nutshell, although holistically distorted models and how-possibly models differ crucially in their characteristics, there is no relevant difference regarding understanding.

Federica Malfatti

On the possibility of understanding without (true) explanation

Is it possible to gain understanding about a certain phenomenon P without having an explanation for P? Is there a path leading to understanding that does not pass through an explanation? In this paper, developing the view endorsed by De Regt and Gijsbers in their "How false theories can yield genuine understanding" (2017), I argue that, although understanding and explaining are closely related to one another, they might come apart at least in the sense that it is possible for a subject S to gain understanding of a real phenomenon P by means of an explanation E (meant to be) about P that is utterly false, i.e. that is not even partially correct. More specifically, I argue that an utterly false explanation of P can provide a certain subject S with understanding of P, granted that: (i) the explanation is intelligible for S; (ii) the explanation fits, in some measure, with what S already believes or endorses about reality, and (iii) S has good reasons to judge the explanation to be true, given the epistemic circumstances she/he happens to live in. Understanding, hence, has an internalistic, holistic and social component that explanation, instead, seems to lack – or so I will argue. Understanding, in my view, is subjected to:

1. an intelligibility constraint;
2. a holistic constraint;
3. a social constraint.

When we judge whether an explanation is adequate or not, we are apt to ask ourselves whether the explanation satisfies a certain semantic constraint: is the explanation true, at least partially so? Does the explanation depict real relations? Is there a correspondence between explanatory structure and a real, dependency structure? When we find ourselves judging whether somebody understands or has understood a phenomenon or not, or whether a certain (tentative) explanation provides one with understanding or not, we are apt to ask ourselves:

- Is the explanation properly embedded in the subject's web of cognitive attitudes, that is, is the explanation intelligible for her/him?
- Is the explanation reasonable, relative to the subject's web of cognitive attitudes, that is, relative to what the subject already believes or endorses about reality?
- Does the subject have good reasons to believe the explanation to be true, or to endorse it, relative to the best of her/his knowledge and relative to the epistemic circumstances?

If the answer to all these three questions is yes, I will argue, we can reasonably attribute understanding to the subject. If I am right, this tells us, roughly, that while explanation answers to a semantic constraint, understanding, instead, answers to a purely epistemological one. A false explanation, hence, can provide one with understanding, granted that the epistemic situation one happens to live in provides her/him with good reasons to believe that the explanation is true.

Interestingly enough, this leaves open the possibility of a true explanation E not providing S with understanding, for example because:

- E is not intelligible to S;
- E does not fit into S's web of cognitive attitudes;
- S fails to have reasons to believe or to endorse E.

Conclusion: While it is an open question whether understanding is possible without an explanation tout court, it seems to be highly plausible that understanding can be gained by means of an explanation that is not correct – not even partially so.

Paul Roth

The Structure of Structure: How Kuhn Establishes that Science Requires Historical Explanation

As is well known, Kuhn restricts a designation of "normal science" to those disciplines with accepted research practices. What makes for normal science, of course, shifts with changes in paradigms on Kuhn's account. Now this way of specifying normal science has a whiff of circularity inasmuch as it defines normal science by reference to "scientific research," but that can be overlooked. Sufficient for my purpose will be to take as a 'science' whatever comes to pass as such. In this respect, given the century old controversy regarding history's status as a science, I propose focusing rather on the question of how whatever passes as "normal science" comes to achieve that status. My argument will be that any answer to a question about how normal science comes to be, i.e., one that develops a non-a priori causal/explanatory account, will have to utilize what I term an "essentially narrative explanation." In other words, my account shows how in SSR Kuhn crafts a narrativized account of normal science. This will count as naturalistic in a minimalist sense inasmuch as it does not begin with any philosophical definition of what is or is not a science, and utilizes in its explanation nothing more than facts narratively ordered so as to explain (in the sense of revealing how a later point time results from earlier ones) how what comes to be called science achieves that status. Understanding Kuhn's work in this way helps naturalize narrative explanation through a form of mutual containment—since narrative helps constitute any understanding of what counts as normal science, that narrative becomes a part of any account that comes to be viewed as science. It would be highly ironic then to reject an explanation form that in fact proves unavoidable for purposes of revealing why what passes as science at a particular time does so.

Luana Poliseli

Explanation and scientific understanding: lessons from a study in pollination service modeling

The notion of explanation has been widely discussed in philosophy of science. However, only recently scientific understanding has become an important target to be addressed. Alas in ecology, such discussions are still shy. In this talk, I consider a case of study from ecology that aimed to construct a set of heuristics to guide the development of an explanatory model of an ecological phenomenon: the functional structure of autochthonous bee's community as well as the maintenance of their pollination services in the agricultural system. Therefore, a communicative bridge between ecology and philosophy of science was constructed reflecting a partnership between two Ph.D. students. In biology, explanations usually use mechanisms to provide an understanding of life phenomena and, in ecology, mechanisms are not only used to derive descriptive explanations but also predictive models of ecological systems. Although the widespread usage, these mechanisms for long have been developed with no solid framework concerning strategies of models and mechanisms constructions. Thus, if mechanisms are used to provide explanations and understanding, how can understanding exist when there is no framework to enable so? How can these predictive models be reliable when the framework needed for its development is absent? This talk aims to investigate how the philosophy of science may help to construct explanations in ecology and, how the understanding of an explanation in ecology is assessed by scientists themselves during the process of model construction. As a result, it is asserted that an explicative model was successfully created using as an instrument the heuristics derived from the theoretical framework between those areas. Looking at their construction and application, these heuristics not only yielded a theoretical framework for model confection as well as provide enough information to acquire parameters to comprehend how the understanding process happened. Thus, the heuristics construction exhibited features on the elaboration of explanandum; while the heuristic application exhibited features suggesting how the understanding (of explanandum and consequently explanans) was achieved by the modeler. These features were translated into different parameters (such as intelligibility, subjectivity, abstraction, ontic-epistemic constraints). Such parameters will be compared with De Regt's idea of intelligibility within the context of the notion of scientific understanding. On the basis of the case study and the comparative analysis between my parameters and De Regt's, I assert 3 points: (i) philosophy of science when combined with scientist's empirical and theoretical knowledge, during an explanation construction, may provide a better understanding of his explanandum as well as explananda; (ii) the understanding processes provided by the heuristics, concerning the phenomenon explanation, is not always related to normativity, on the contrary, is also deeply related to abstraction and subjectivity, and (iii) the interdisciplinary demanded on this research contribute to improve the understanding of a scientist's subject as well as to improve his own skill as a researcher by means of self-reflection.

Anne Ruth-Mackor

Explanation and understanding in criminal law. Some reflections on the likeliness and the loveliness of scenario's

In criminal trials, the judge (or the jury) has to decide whether it has been proven beyond reasonable doubt that defendant has committed the criminal fact he is accused of. The last few decades, two theories about how the judge can rationally decide about this question – the scenario theory and Bayesian probability theory - have gained increasing but not uncontested influence.

The scenario theory has been developed in psychology - among others by Pennington and Hastie - as a descriptive theory about how people reason about actions and events. Wagenaar Van Koppen and Crombag have given a normative twist to the theory, claiming their theory offers a model of how judges should reason if they want to make rational decisions about criminal proof. The scenario theory and Bayesian probability theory have characteristics in common. However, they also differ in several respects.

One of the respects in which the theories differ is particularly interesting in the light of the theme of the conference, viz. the relation between understanding and explanation. On a Bayesian view, explanation boils down to applying Bayes' rule to assess an compare the probability of different hypotheses at hand. The scenario theory emphasizes that specific characteristics of scenario's such as their coherence, the level of detail, the fit with background knowledge and the number and nature of evidence gaps, determine how good and acceptable a scenario is.

A question is whether the features of scenario's through which we 'understand' what has happened, can be explicated in probabilistic terms. We can also formulate this question in terms of Lipton's famous distinction: should the judge or jury opt for the scenario which is the most likely, or should they opt for the scenario which is the most lovely? In order to answer this question, we will discuss the questions what makes a scenario lovely and how does the loveliness of a scenario relate to its likeliness?

Emily Sullivan-Mumm

Understanding from machine learning models

There has been increasing work on the way that minimal models and simulations explain and what sort of understanding we can gain from them. One common view is that simple idealized models can provide more understanding than more complex or hyper-realistic models (Bokulich 2008, Kuorikoski and Ylikoski 2015, and Strevens 2008). For starters, being able to understand the model itself makes the understanding gained from the model deeper. Second, simpler models are more tractable, are said to answer more what-if or w-questions, and can highlight salient difference-makers. However, as philosophers are gaining better insight into minimal models, an increasing number of scientists are going in the opposite direction by utilizing machine learning algorithms using big data to make predictions and draw inferences. If we take the lessons learned from understanding and minimal models, it suggests that scientists are curiously opting for models that have less potential for understanding. Machine learning algorithms are opaque to modelers, they are increasingly complex and have less modeler control, and the amount of w-questions are seemingly limited. Are scientists trading understanding for some other epistemic or pragmatic good, when they choose a machine learning model? Or are the assumptions behind why minimal models provide understanding misguided? In this paper I argue that it is latter. In particular, I argue that it is not in virtue of the complexity or opaqueness of the model that limits how much understanding the model provides. Instead, it is how the findings from the model are supported by other independent or related scientific studies and how the model fits within a background of epistemic or normative commitments that determines the understanding that they provide. To make my argument, I identify two ways in which machine learning models can be opaque to modelers: low-level illegibility and algorithm illegibility. Low-level illegibility occurs the low-level details of how the model works is obscured, but the higher-level details are intact. For example, a modeler could understand the outline of the algorithm structure without understanding how each step is exactly implemented. Algorithm illegibility occurs when the algorithm itself changes through the execution of the program. This means that even higher-level details can become opaque to the modeler because the computer updates the algorithm in the course of the program being run in order to optimize for the intended result. Machine learning algorithms are unique in that they not only have low-level illegibility, but also algorithm illegibility. I argue that a lack of understanding from machine learning models is not from either types of illegibility in itself, but from a lack of supporting evidence from other scientific studies on the topic or due to a mismatch with background epistemic or normative commitments. I focus on two machine learning models to make this argument: anomaly detection in medical science and machine learning sexual orientation detection through facial recognition.

Philippe Verreault-Julien and Vaios Koliofotis

Hamilton's rule: understanding the disagreement about its explanatoriness

More than half of a century since its initial development (Hamilton 1964a; Hamilton 1964b), inclusive fitness theory remains controversial in evolutionary biology. Nowak, Tarnita, and Wilson's (2010) Nature article resulted in a fierce response (e.g. Abbot et al. 2011) and a heated and ongoing debate over the explanatory status of Hamilton's rule (e.g. Marshall 2015; Okasha and Martens 2016; Veelen et al. 2017). In a nutshell, Nowak, Tarnita, and Wilson argue that the general form of Hamilton's rule (HR) does not explain nor afford understanding. In this paper, we argue that the distinctions between how-possibly (HPEs) and how-actually (HAEs) explanations, on the one hand, and between causal and mathematical explanations, on the other hand, illuminate the source of the disagreement between the critics of HR and its supporters. Furthermore, it may serve as a cautionary note to not over or understate whether HR affords understanding. First, while Nowak et al. consider HR only supplied "hypothetical explanations" (2010, 1058), Abbot et al. consider it was successful in "explaining a wide range of phenomena" (Abbot et al. 2011, E1). In other words, there is fundamentally a dispute over whether HR provides an HAE of phenomena such as the evolution of eusociality, or only a HPE. Second, one way of interpreting Nowak et al.'s (2010; 2011) charge against HR is that not only is it a HPE, but that it is a mathematical HPE. Whether or under what conditions HR can receive a causal interpretation is nebulous (Allen, Nowak, and Wilson 2013; Birch 2014; Okasha and Martens 2016). Mathematical HAEs typically show how an explanandum holds out of mathematical necessity (Lange 2013; Pincock 2015). Interpreting HR in the light of the literature on mathematical explanation allows to see that Nowak et al.'s critique is twofold: HR fails 1) to identify necessary mathematical conditions and 2) to map physical structures similar to those the rule depicts (Bueno and Colyvan 2011; Pincock 2007). Therefore, it does not provide a mathematical HAE, but at best a HPE.

Beyond illuminating the source of the disagreement over HR, an important benefit of using these distinctions is that it allows to temper both the critics and the supporters of HR in their evaluation of whether HR affords understanding. It does so for two reasons. First, even if we accept that HR fails as a HAE of social behaviour like cooperation, it does not follow that it can't afford understanding (see Kuorikoski and Ylikoski 2015; Rice 2016). But even if HPEs may afford understanding, proponents of HR should also be careful when claiming it actually explains. Second, under suitable conditions, a lack of causal interpretation does not imply HR can't be explanatory (cf. Birch 2014). A mathematical HAE, or even a HPE, may afford understanding of empirical phenomena. While the critics may be too quick to dismiss HR on the ground that it is a mathematical explanation, its proponents should also be wary of not jumping too fast from the mathematical to the causal interpretation.

Naftali Weinberger

Explaining how

, Intuitively, learning how a cause brings about its effect helps one better understand the relationship between these variables. If, for example, being exposed to news stories about immigrants causes one to be more likely to have anti-immigrant attitudes, it is illuminating to learn that one path by which stories influence one's attitudes is via increasing one's anxiety level. Additionally, explanations that appeal to mediators such as anxiety seem to provide better explanations than those that do not. But what do we better understand when we measure mediators, and how does doing so enhance explanation? There are several philosophical areas in which one might expect to find an answer to this question. To mention just one, mechanistic explanations spell out the way that a mechanism's components interact to produce a mechanistic phenomenon, and thus ought to have something to say about why finer grained descriptions involving intermediate factors are more explanatory than those without such factors. The thesis of my talk is that without a more sophisticated understanding of the contributions of direct and indirect paths to a net effect, philosophical discussions of the explanatory role of mediators will apply only to the simplest causal scenarios. Philosophical discussions often involve a single mediator along a single path. One might imagine that being exposed to news stories influences participant's attitudes only via producing anxiety. Here the contribution of anxiety is simple: prevent the increase in anxiety and you will prevent the effect of news exposure on attitudes. But in a more realistic case, there will be other paths by which news-story exposure influences attitudes. Plausibly, the news stories also contain information regarding the economic effects of immigration, and this information also affects attitudes. Moreover, the mediators along different paths could interact in their effects. Perhaps the amount by which anxiety promotes anti-immigrant attitudes depends on the information one received. In cases like this with multiple mutually interacting causal paths, questions about the causal or explanatory "contribution" of a single mediator such as anxiety to the net effect are highly ambiguous. In my talk, I explain how modern non-parametric mediation techniques enable one to distinguish among the various counterfactual questions one could be considering when asking about the contribution of a mediator to a net effect. I then suggest that measuring mediators facilitates explanation by enabling one to answer a wider range of what-if-thing-had-beendifferent questions. While the idea that explanations with broader counterfactual scope are better is itself commonplace in philosophy of science, the challenges arising in specifying counterfactuals about mediators in models with multiple paths are generally not well appreciated. I hope to convince the audience that the challenges that arise in the multi-path case are philosophically interesting, and need to be addressed to understand the contribution of mediators to explanation and understanding.

Martin Zach

Factive understanding with model sketches

It has long been argued that idealized model sketches cannot provide us with factive scientific understanding, precisely because these models employ various idealizations; hence, they are false, strictly speaking (e.g. Elgin 2004, 2007, Potochnik 2015). Others espouse the view that understanding is quasi-factive (e.g. Mizrahi 2012), acknowledging the role of simplifying assumptions and the need to relax the standards, though there are well known issues surrounding this position. Few have defended (in one way or another) the factive understanding account despite the objections raised against it (e.g. Reutlinger et al 2017, Rice 2016). In this talk I argue for the claim that there is a way in which we can maintain the position of factive understanding. All it takes is to accept that there are different "levels of abstraction" which still can (and do) give us factive understanding. In addition to that, it should be noted that the whole debate on factive understanding also suffers from an inadequate distinction between the processes of abstraction and idealization (see Godfrey-Smith 2009 for a widely held account). As an example, consider a mechanistic model of an enzyme regulation, specifically the way in which the product of a metabolic pathway feeds back into the pathway and inhibits it by inhibiting the normal functioning of an enzyme. It can be said that such mechanistic model abstracts away from various key details, or alternatively that it idealizes various factors. For instance, it ignores the distinction between competitive and non-competitive inhibition which can be considered both as abstraction and idealization. Furthermore, a simple model often disregards the role of molar concentration. Yet, models such as these do provide us with factive understanding when they tell us something true about the phenomenon, namely the way in which it is causally organized, i.e. by way of negative feedback (see also Glennan 2017). This crucially differs from the views of those (e.g. Strevens 2017) who argue that idealizations highlight causal irrelevance of the idealized factors. For the phenomenon to occur, it makes all the difference precisely what kind of inhibition is at play and what the molar concentration of the product is (see also Love and Nathan 2015 who consider ignoring concentrations as a case of idealization). Finally, I will briefly distinguish my approach to factive understanding from those of Reutlinger et al (2017) and Rice (2016). I take it that in the terminology of Reutlinger et al (2017), a specific kind of models, embedded toy models, give us how-actually understanding (i.e. factive understanding) even though these models are highly idealized and simple. Their notion rests on the need for theory-driven de-idelization of the assumptions, however, and as such it importantly differs from my view which is free of such need. Rice (2016) suggests that optimization models give us factive understanding by providing us with true counterfactual information about what is relevant and irrelevant, which, again, is not the case in the example discussed above. Note that I am not disputing their views, but rather I am adding another type of cases in favor of factive understanding.