Inferentialism, Bayesianism and Scientific Explanation (2015-2018, DFG)

Principal Investigators: Stephan Hartmann (LMU Munich, Germany) and Marcel Weber (University of Geneva, Switzerland)

Project Description

1 State of the art and preliminary work

Explanation constitutes a cornerstone of scientific rationality. But what is valid explanatory reasoning? And when is valid explanatory reasoning a guide to truth? This project tackles such questions by offering a novel account of explanation, which draws on and unifies two original views on explanation, viz. the 'inferentialist view' and the 'Bayesian view'.

The received view(s). Philosophical accounts of scientific explanation study the dependence of explanandum on explanans. The existing variety of accounts may be usefully classified into two broad categories, depending on whether the dependence is regarded as essentially ‘epistemic’ or ‘ontic’. *Epistemic* accounts ascribe an essential role to the vehicles (arguments, models, texts) by which the explanation is conveyed. The two best-known examples are the Deductive-Nomological account (Hempel, 1962, 1965; Hempel and Oppenheim, 1945, 1948), now dismissed by most authors (see, e.g., van Fraassen, 1980; Ruben, 1992), and the unificationist account (Friedman, 1974; Kitcher, 1981; see also Bartelborth, 2002). In both accounts, an explanation is an argument where the explanandum is deduced from the explanans. In both, an explanation proceeds by subsumption under a universal generalization. In the Deductive-Nomological account, explaining something (e.g., the expansion of a heated gas) consists in showing how the explanandum follows from a law (Charles' law) and background conditions (constant pressure). In unificationist accounts, explaining something (e.g., Brahe's astronomical observations) consists in showing how an explanandum which falls under certain laws (Kepler's laws of planetary motion) can be subsumed under laws of wider scope (Newton's law of universal gravitation). *Ontic* accounts, instead, downplay the vehicles of explanation in favor of the worldly objects those vehicles refer to. Ontic accounts are typically causal: an explanation tracks the causal process or mechanism (e.g., photosynthesis) that brings about, or makes a difference to, the explanandum (the plants' production of oxygen). How this tracking is conveyed is inessential, or at most ancillary, to explanation. This ontic view stems from (Salmon, 1984, 1989) and (Lewis, 1986). It is currently most popular in the guise of mechanistic accounts (Machamer et al., 2000; Machamer, 2004; Glennan, 2002; Craver, 2007, 2014; two exceptions are Bechtel and Abrahamsen, 2005, and Wright, 2012), the interventionist account (Woodward, 2003; see also Weber, 2008), and the kairetic account (Strevens, 2008).

Unification and causation, which are central to the currently most prominent accounts of explanation, are clearly important. However, it is doubtful whether unificationist accounts or causal accounts provide a *general* model of explanation. On the one hand, unificationist accounts struggle to account for causal and statistical explanations that do not seem to have an argument structure (van Fraassen, 1980). That is, the unificationist view is too narrow, because it neglects non-deductive vehicles. On the other hand, causal accounts struggle to find what causal processes or decomposable mechanisms are tracked by explanations where the explanandum is deduced from mathematical facts (Nerlich, 1979; Sober, 1983; Batterman, 2001; Batterman and Rice, 2014; Huneman, 2010; Lange, 2013). That is, the causal view is too narrow, because it neglects certain explanatory vehicles. Not to mention that explanations in the social sciences are often based on neither laws nor causal relations but intentions and social norms. Besides, it is widely acknowledged that contextual considerations (e.g., the context of a why-question) are crucial to selecting the explanatorily relevant factors (Nagel, 1961; van Fraassen, 1988; Garfinkel, 1981; Achinstein, 1983; Faye, 2007). Such contextual considerations are typically neglected by allegedly general models of explanation. All this suggests that there may not be a unique standard to assess the validity of scientific explanations.

Another topic philosophers are focusing more and more on is the connection between explanation and truth. Explanation, it is sometimes argued, is a 'guide' to truth. For instance, we infer to the existence of electrons, the heliocentric model of the solar system, and the laws of inheritance from the explanations that they provide for a variety of phenomena. The 'beauty' of the connection between explanans and explanandum is a reason to infer to the truth of the explanans. This use of explanation is called 'inference to the best explanation', or IBE (Harman, 1965; Lipton, 2004; Douven, 2011; see also Weber, 2009). Yet, to date it is still unclear how to rationalize IBE.

Examples. Certain phenomena are particularly resistant to being captured by existing models of explanation. In biology or economics, for instance, explanations seldom proceed by reference to laws (Mendel's laws of inheritance and the law of supply and demand are notable exceptions). And although causal relations are widespread in the biological and economic domain (e.g., the relation between reproduction and selection, or between inflation and unemployment), it is hard to explain all of biological or economic phenomena in purely causal terms.

Biological phenomena are often explained by their ends, too, and not only by their causes. For instance, a biological function (e.g., pumping blood), which is meant to explain the existence of the function's bearer (the heart), is often defined as a capacity that was selected in the organism's evolutionary history, that is, by its etiology (Wright, 1973; Millikan, 1984). A problem with this view is that it does not allow that something is a function if it has arisen anew, for instance, by spontaneous mutation. Also, biologists sometimes attribute functions without knowing the bearer's evolutionary past. Either way, the identification of functions would seem to rely on teleological criteria. However, teleology is regarded as dubious in natural sciences. Some propose to naturalize the definition by explaining function as (causal) contribution to self-reproduction, that is, maintenance of an organism's identity by a continuous replacement of its parts (Maturana and Varela, 1980; McLaughlin, 2001; Mossio, Saborido and Moreno, 2009). This explanation faces two *prima facie* problems. First, the definition seems to involve downward causation, insofar as the organ's function depends on the whole organism's function. Second, the definition seems to involve causal holism, insofar as it mentions the contribution to self-reproduction of all parts of the organism. Weber, (2005a) has recently made an alternative proposal: biological functions are defined in terms of the *coherence* between the capacity for the target function and the capacities for all other functions, such that the whole system of capacities provides the best explanation for the system's self-reproduction. The dependence among the functions themselves is non-causal (Cummins, 1975), and so is the dependence of the target function on self-reproduction. The explanatory value of coherence has been largely neglected (a notable exception is Thagard (1989); see also Colombo, Hartmann and van Iersel (2014)). Weber proposes to understand the coherence in question as analogous to the coherence among sentences. How exactly the coherence constraint is to be understood in this context is an open issue.

Similarly, some economic phenomena seem to depend on intentions and rules. For instance, when economists explain price fluctuations in the stock market, they cite the rule that one should buy underpriced assets and sell overpriced ones, the (rational) belief that price will revert to its fundamental value, the intention to exploit temporary asymmetries in information on the difference between fundamental value and actual price, etc. It is unclear whether such explanations depend on laws (the 'laws' of rational choice?) or causes (the agents' preferences and expectations?). Furthermore, there are other economic phenomena, such as the "stylized facts" of financial time series (fat-tails, volatility clustering and persistence), that seem best explained by neither laws, nor causes, nor rational expectations (Kuhlmann, 2011; Rickles, 2011; Casini, 2014a). Several agent-based models (Lux and Marchesi, 1999, 2000; Arthur et al., 1997; LeBaron et al., 1999) explain the stylized facts in terms of their robust dependence on structural properties of the market – such as the agents' heterogeneity – irrespective of the market's 'laws' and causal details. Robustness analysis (Woodward, 2006; Weisberg, 2006; Kuirikoski, Lehtinen and Marchionni, 2010) is typically used to identify potential explanantes in such models. How exactly the reasoning behind robustness analysis is explanatory is an open question. On the face of it, claiming that a phenomenon obtains *because of* laws, causes, or expectations, is intuitively different from claiming that it obtains *irrespective of* them. Also an open issue is whether the explanations of the stylized facts ought to be accepted as true. The agent's heterogeneity seems to be the best explanation of the stylized facts. Should we then infer to the truth of the corresponding hypothesis? This question is all the more pressing, since neoclassical macroeconomics offers no alternative, equally satisfying, explanation (Kirman, 2010).

A novel approach. In the light of the above general considerations and problematic examples, what one would need is an approach that takes the diversity and contextuality of scientific explanations seriously, while simultaneously being informative on what different explanations have in common, or how they can be more or less appropriate. In this project, we develop such an approach by reviving the epistemic view that explanations are arguments and by combining two philosophical frameworks – inferentialism and Bayesianism – that help each other in achieving our proposed goal. Our analysis will be informed by, and tested against, the case studies of biological functions and the stylized facts of finance. The state-of-the-art of the research regarding the inferentialist and the Bayesian approaches to explanation and IBE are outlined below, together with the research questions we wish to address in this project.

Inferentialism (Harman, 1999; Brandom, 1994, 2008) is a pragmatist approach to semantics (Sellars, 1953; Wittgenstein, 1956). It has been applied to the interpretation of logical rules (Dummett, 1991; Peregrin, 2006; Prawitz, 2006; Steinberger, 2011), causal claims (Reiss, 2011, 2012; Casini, 2012), and the norms governing the acceptance of scientific hypotheses (Zamora-Bonilla, 2006). Inferentialism explicates the meaning of linguistic expressions in terms of their role in 'language games' where we give and ask for *reasons*. Such reasons are given in terms of the assertions' *inferential connections* with the circumstances of their appropriate application (the premises) and the consequences of their appropriate application (the conclusions). The game is governed by social norms that dictate (depending on context-dependent factors such as available evidence, shared knowledge, or intended purpose) which conclusions one should draw.

Inferentialism supports an epistemic view of explanation. An explanation is an argument offering a (deeper) reason for a conclusion, which is already believed but not so well understood or justified (de Donato-Rodriguez and Zamora-Bonilla, 2009a,b; Zamora-Bonilla and de Donato-Rodriguez, 2013; see also Walton, 2004). A reason can be explanatory in several ways. For instance, it can make the explanandum *more credible*, by increasing its plausibility or fruitfulness. One way to increase plausibility is to produce a warrant that entitles one to a commitment. For instance, knowing of an advance in the technology used by a software firm explains the evidence that the firm had a good year by making the evidence more expectable. One way to increase fruitfulness is to increase the ratio of successful to unsuccessful inferences and, in so doing, one's entitlement to certain actions and predictions. For instance, explanations by unification make the hypotheses in one theory translatable into the vocabulary of another theory and applicable to the latter's domain, too, thereby increasing the ratio of successful to unsuccessful inferences. Yet another, less obvious way for a claim to be explanatory is to make previous commitments more "enlightening". This happens when an incompatibility between the commitment to one claim and the entitlement to another is made explicit and removed, thereby smoothing the flow of inferences. This move makes the explanandum *more coherent*. Here the following questions naturally arise: Is there something all explanations have in common, in spite of this apparent diversity? And can inferentialism be used to account for both the diversity and the commonality among scientific explanations?

Whilst in line with the view that explanations are arguments, the inferentialist view on scientific explanation differs from current argument-based models in that not only formal but also 'material' rules of inference are accepted as valid. Material inferences such as 'Lightning. Therefore, thunder.' are valid in virtue of their content rather than their logical form (Sellars, 1953). It has been argued that the category of material inferences is broad enough to encompass different kinds of explanation, depending on the kind of concepts appealed to (Brigandt, 2010): the explanans may grant the inference to the explanandum because it mentions a cause (causal explanation), or a law (DN, unification), or intentions/norms (intentional action explanation), etc. In this regard, the following question arises: Can material inferences be employed to argue that the inferentialist account is superior to available models in accounting for the validity of explanations?

Another concern one may have is how to decide whether one explanation is better than another. Inferentialism is of little help when it comes to quantitative aspects of explanations, such as measuring their strength. For this, we turn to *Bayesianism*. Bayesianism is an overarching framework for formal epistemology (Howson and Urbach, 2006). According to Bayesianism, an agent's degrees of belief are represented by a probability function. The Bayesian argues that in order to avoid sure loss irrespective of the outcome to which probabilities are associated (i.e. to avoid a 'Dutch book') the agent's degrees of belief ought to be probabilistically consistent, and that her changes in degrees of belief upon learning new evidence ought to obey the principle of conditionalization, i.e., the new unconditional probability of a hypothesis equals its old conditional probability given the evidence. (The latter probability, also called 'posterior' probability, is calculated according to Bayes' rule as equal to its 'prior' probability times the 'likelihood' of the evidence given the hypothesis divided by the probability of the evidence.)

Although Bayesianism alone does not offer a substantive account of what an explanation is, it lends itself naturally to the task of quantifying the strength of explanations. Recent work (Schupbach and Sprenger, 2011; Crupi and Tentori, 2012) has exposed an important relation between explanation and confirmation: a good explanans often increases the expectedness of the explanandum. By exploiting this relation, Bayesianism has been used to quantify the strength of explanations through measures that best satisfy general probabilistic constraints and intuitive adequacy criteria. Such measures have then been tested against case studies (McGrew, 2003; Myrvold, 2003) or psychology experiments (Schupbach, 2011a). They certainly capture the quantitative aspect of one important kind of explanations. A shortcoming is that not all cases of good explanation result in an increase in expectedness. From an inferentialist point of view, other explanatory patterns exist, such as those that, by removing incompatibilities, increase the coherence of the overall framework. Since inferentialism itself does not quantify their strength, the question arises: Are there Bayesian measures of coherence-increasing explanations, too?

Another issue concerns the capacity of Bayesianism to account for epistemic uses of explanations, such as *IBE*. Here, it is inferentialism that aids Bayesianism. Through IBE one infers to the high(er) probability of a hypothesis from its high(er) explanatory power, i.e. from the hypothesis' high(er) capacity to explain some evidence. Bayesian conditionalization, too, may be used to infer the high(er) probability of a hypothesis given some evidence. Yet, for the latter kind of inference the assessment of explanatory power appears inessential, in the sense that Bayes' rule may allow us to infer a hypothesis' high(er) posterior probability whether or not the hypothesis explains the evidence. This is so because, although probabilistic features might indicate the existence of explanatory features, the link between the former and the latter is not straightforward. For instance, one cannot simply equate high explanatory power with high probability, as there are cases of highly explanatory hypotheses with low probability and cases of hypotheses with high probability and low explanatory features (cf. van Fraassen 1980, 1989). An important consequence is that the Bayesian framework is by itself unable to provide a full account for the apparent rationality of IBE.

The problem emerges in full force whenever one's inference to the high(er) probability of a hypothesis – through IBE – cannot be backed by probabilistic judgments because there are insufficient data for setting the probability values that could be considered evidence of explanatory features. For instance, in the lightning/thunder case, one can imagine that only a few observations and no alternative hypotheses are available. Still, one may want to infer to lightning as the best available explanation of thunder. If such an inference is rational at all, this does not seem to depend on Bayesianism. Here, inferentialism comes to rescue: material rules of inference are applicable also when only one hypothesis is available – it seems sufficient to observe lightening to infer thunder. Inferentialism seems to ground the rationality of IBE in these and other cases. When material inferences from the explanans to the explanandum appear correct, it would be desirable to apply Bayes' rule to turn the originally qualitative analysis into a quantitative one, thereby increasing its accuracy and precision. So the question arises: How exactly, if at all, does inferentialism complement Bayesian rationality and rationalize IBE?

2 Objectives and work programme

2.1 Objectives

The state of the art of the literature on scientific explanation reveals that available models of explanation can hardly encompass the variety of kinds of scientific explanation. At the same time, two original research programs are being put forward. The inferentialist approach attempts to flesh out and systematize the intuition that explanation is diverse and contextual in terms of the role that explanations play in our reasoning practices. The Bayesian approach focuses instead on modeling explanatory reasoning in probabilistic terms. This project aims to bring together the resources of the two approaches to tackle the traditional task of interpreting scientific explanation from an innovative perspective. In particular, our project aims to answer the following research questions:

1. Can a contextual account of explanation such as the inferentialist account be *compared* – and possibly proved *superior* – to current models of explanations? In particular:
   - Can the notion of material inference be used to show that the inferentialist account is superior to traditional argument-based accounts?
   - Can one classify all various kinds of explanation under the umbrella 'inference facilitation'?
   - Can one account for causes as particular kinds of reasons, and causal explanations as reason-giving explanations?

2. Can one single out a different Bayesian measure of explanatory power for each kind of explanation? In particular:
   - Are coherence-based measures of explanatory power irreducible to confirmation-based measures of explanatory power?
   - Are there distinct measures of causal explanatory power, corresponding to different kinds of causal explanation, or are they reducible to one?
   - Are coherence- and confirmation-based measures reducible to causal ones?

3. Are the inferentialist and the Bayesian account to explanation compatible and mutually integrating? In particular:
   - Can material inference be used to justify Bayesian rationality?
   - Can such a justification be used to understand the rationality of IBE?

4. How faithful to *scientific practice* is our conception of explanation? In particular:
   - How should the notion of 'inference facilitation' be interpreted if one is to make sense of robustness explanation of the stylized facts of finance?
   - How should the notion of 'inference facilitation' be interpreted if one is to define, and thereby explain, biological functions in terms of coherence?
   - In view of our inferentialist-Bayesian reconstruction of IBE, what would justify the inference to

the truth of the best explanation of the stylized facts?

2.2 Work programme incl. proposed research methods

Work programme

Our project will take place at two universities, LMU and Geneva. Our research program will employ a *mixed methodology*, comprising conceptual analysis and formal modeling (LMU), and two case studies on biological functions and the stylized facts of finance (Geneva).

We will investigate the meaning and pragmatics of explanation within the philosophical framework of inferentialism. Since inferentialism is primarily a theory of meaning, it lends itself naturally to the task of conceptual analysis. For the quantitative analysis of changes in degrees of belief in explanations, we will employ Bayesian modeling. While Bayesianism and inferentialism independently have much to offer towards a better philosophical analysis of explanation, there are interesting connections between them which we want to exploit in this project. Both approaches adopt a pragmatic attitude: inferentialism with respect to the explication of meaning, Bayesianism with respect to the probabilistic consistency of degrees of belief. Also, both stress the importance of coherence between background beliefs/commitments and available evidence for successful explanatory claims. We will draw on these parallels to develop an integrated approach to explanation, overcome the limitations of available models of explanation, and understand the rationality of IBE.

To study the implications of our proposed integration of Bayesianism and inferentialism, we will examine two case studies, one from biology, viz. the explanation of biological functions, and one from economics, viz. the explanation of the stylized facts of finance. We chose these two examples because their candidate explanations, based on respectively coherence and robustness, are not well accounted for by traditional models of explanation. Since our project aims to provide, among other things, an account of explanations that rely on coherence and robustness, our contribution purports to be original. In addition, if our account can be applied to explanations as problematic as those of biological functions and stylized facts, this seems good evidence that it will be applicable elsewhere to less or equally problematic cases. For these reasons, we believe that our choice of case studies make a good test for the proposal we put forward.

The project is split into four subprojects, each in turn subdivided into several studies. Subprojects 1-3 will be conducted at LMU. Subproject 4 will be conducted at Geneva. Investigators at LMU and Geneva will continually interact during the project.

Subproject 1 (LMU): Inferentialism and scientific explanation. None of the standard accounts of explanation is exempt from counterexamples. Also, we increasingly appreciate how much the character of scientific explanations varies from discipline to discipline. This casts doubt on the possibility to capture the 'logic' of explanation by means of monistic accounts. Yet, it would be desirable to have a framework that explains why different contexts require different kinds of explanation. This project paves the way for such a framework by offering an inferentialist interpretation of scientific explanation.

Study 1.1: Explanation and material inference. Drawing on the classical works by Peirce (1931) and Hempel (1965), inferentialism interprets explanations as *reasons*: once an explanation is given, the explanandum makes more sense. In order to defend this epistemic view against critics such as Salmon (1984, 1989) and Lewis (1986), it is crucial to have an adequate account of inference, which makes the inferential account superior to alternative argument-based accounts, namely the Deductive-Nomological and the unificationist accounts. The aim of this study is then twofold.

First, we will provide an analysis of the notion of material inference. To this end, we will make use of the Sellarsian notion of ‘material' – as opposed to ‘formal' – mode of inference (Sellars, 1953; Norton, 2003; Brigandt, 2010). Material inferences (e.g. 'Smoking. Therefore cancer.') are defeasible (not all cases of smoking result in cancer) and non-enthymematic (there is no additional true premise that would make the conclusion follow as a matter of logical necessity). Their validity depends on contextual considerations. Reference to material inferences will allow us to justify the claim that explanations may be good, yet contextual: contrary to other argument-based accounts, explanations don't always proceed by reference to universal generalizations.

Secondly, we will develop an inferential account of the meaning of ‘explanation’. We will argue that explanations are ‘inference facilitators’, argumentative moves where one claim is put forward to facilitate inferences to other claims already accepted, but not well understood or justified (de Donato and Zamora-Bonilla, 2009a,b). Crucially, the facilitation can obtain in various ways. This explains the existing diversity of explanations: an explanation may increase the expectedness of the explanandum; or it may remove/reduce the incompatibility among previous commitments, thereby increasing their coherence; or it may make previous commitments more fruitful; etc. Inferentialism can make sense of both the diversity and the similarity among explanations, thereby providing a contextual, yet informative account of explanation. An explanation varies depending on features of the context, such as the information available to, and the purpose of, the explainee and the explainer. At the same time, in any given context the view that explanations are inference facilitators serves as the interpretive key for evaluating the explanatory value of the arguments where the explanandum is inferred.

For both the analysis of material inference and the analysis of explanation, we plan to exploit the intuitive connection between material inferences and nonmonotonic logic. Nonmonotonic logic allows one to do justice to the intuition that meaning depends on inferences whilst accepting that a number of them are not universally valid (Makinson, 2005; for an application to scientific theories, see Andreas, 2011, 2012). Whereas the account of material inference will serve to analyze the logic of *particular* explanatory patterns, the inferential account of explanation will serve to analyze the concept of explanation *as such*, on the basis of a (possibly open-ended) hierarchy of conditions for inference facilitation. Since the conditions may change from one context to another, explanation is analyzed as a "cluster concept", such that no set of criteria in the cluster may be sufficient or necessary.

As regards the choice of nonmonotonic logics, two possibilities come to (our) mind. One possibility to use *default* logic (Reiter, 1980), which is rule-based rather than axiom-based, and thus particularly well-suited to our task of explicating explanation in terms of material inferences. The application of rules in default logic depends on consistency constraints on the conclusions of such rules. Because of such constraints, applications of default rules may have to be later retracted when further information becomes available. *Prioritized* default logic introduces a hierarchy of default rules. Importantly, since the correct applicability of default rules may be either tacit or unknown, the validity of material inferences doesn't depend on making explicit the deductive form of the argument. This makes it possible to vindicate the legitimacy of explanations in disciplines (e.g., sociology) where axiomatized theories are simply not available. Alternatively, another possibility is to use the preferred subtheories system (Brewka, 1991; see also Andreas and Casini, 2014), which has the advantage of being of easier applicability. This is an inference system for paraconsistent and nonmonotonic reasoning that provides a particularly intuitive model of belief changes in the sense of belief revision theory. The system is based on the application of a hierarchy of axioms, ranked by their reliability/accuracy.

Study 1.2: Causes as conditional reasons. It has often been argued (Salmon, 1984; Dowe, 2000; Machamer et al., 2000; Machamer, 2004; Lewis, 1986; Woodward, 2003; Craver, 2007, 2014) that explanations point to causal relations, or causal mechanisms, and that any understanding of explanation that cuts this intimate connection would be misguided. At the same time, following the lead of psychologists, philosophers are realizing that subjective factors are an uneliminable component in judgments of 'actual' causation, that is, judgments on which causal relation is explanatorily relevant in a given context (Hitchcock and Knobe, 2009; Halpern and Hitchcock, 2010; see also Hilton and Erb, 1996; Lombrozo and Carey, 2006; Lombrozo, 2010). We think the best way to reconcile these intuitions is to interpret causes as *reasons*.

Causes will be conceived of as a particular kind of reasons, viz. "conditional" reasons that obey certain probabilistic constraints (Spohn, 1993, 2012). Building on previous work (Casini, 2012, 2013) and the results of study 1.1, we will argue for two claims. First, interpreting causes as reasons is compatible with viewing causal claims as objective, thereby accounting for a crucial mechanistic intuition. The idea is that, although causal explanations are contextual (for the inferentialist, they lack necessary and sufficient truth conditions), relative to a given context material inferences may be more appropriate (causal claims act as default inference rules/axioms) or less appropriate (further information invalidates the inference). Second, interpreting causes as reasons explains why one's reasons for the explanandum are not always reducible to knowledge of causal relations.

Particular attention will be paid to Williamson (2013)'s argument against the inferentialist interpretation of causal explanation. Williamson claims that inferentialism is either made redundant by the existence of matters of fact which explain by making causal claims true, and on which the correctness of the inferences parasitically depends, or is unable to account for systematic mistakes in our linguistic practice. He then goes on to defend a mechanistic account of explanation, which proceeds by identifying the truth-maker of the explanandum. We will defend the inferentialist account by arguing that: first, since contextual considerations are ineliminable, reasons do not reduce to matters of facts; and second, mistakes in explanatory practices should be regarded as possible only against the rules of particulars communities. Next, we will take issue with interpreting mechanisms as having explanatory value in virtue of constituting the truth-maker of the explanandum. Colombo, Hartmann and van Iersel (2013) have recently suggested that a mechanism may explain by making our beliefs more coherent, and coherence need not be truth-conducive. We will give to this claim an explicitly inferentialist interpretation: mechanisms explain because they facilitate the inference to the explanandum. One way to do this is to make the explanandum 'true', or more expected. Another way is to increase the explanandum's coherence with our previous commitments.

Subproject 2 (LMU): Bayesianism and scientific explanation. Bayesianism has so far been applied to a variety of topics from confirmation theory, to causal discovery, the coherence theory of justification, and testimony (Hartmann and Sprenger, 2010). This project extends its scope to scientific explanation. More specifically, we will use Bayesianism to propose measures of explanatory power along the lines of (Schupbach and Sprenger, 2011) and (Crupi and Tentori, 2012).

Study 2.1: Coherence and explanatory power. Inferentialism helped us to substantiate the claim that explanations are reasons. In order to quantify how our degrees of belief in an explanatory hypothesis should change in the light of new evidence, we turn to Bayesianism. Intuitively, many explanations are confirmation-increasing. Schupbach and Sprenger (2011) and Crupi and Tentori (2012) have recently explored the parallels between explanation and confirmation by relating Bayesian confirmation measures to measures of explanatory power. They focus on measuring the power of explanations where the explanans makes the explanandum more expected. We are now interested in another intuitive feature of explanations, namely making a set of propositions more acceptable as a whole. This qualifies as a Bayesian explication of explanatory coherence.

The driving intuition is that an explanation that proceeds by increasing the coherence between evidence and hypotheses need not increase the overall confirmation of the evidence. (An orthogonal claim is made by Shogenji (2005) and Wheeler and Scheines (2013), viz. more coherent evidence does not necessarily increase the confirmation of a hypothesis.) For instance: two competing software firms, using different systems, had a good year. This is partly explained by a technological progress in their respective systems, but not so well explained if we also consider the firms' competition. In fact, progress in one system is positively relevant to one firm but negatively relevant to the other. So, the observation that *one* firm did well is less supported by the progress of *both* systems than it is by the progress of the system used by the firm in question. In this context, an extra hypothesis, for instance that a third firm had a bad year, may partly remove the incompatibility, by making competition less fierce. Thus, the extra hypothesis explains the good year of each firm by increasing the coherence of the good year with the technological progress in both systems. Importantly, depending on the exact probabilities, the extra hypothesis may actually *reduce* the chance that *both* firms had a good year. So, explanation by coherence increase does not necessarily reduce to explanation by confirmation increase.

Building on existing research (Thagard, 1989; Thagard and Verbeurgt, 1998), we ask what adequacy criteria should be satisfied by coherence-increasing explanations, and what measures of explanatory power satisfy them best. To this end, we will investigate the relation between explanatory power and existing probabilistic measures of coherence (Olsson, 2002; Shogenji, 1999; Schupbach, 2011b; Fitelson, 2003). Finally, we will investigate under what conditions the identified measures of explanatory power by coherence increase reduce to the measures of explanatory power by confirmation increase.

Study 2.2: Causation and explanatory power. The idea that causal inference can be modeled by means of Bayesian probabilities and Bayesian networks has been extensively investigated (Spirtes et al., 1993; Pearl, 2000; Williamson, 2005; Gopnik and Schulz, 2007; see also Casini et al., 2011; Casini, 2014b). Recently, the Bayesian approach has been extended to comparing/selecting probabilistic measures of *causal power*. Fitelson and Hitchcock (2011) discuss the adequacy of existing measures (e.g., Good's measure, Suppes' measure, Cheng's measure) in different contexts. Korb et al. (2011), instead, put forward a new measure based on information theory and causal Bayesian networks, and argue that it is superior to existing measures. Both studies are clearly relevant to the present project, since they suggest an intuitive connection between the power of a cause and the power of the explanation, which cites that cause.

Whilst believing that Bayesianism may be a powerful tool to analyze causal reasoning, we remain uncommitted on whether causation can be fully explicated in terms of Bayesian networks. We adopt an analogous pragmatic stance with regard to explanation. This study sets up a Bayesian reinterpretation of causal explanation in line with the inferentialist view of causes as particular kinds of reason elaborated in study 1.2. The study is informed by the idea that a causal explanation is grounded in objective properties of the world, but essentially relative to pragmatic factors, such as the state of knowledge and the epistemic interests of an agent. This move has two advantages. First, it makes explicit the (in our view: unavoidable) interest-relative factors that enter explanation, such as the selection of a relevant contrast. Second, it makes the concept of explanation amenable to a Bayesian explication in terms of positive relevance relations.

In this study, we look for measures of explanatory power that best satisfy general adequacy criteria for *causal* explanations, each set of criteria corresponding to a different intuition about causation (as a necessary condition to the effect, or a sufficient condition to the effect, etc.). The selection of these measures will draw on existing work on measures of causal power. We will (re-)classify existing measures of causal power as measures of causal explanation that satisfy different intuitions. Finally, we will compare the adequacy of the measures, and consider whether some (e.g., Korb et al.'s measure of causal power) are systematically better than others.

Study 2.3: Confirmation, coherence and causation. Wheeler and Scheines (2013) have recently claimed that the interplay between coherence and confirmation is fully captured in terms of causal knowledge. More precisely, they argue that causal knowledge (embedded in causal Bayesian networks) suffices – in certain cases, at least – to explain why coherence and confirmation sometimes track each other (Wheeler, 2009) and sometimes don't. This may have the following consequence: if the logic of coherence and confirmation were reducible to causation, arguably so would explanation by confirmation increase and explanation by coherence increase. In turn, this would undermine our conjecture that explanation, whether causal or not, is diverse and not reducible to causation. We will test Wheeler and Scheines' claim against the cases of explanation emerged in 2.1 and 2.2.

Subproject 3 (LMU): Bayesianism, inferentialism, and scientific explanation. Subprojects 1 and 2 have shed light on two important aspects of scientific explanation: inferentialism helps us to substantiate the claim that explanations are reasons; Bayesianism helps us to measure explanatory power. Guided by the intuition that inferentialism provides the semantic underpinning for Bayesian machinery, this study will explore the compatibility of the two frameworks, in particular how they aid each other in accounting for the semantics of 'explanation' and in vindicating the rationality of an important epistemic use of explanation, viz. IBE.

Study 3.1: Bayesianism meets inferentialism**.** One may worry that, in the same way Bayesianism has in the past been used to (try to) provide a (formal) logic of confirmation, coherence, causation, etc., the use of Bayesianism in our project is driven by the implicit goal of providing a logic of explanation. This is, however, not so. There seems to be more to the rationality of belief change than Bayesian conditionalization – and inferentialism is helpful in explaining *why*. For instance, Salmon (2005) suggests that rationality is split into a *kinematic* kind, which follows Bayesian conditionalization, and a *dynamic* kind, which goes beyond conditionalization insofar as it requires the revision of the set of hypotheses on which to conditionalize. Inferentialism bridges the gap between kinematic and dynamic rationality.

Reasoning – explanatory reasoning included – depends on *material* connections between premises and conclusions. The correctness of material inferences cannot be reduced to an algorithm – whether Bayesian or else. Indeed, inferentialism leads naturally to regarding explanation as at most definable as a cluster concept in a network of other concepts (e.g., confirmation, coherence), which may themselves be cluster concepts. Yet, given certain contextual constraints, inferentialism still tells what counts as a correct inference. For instance, the adequacy of an explanation may depend on the particular way the explanans should facilitate the inference to the explanandum – by increasing its expectedness, or its coherence, or… The hypothesis that contributes to the kind of inference facilitation most appropriate in some context is the explanation one ought to endorse. In that context, Bayesian conditionalization is vindicated as a tool which further supports or undermines the endorsement of the explanation in the light of the evidence.

An important consequence of this inferentialist foundation is that Bayesianism, too, can contribute to elucidate what 'explanation' means in some context, by determining the strength of the inferential connections between the concept of explanation and other concepts (confirmation, coherence, etc.). The elucidation obtains because the Bayesian measures of explanatory power help in reconstructing the rationality of an explanation in terms of how much it depends on, say, coherence considerations, or by showing that it depends more on coherence rather than, say, confirmation, or on an interpretation of coherence rather than another, etc.

Study 3.2: IBE. The inferential foundation also helps Bayesianism rationalize IBE. In the literature there is a debate on whether Bayesian inference and IBE are *compatible*. Van Fraassen (1989) argues that they are not, because explanatory power is not reducible to posterior probability. Hence, Bayesianism cannot be used to support IBE. Others (e.g.: Okasha, 2000; Lipton, 2004; Weisberg, 2009; see also Weber, 2009) have replied that this is to misread the relation between Bayesianism and IBE. IBE does not compete with Bayesianism. Rather, explanatory considerations are often reflected in priors and likelihoods, or they act as heuristics for determining priors and likelihoods.

We conjecture that inferentialism – by motivating such explanatory considerations – can be used to aid IBE, thereby supporting the 'compatibilist' view. From an inferentialist perspective, a theory is explanatory if it facilitates inferences among the shared commitments of the community. This is evaluated by considering to what extent default rules are successfully applied, that is, applied without later retraction. These considerations affect prior probabilities and/or likelihoods. From a Bayesian perspective, a theory – if explanatory – should make the set of shared beliefs in the light of the evidence more confirmed, or coherent, etc. This is evaluated in terms of posterior probabilities. The Bayesian reconstruction of IBE, then, works as follows: once one has established which one among a set of hypotheses has the highest explanatory value (according to the appropriate measure), one is entitled to conditionalize on the evidence to infer to the hypothesis with highest probability within the set, that is, to apply IBE. Due to the connections between explanation on the one hand, and confirmation and coherence on the other, we expect IBEs to be typically truth- or at least coherence-conducive.

Subproject 4 (Geneva): Application to the case studies. The Bayesian-inferential framework developed in subprojects 1-3 has made two main contributions to the analysis of scientific explanation. It has offered a novel interpretation of 'explanation' and it has proposed a novel reconstruction of IBE. In this subproject, we apply these results to two case studies, namely the explanation of biological functions and of the stylized facts of finance.

Study 4.1: Causal vs non-causal explanation. What is the difference between causal and non-causal explanations? We will use the case study on the explanation of the stylized facts of finance (Casini, 2014) to investigate the difference between causal and non-causal explanations in terms of the different inferences they facilitate. We expect the category of causal explanations to be diverse, tracking the diversity of our causal intuitions in different contexts – viz. causes are necessary to their effects (as in counterfactual accounts of causation), or sufficient to them (regularity accounts), or probability-raisers (probabilistic accounts), etc. (cf. Fitelson and Hitchcock, 2011). Because of this, we conjecture that non-causal explanations differ from causal explanations to different extents, and may be close-to-causal in one respect but not in another.

More precisely, our investigation will contrast robustness analysis, which is the explanatory reasoning employed in the ABMs of asset pricing, and other kinds of explanations. In the asset pricing case, robustness analysis contributes to explanation by identifying the (seemingly non-causal) determinants of the stylized facts, namely the heterogeneity and the bounded rationality of all agents, which make them react to market moods (Lux and Marchesi, 1999, 2000) or learn inductively (Arthur et al., 1997; LeBaron et al., 1999). This kind of explanation will be compared with the (causal) explanation based on invariances under intervention (Woodward, 2003). Both robustness and invariance rely on the stability of a relation across certain changes. However, they also differ. Invariance relies on stability across background contexts. Robustness relies on stability across micro-specifications. Invariance involves sensitivity: the explanandum is explained if changes in the explanans make a difference to it. Robustness involves insensitivity: the explanandum is explained if changes in non-explanatory factors do not make a difference to it. In addition, robustness has similarities and differences with the explanatory (and also seemingly non-causal) notions of topological invariance (Huneman, 2010) and universality (Batterman, 2001). The study aims to find a common ground to interpret all of the above kinds of explanation.

Study 4.2: Explanation of biological functions**.** What explains biological functions? Following (Weber, 2005a), we conjecture that an organ has a function if the (causal) capacity to perform that function helps the other capacities to perform their own functions, so as to maximize the overall coherence of the entire system of capacities. We have two kinds of relations here, causal relations among those capacities that best explain the self-reproduction of the system, and functional relations among functions.

Our hunch is that, causally, an organism may be interpreted as a homeostatic mechanism for self-reproduction, sustained by the interactions among its capacities. This may be modelled by means of, say, an equilibrium Bayesian network (Spirtes, 1995; Pearl and Dechter, 1996; see also Clarke et al., 2014). The network should be the best explanation of the equilibrium state, in the sense that it maximizes a measure of causal explanation proposed in 2.2. Functionally, instead, an organism may be interpreted as a network of functions that 'supervene' (in a sense to be qualified) on the capacities.

To each function corresponds a theory, made of sentences that state the dependence relations between the target function and the other functions. Because of the overlap among the theories' content, the functions probabilistically depend on each other. Since each capacity may ground several functions, to the mechanism for self-reproduction correspond many possible networks of theories. Relative to a fixed system of other functions, a target function will then be singled out as the theory that contributes best to the coherence of the system of theories to which it belongs, in the sense that it maximizes a measure of explanatory coherence proposed in 2.1. This study will aim to exploit this intuition, by studying what kind of constraints the causal network imposes on the functional network, so that to the stability of the self-reproduction state correspond a unique set of functions.

Study 4.3: The best explanation of the stylized facts. Can we use the Bayesian-inferentialist reconstruction of IBE to rationalize the inference to the truth of the (only) explanation of the stylized facts of finance? We will consider whether our reconstruction of IBE can be used to embed the rationality of "no-alternatives arguments" (NAAs) in the wider framework of IBE. In particular, we will ask whether the Bayesian-inferentialist justification of IBE extends to the kind of explanation provided by the ABMs of asset pricing, too. NAAs – popular in the string theory community – start from the observation that a scientific community has not yet found an alternative to a proposed theory, and take this fact as evidence in favor of the theory in question, even in the absence of the possibility of direct confirmation of the theory (Dawid, Hartmann and Sprenger, 2013).

Despite their negative nature, NAAs resemble IBE arguments. The reasoning involved in the explanation of financial time series offered by computational economics seems underpinned by a NAA. ABMs are notoriously hard to validate (Fagiolo et al., 2007). In the case of asset pricing, it is hard to prove that the mechanism that generates the financial time series necessarily possess certain features (interactions, learning, asymmetry of information, etc.). Yet, ABMs of asset pricing based on the general assumption that agents are heterogeneous provide a convincing explanation of phenomena for which neoclassical macroeconomics seems unable to provide an alternative, equally good explanation (Casini, 2014a).

3 Bibliography

Achinstein, P. (1983). *The Nature of Explanation*. Oxford: Oxford University Press.

Andreas, H. (2011). A Structuralist Theory of Belief Revision. *Journal of Logic, Language and Information,* 20:205-232.

Andreas, H. (2012). *Thinking in Frames. A Structuralist Approach to Belief Revision*. Postdoctoral Thesis, in preparation for publication.

Andreas, H. and Casini, L. (2014). Explanatory Conditionals. Manuscript in preparation for submission to the *British Journal for the Philosophy of Science*.

Arthur, W. B., LeBaron, B., Palmer, B., and Taylor, R. (1997). Asset Pricing under Endogenous Expectations in an Artificial Stock Market. In Arthur, W. B., Durlauf, S. N., and Lane, D. A., editors, *Economy as an Evolving Complex System II,* vol. XXVII, pp. 15-44. Santa Fe Institute Studies in the Science of Complexity, Reading, MA: Addison-Wesley.

Bartelborth, T. (2002). Explanatory Unification, *Synthese,* 130(1):91-107.

Batterman, R. (2001). *Devil in the Details*, Oxford University Press.

Batterman, R. W. and Rice, C. C. (2014). Minimal Model Explanations. *Philosophy of Science*, 81:349-376.

Bechtel, W. and Abrahamsen, A. (2005). Explanation: A Mechanist Alternative. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 36:421-441.

Bovens, L. and Hartmann, S. (2003). *Bayesian Epistemology*. Oxford: Clarendon Press.

Brandom, R. B. (1994). *Making It Explicit: Reasoning, Representing, and Discursive Commitment*.

Harvard University Press.
Brandom, R. B. (2008). *Between Saying and Doing. Towards an Analytic Pragmatism*. Oxford: Oxford University Press.
Brigandt, I. (2010). Scientific Reasoning is Material Inference: Combining Confirmation, Discovery, and Explanation. *International Studies in the Philosophy of Science*, 24:31-43.
Brewka, G. (1991). Belief Revision in a Framework for Default Reasoning, *Proceedings of the Workshop on the Logic of Theory Change*, London: Springer, pp. 602-622.
Casini, L. (2012). Causation: Many Words, One Thing? *Theoria*, 27(74):203-219.
Casini, L. (2013). *Causality in Complex Systems. An Inferentialist Proposal*. PhD thesis, Department of Philosophy, University of Kent.
Casini, L. (2014a). Not-so-minimal Models: Between Isolation and Imagination. *Philosophy of the Social Sciences*, 44(5):646-672
Casini, L. (2014b). Can Interventions Rescue Glennan's Mechanistic Account of Causality? *British Journal for the Philosophy of Science*, forthcoming.
Casini, L., Illari, P., Russo, F., and Williamson, J. (2011). Models for Prediction, Explanation and Control: Recursive Bayesian Networks. *Theoria*, 26(70):5-33.
Clarke, B., and Leuridan, B., and Williamson, J. (2014). Modelling Mechanisms with Causal Cycles, *Synthese*, 191:1651-1681.
Colombo, M. and Hartmann, S. and van Iersel, R. (2014). Models, Mechanisms and Coherence. *British Journal for the Philosophy of Science*, doi: 10.1093/bjps/axt043.
Craver, C. F. (2007). *Explaining the Brain*. Oxford: Oxford University Press.
Craver, C. (2014). The Ontic Account of Scientific Explanation. In Kaiser, M.I., Scholz, O.R., Plenge, D., and Hüttemann, A. (eds), *Explanation in the Special Sciences: The Case of Biology and History*, Dordrecht: Springer.
Crupi, V. and Tentori, K. (2012). A Second Look at the Logic of Explanatory Power (with Two Novel Representation Theorems). *Philosophy of Science*, 79:365-385.
Cummins R. (1975). Functional Analysis. *Journal of Philosophy*, 72:741-765.
Dawid, R., Hartmann, S., and Sprenger, J. (2014). No-Alternative-Arguments. *British Journal for the Philosophy of Science*, doi: 10.1093/bjps/axt045.
de Donato-Rodriguez, X. and Zamora-Bonilla, J. (2009a). Credibility, Idealisation, and Model Building: An Inferential Approach. *Erkenntnis*, 70(1):101-118.
de Donato-Rodriguez, X. and Zamora-Bonilla, J. (2009b). Explanation and Modelization in a Conprehensive Inferential Account. In *EPSA09: 2nd Conference of the European Philosophy of Science Association* (Amsterdam, 21-24 October, 2009).
Douven, I. (2011). Abduction. In Zalta, E. N. (ed.), *The Stanford Encyclopedia of Philosophy*.
Dowe, P. (2000). *Physical Causation*. Cambridge: Cambridge University Press.
Dummett, M. (1991). *The Logical Basis of Metaphysics*. Harvard University Press.
Fagiolo, G., Moneta, A., and Windrum, P. (2007). A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems. *Computational Economics*, 30:195-226.
Faye, J. (2007). The Pragmatic-rhetorical Theory of Explanation. In Persson, J. and Ylikoski, P. (eds), *Rethinking Explanation*, Dordrecht: Springer, pp. 43-68.
Fitelson, B. (2003). A Probabilistic Measure of Coherence. *Analysis*, 63: 194-199.
Fitelson, B. and Hitchcock, C. (2011). Probabilistic Measures of Causal Strength. In Illari, P., Russo, F., and Williamson, J. (eds), *Causality in the Sciences*, pp. 600-627. Oxford: Oxford University Press.
Friedman, M. (1974). Explanation and Scientific Understanding. *The Journal of Philosophy*, 71(1):5-19.

Garfinkel, A. (1981). *Forms of Explanation: Rethinking the Questions of Social Theory*. Yale University Press.

Glennan, S. (2002). Rethinking Mechanistic Explanation. *Philosophy of Science*, 69(3):S342-S353.

Gopnik, A. and Schulz, L. (eds) (2007). *Causal Learning. Psychology, Philosophy, and Computation*. Oxford: Oxford University Press.

Halpern, J. and Hitchcock, C. (2010). Actual Causation and the Art of Modeling. In Dechter, R., Geffner, H. and Halpern, J. (eds), *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, pp. 383-406. London: College Publications.

Harman, G. H. (1965). The Inference to the Best Explanation. *Philosophical Review*, 74:88-95.

Harman, G.H. (1999). *Reasoning, Meaning and Mind*. Oxford: Clarendon Press

Hartmann, S. and Sprenger, J. (2010). Bayesian Epistemology. In Bernecker, S. and Pritchard, D. (eds), *The Routledge Companion to Epistemology*, chapter 55. London: Routledge.

Hempel, C. G. (1962). Two Basic Types of Scientific Explanation. Reprint. In Cover, J. A. and Curd M. (1998). *Philosophy of Science: The Central Issues*. New York: W. W. Norton & Company, pp. 685-694.

Hempel, C. G. (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.

Hempel, C. and Oppenheim, P. (1945). A Definition of "Degree of Confirmation". *Philosophy of Science*, 12(2):98-115.

Hempel, C. G. and Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, 15(2):135-175.

Hilton, D. J. and Erb, H. (1996). Mental Models and Causal Explanation: Judgments of Probable Cause and Explanatory Relevance. *Thinking and Reasoning*, 2:273-308.

Hitchcock, C. and Knobe, J. (2009). Cause and Norm. *The Journal Of Philosophy*, 106(11):587-612.

Howson, C. and Urbach, P. (2006). *Scientific Reasoning: The Bayesian Approach*. La Salle: Open Court.

Huneman, P. (2010). Topological Explanations and Robustness in Biological Sciences. *Synthese*, 177(2):213-245.

Huneman, P. (2013) (ed.). *Functions: Selection and Mechanisms*. Dordrecht: Springer.

Kirman, A. (2010). The Economic Crisis is a Crisis for Economic Theory. *CESifo Economic Studies*, 56(4):498–535.

Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science*, 48(4):507-531.

Korb, K., Nyberg, E. and Hope, L. (2011). A New Causal Power Theory. In Illari, P., Russo, F., and Williamson, J. (eds), *Causality in the Sciences*, pp. 628-652. Oxford: Oxford University Press.

Kuhlmann, M. (2011). Mechanisms in Dynamically Complex Systems. In Illari, P., Russo, F., and Williamson, J. (eds), *Causality in the Sciences*, pp. 880-906. Oxford: Oxford University Press.

Kuorikoski, J., Lehtinen, A., and Marchionni, C. (2010). Economic Modelling as Robustness Analysis. *British Journal for the Philosophy of Science*, 61:541-567.

Lange, M. (2013). What Makes a Scientific Explanation Distinctively Mathematical? *British Journal for the Philosophy of Science*, 64:485-511.

LeBaron, B. D., Arthur, W. B., and Palmer, R. G. (1999). Time Series Properties of an Artificial Stock Market. *Journal of Economic Dynamics and Control*, 23:1487-1516.

Lewis, D. (1986). Causal Explanation. In D. Lewis, *Philosophical Papers*, vol. II, pp. 214-240. Oxford: Oxford University Press.

Lipton, P. (2004). *Inference to the Best Explanation*. London: Routledge.

Lombrozo, T. (2010). Causal–explanatory Pluralism: How Intentions, Functions, and Mechanisms Influence Causal Ascriptions. *Cognitive Psychology*, 61:303-332.

Lombrozo, T. and Carey, S. (2006). Functional Explanation and the Function of Explanation. *Cognition*, 99:167-204.

Lux, T. and Marchesi, M. (1999). Scaling and Criticality in a Stochastic Multiagent Model of a Financial Market. *Nature*, 397:498-500.

Lux, T. and Marchesi, M. (2000). Volatility Clustering in Financial Markets: A Microsimulation of Iinteractive Agents. International *Journal of Theoretical and Applied Finance*, 3(4):675-702.

Machamer, P., Darden, L., and Craver, C. (2000). Thinking about Mechanisms. *Philosophy of Science*, 67:1-25.

Machamer, P. K. (2004). Activities and Causation: The Metaphysics and Epistemology of Mechanisms. *International Studies in the Philosophy of Science*, 18(1):27-39.

Makinson, D. (2005). *Bridges from Classical to Nonmonotonic Logic*. London: King's College Publications.

Maturana, H. and Varela, F. (1980). *Autopoiesis and Cognition: the Realization of the Living*. Dordecht: D. Reidel Publishing Co.

McLaughlin P. (2001). *What Functions Explain: Functional Explanation and Self-Reproducing Systems*. Cambridge: Cambridge University Press.

Millikan R.G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge: MIT Press.

Mossio, M., Saborido, C., and Moreno, A. (2009). An Organizational Account of Biological Functions. *British Journal for the Philosophy of Science*, 60(4):813-841.

McGrew, T (2003). Confirmation, Heuristics, and Explanatory Reasoning. *British Journal for the Philosophy of Science* 54:553-567.

Myrvold, W. C. (2003). A Bayesian Account of the Virtue of Unification. *Philosophy of Science* 70:399-423.

Nagel, E. (1961). *The Structure of Science*. New York: Hartcourt Brace.

Nerlich, G. (1979). What Can Geometry Explain? *British Journal for the Philosophy of Science*, 30:69–83.

Norton, J. D. (2003). A Material Theory of Induction. *Philosophy of Science*, 70(4):647-670.

Okasha, S. (2000). Van Fraassen's Critique of Inference to the Best Explanation, *Studies in History and Philosophy of Science*, 31:691-710.

Olsson, E. J. (2002). What is the Problem of Coherence and Truth? *The Journal of Philosophy*, 99:246–272.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.

Pearl, J., and Dechter, R. (1996). Identifying Independencies in Causal Graphs with Feedback. *Proceedings of the twelfth conference on Uncertainty in artificial intelligence*, pp. 420-426, San Mateo, CA: Morgan Kaufmann.

Peirce, C. (1931). *The Collected Papers Vol. I: Principles of Philosophy*. Cambridge: Harvard University Press.

Peregrin, J. (2006). Meaning as Inferential Role. *Erkenntnis*, 64:1-36.

Prawitz, D. (2006). Meaning Approached Via Proofs. *Synthese*, 148(3):507-524.

Reiter, P. (1980). A Logic for Default Reasoning. *Artificial Intelligence*, 13:81-132

Reiss, J. (2011). Third Time's a Charm: Causation, Science and Wittgensteinian Pluralism. In Illari, P., Russo, F., and Williamson, J. (eds), *Causality in the Sciences*, pp. 907-927. Oxford: Oxford University Press.

Reiss, J. (2012). Causation in the Sciences: An Inferentialist Account. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(4):769-777.

Rickles, D. (2011). Econophysics and the Complexity of Financial Markets. In Collier, J. and Hooker, C. (eds), *Handbook of the Philosophy of Science, vol. 10: Philosophy of Complex Systems*, pp. 527-561. North Holland: Elsevier.

Ruben, D. (1992). *Explaining Explanation. Problems of Philosophy*. London: Routledge.

Salmon, W. C. (1971). *Statistical Explanation & Statistical Relevance*. University of Pittsburgh Press.

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.

Salmon, W. (1989). *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.

Salmon, W. C. (2005). *Reality and Rationality*. Oxford: Oxford University Press.

Schupbach, J. N. (2011a). Comparing Probabilistic Measures of Explanatory Power. *Philosophy of Science*, 78(5):813-829.

Schupbach, J. N. (2011b). New Hope for Shogenji's Coherence Measure. *British Journal for the Philosophy of Science*, 62(1):125-142.

Schupbach, J. N. and Sprenger, J. (2011). The Logic of Explanatory Power. *Philosophy of Science*, 78(1):105-127.

Shogenji, T. (1999). Is Coherence Truth-conducive? *Analysis*, 59:338-345.

Shogenji, T. (2005). The Role of Coherence of Evidence in the Non-Dynamic Model of Confirmation. *Erkenntnis,* 63(3):317-333

Sellars, W. (1953). Inference and Meaning. *Mind*, 62(247):313-338.

Sober, E. (1983). Equilibrium Explanation. *Philosophical Studies*, 43 (2):201-210.

Spirtes, P. (1995). Directed Cyclic Graphical Representations of Feedback Models. In P. Besnard & S. Hanks (eds), *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, pp. 491-498, San Francisco: arXiv.org.

Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, Prediction, and Search*. MIT Press, Cambridge MA, second (2000) edition.

Spohn, W. (1993). Causal Laws are Objectifications of Inductive Schemes. In Dubucs, J. (ed.), *Philosophy of Probability*, pp. 223-252. Dordrecht: Kluwer.

Spohn, W. (2012). *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press.

Steinberger, F. (2011). What Harmony Could and Could Not Be, *Australasian Journal of Philosophy*, 89(4):617-639.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.

Thagard, P. (1989). Explanatory Coherence. *Behavioral and Brain Sciences*, 12:435-502.

Thagard, P. and Verbeurgt, K. (1998). Coherence and Constraint Satisfaction. *Cognitive Science*, 22(1):1-24.

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.

van Fraassen, B. (1988). The Pragmatic Theory of Explanation. In Pitt, J. (ed.), *Theories of Explanation*, pp. 51-74. Oxford: Oxford University Press.

van Fraassen, B. (1989). *Laws and Symmetry*. Oxford: Oxford University Press.

Walton, D. (2004). A New Dialectical Theory of Explanation. *Philosophical Explorations*, 7:71-89.

Weber, M. (2005a). Holism, Coherence and the Dispositional Concept of Functions, *Annals of the History and Philosophy of Biology*, 10:189-201.

Weber, M. (2005b). *Philosophy of Experimental Biology*, Cambridge: Cambridge University Press.

Weber, M. (2008). Causes without Mechanisms: Experimental Regularities, Physical Laws, and Neuroscientific Explanation. Philosophy of Science, 75: 995-1007.

Weber, M. (2009). Crux of Crucial Experiments. British Journal for Philosophy of Science 60(1): 19-49.

Weisberg, J. (2009). Locating IBE in the Bayesian Framework, *Synthese*, 167:125-143.

Wheeler, G. (2009). Focused Correlation and Confirmation. *The British Journal for the Philosophy of Science*, 60(1), pp. 79-100.

Wheeler, G. and Scheines, R. (2013). Coherence and Confirmation Through Causation. *Mind*, doi: 10.1093/mind/fzt019.

Williamson, J. (2005). *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press.

Williamson, J. (2013). How Can Causal Explanations Explain? *Erkenntnis*. 78:257-275.

Wittgenstein, L. (1956). *Bemerkungen über die Grundlagen der Mathematik*. English Trans. *Remarks on the Foundations of Mathematics*, 1978, Oxford: Basil Blackwell.

Woodward, J. (2003). *Making Things Happen. A Theory of Causal Explanation*. New York: Oxford University Press.

Wright, C.D. (2012). Mechanistic Explanation Without the Ontic Conception. *European Journal of Philosophy of Science*, 2: 375-394.

Wright, L. (1973). Functions. *Philosophical Review,* 82:139-168.

Zamora-Bonilla, J. (2006). Science as a Persuasion Game: An Inferentialist Approach. *Episteme*, 2:189-201.

Zamora-Bonilla, J. and de Donato-Rodriguez, X. (2013). Scientific Explanation and Representation: An Inferentialist Viewpoint. In Agazzi, E. (ed.), *Representation and Explanation in the Sciences*, pp. 81-93. Milano: Franco Angeli.