Rethinking Explanation

Persson, Johannes; Ylikoski, Petri

2007

CONTENTS

CONTRIBUTORS

Alexander Bird (University of Bristol)
Jan Faye (University of Copenhagen)
Olav Gjelsvik (University of Oslo)
Bengt Hansson (Lund University)
Henrik Hållsten (Stockholm University)
Mika Kiikeri (University of Tampere)
Tomi Kokkonen (University of Helsinki)
Jaakko Kuorikoski (University of Helsinki)
Johannes Persson (Lund University)
Stathis Psillos (University of Athens)
Rebecca Schweder (Lund University)
Jeroen Van Bouwel (Ghent University)
Robin Stenwall (Lund University)
Annika Wallin (SCAS and Lund University)
Erik Weber (Ghent University)
Petri Ylikoski (University of Helsinki)

ACKNOWLEDGEMENTS

We want to express our gratitude to Matti Sintonen. He has supported this project in so many ways. The editors also want to thank NOS-H for funding the project.

PREFACE

This volume is a product of the international research project *Theory of Explanation*, which was funded by the *Joint Committee for Nordic Research Councils for the Humanities and the Social Sciences* (NOS-HS). The project started in 2001 and operated for a period of three years by organizing a number of workshops on scientific explanation in Norway, Iceland, Sweden and Finland. The workshops included presentations by people involved in the project and by invited guests. Both groups are represented in this volume, which brings together some of the papers presented in these meetings.

The central theme of the research project was scientific explanation, but it was approached from many different angles. This plurality of approaches is also visible in the present volume. The authors share a joint interest in explanation, but not the same theoretical or methodological assumptions. As a whole, this volume shows that, although the theory of explanation has been a major industry within philosophy of science, there are still both conceptual problems to be solved and fresh philosophical ideas to explore.

The papers in this volume have been divided into two broad groups. Part 1 consists of papers dealing with general issues in the theory of explanation, while the papers in Part 2 focus on some more specific problems.

Part 1: Theory of Explanation

The book opens with Bengt Hansson's chapter on what explanation is and is not. He argues that an explanation is not a logical structure¸ and that it cannot be characterized in syntactic terms. It is rather an epistemological structure, a structure organising conceptual content. In order to show this, Hansson begins with a simple example and develops his argument by systematically exploring the effects of making the premises more general or more specific, thereby exposing the implicit suppositions and consequences of some commonly held views.

In the next chapter, Henrik Hållsten argues that any theory of explanation must do justice to the intuition that a putative explanation can be irrelevant in at least two ways: relative to the epistemic situation and due to objective irrelevance. He argues that the latter sort of irrelevance cannot be relativized to the epistemic situation and that this poses an important, but often neglected, challenge for many theories of explanation.

In chapter three, Petri Ylikoski discusses the idea of contrastive *explanandum*. He begins by presenting the intuitive idea behind the contrastive questions, and suggests a novel way to see the difference between scientific and everyday explanatory questions. He argues that all *explananda* can be analysed as contrastive and that this is a fruitful approach in making explanatory questions more explicit. In the latter part of the paper he defends the contrastive view against the major criticisms presented against it.

Jan Faye focuses on pragmatic theories of explanation. According to him, these theories make a problematic association between the relevance and the correctness of an explanation. An answer to an explanation-seeking question can be false, but may still be relevant, and therefore be an explanation. Faye suggests a pragmatic-rhetorical theory of explanation, where an explanation is grounded in a rhetorical practice of raising questions and answering them within a certain discourse. He also argues that his view has an interesting implication, according to which the distinction between scientific and everyday explanations turns out to be arbitrary.

In chapter five, Olav Gjelsvik examines David Lewis's thesis that to explain an event is to provide some information about its causal history. According to Gjelsvik this requirement is too weak to provide a fruitful starting point for a theory of explanation. He then proceeds to develop an account of his own by considering a number of suggestions to strengthen Lewis's account of causal relevance.

Stathis Psillos discusses causal explanation too. His focus is on the manipulationist account of causal explanation developed by James Woodward. According to this theory, *c* causally explains *e* on condition that, if *c* were to be (actually or counterfactually) manipulated, *e* would change too. Psillos discusses the role of laws in explanations of this kind. He claims that the very possibility of experimental counterfactuals requires that the laws are more than just invariances among magnitudes.

In chapter seven, Erik Weber and Jeroen van Bouwel defend the view that the criteria by which the explanatory value of explanations is to be judged are context-dependent. Explanation-seeking questions have different motivations and the criteria for evaluating the explanatory power depend on the motivation. Weber and van Bouwel argue that in most contexts *a posteriori* probability is important. They also argue that explanatory depth (going back further and further in time) is relevant in only one context. In most contexts there are further explanatory criteria (e.g. familiarity or manipulability), but interestingly enough unification is not among these.

In the last chapter of the Part 1, Rebecca Schweder argues that there are at least three different ways to understand the notion of probabilistic explanation. She proposes that these differences reflect real metaphysical differences. According to

Schweder, this constitutes a problem for the unificationist account of explanation, since not all three would count as explanations.

Part 2: Issues in explanation

Alexander Bird starts Part 2 by examining selection explanations. He argues that familiar raven cases provide counterexamples to the D-N model of explanation, not only to the hypothetic-deductive model of confirmation. Thus not everything deducible from a law is explained by that law. In selection cases the law operating is often negative, and the contrapositives of such laws do not explain their instances. However, the physical sciences tend not to focus on such negative general truths but on their (particularizable) contrapositives. What makes selection explanations different, Bird argues, is their use of negative properties. Consequently, in selection explanation it is the negative generalization that has a particularizable explanation while it is the logically equivalent positive generalization that is non-particularizable. Selection explanations, Bird concludes, do not appear fundamentally different from other nomological explanations.

Johannes Persson looks into the relations between explanation and inference. Sometimes explanation clearly comes before inference (EBI). But are such cases also instances of Inference to the best explanation (IBE)? Persson proceeds to answer this question by examining some of the limiting cases of IBE—such as when no hypothesis meets the minimal requirement for being an explanation, or where we have competing concepts of explanation, or where the possibility of an IBE coexists with the applicability of other inferential tools. Persson finds that, sometimes, essentially explanatory considerations nevertheless belong to the early phases of such inquiries, for instance involving concept formation. Later phases of such enterprises merely appear to be governed by IBE.

According to Jaakko Kuorikoski, equilibrium explanations, which are quite pervasive in economics and biology, have not received the philosophical attention they deserve. Elliot Sober once claimed that equilibrium explanations constitute a counter-example to "the causal thesis of explanation", because they do not require the exposition of the actual causal history of the event to be explained. Kuorikoski shows that this claim is mistaken and based on the confusion about relation between causal and structural explanations. On a closer view, Kuorikoski argues, the explanatory uses of equilibrium models require that the models pick out causally relevant properties of the system.

In the next chapter, Annika Wallin investigates the explanatory role of the environment in psychology and the cognitive sciences. She begins by considering different types of adaptation to the environment in psychology. Depending on the notion of adaptation adopted, different sorts of explanations of psychological phenomena become viable. Wallin describes how psychologists use methodological shortcuts to avoid some of the most difficult conceptual questions related to the differences between these notions of adaptation. For instance, using representative sampling (i.e. making sure that the environment in which measures are taken is typical for the everyday life of participants) gives psychologists an idea of how a cognitive process functions in its normal environment, even when they are unable to

characterise this environment more precisely. She concludes by discussing how psychologists can rely on such methodological shortcuts when explaining behaviour, and what their environment-oriented explanations require to be theoretically solid.

Mika Kiikeri and Tomi Kokkonen start by comparing the empiricist and nativist accounts of language acquisition. Nativist explanations of language acquisition have assumed that all details that cannot be accounted for on the cognitive level are innate, based on the biological properties of an organism. At the same time, the details of biological implementation are left to future research. Innateness has thus remained as a kind of explanatory black-box for linguistics and cognitive science. Kiikeri and Kokkonen point out that the concept of innateness has somewhat different explanatory roles in biological and psychological contexts, making the assumption of neat disciplinary division of labour untenable. However, they argue that an adequate biological analysis of psychological innateness could still help to re-evaluate the role of innate information in language acquisition.

This book ends with Robin Stenwall's interesting study on explanations involving natural kinds. These explanations can be structurally complicated. For instance, in explaining why zinc sulphate is easily soluble in water, we make references to solubility being a characteristic of metallic *sulphates*, but when an explanation of the positive effect of zinc sulphate on the immune system is called for, we may point out that it is a *zinc* compound. Stenwall argues that when we explain in this manner, we are almost always doing this in relation to some contingent property of the natural kinds involved. Aspect kinds are kinds defined in terms of their contingent properties.

Johannes Persson and Petri Ylikoski