THE ROLE OF UNIFICATION IN EXPLANATIONS OF FACTS

Erik Weber, Jeroen Van Bouwel & Merel Lefevere[1]

Centre for Logic and Philosophy of Science
Ghent University (UGent)
Blandijnberg 2
B-9000 Gent
Belgium
Phone: +32-9-2643785
Fax: +32-9-2644187
Erik.Weber@UGent.be
Jeroen.VanBouwel@UGent.be
Merel.Lefevere@UGent.be

Jeroen Van Bouwel is a Postdoctoral Fellow of the Research Foundation (FWO) Flanders.

[1] The authors thank Raoul Gervais, Jan Willem Wieland and the audience at EPSA09 for their comments on earlier versions of this paper. Jeroen Van Bouwel is a Postdoctoral Fellow of the Research Foundation (FWO) Flanders

THE ROLE OF UNIFICATION IN EXPLANATIONS OF FACTS

Abstract
In the literature on scientific explanation, there is a classical distinction between explanations of facts and explanations of laws. This paper is about explanations of facts. Our aim is to analyse the role of unification in explanations of this kind. We discuss five positions with respect to this role, argue for two of them and refute the three others.

Keywords
Aim of explanation, causation, explanatory virtues, unification.

1. Introduction

In the literature on scientific explanation, there is a classical distinction between explanations of facts and explanations of laws. This paper is about explanations of . Our aim is to analyse the role of unification in explanations of this kind. With respect to this role, five positions can be distinguished:
(1) There is no place for unification in the explanatory process of singular facts. This view is defended in Halonen & Hintikka (1999). They do not deny the value of unification, but they locate this value outside the explanatory process; unification is a criterion for theory choice.
(2) There are cases in which unification is an explanatory virtue, while the aim of explanation is not unification. In these cases the *aim* of explanation is to provide information about the causes of the explanandum, but it is *better* or even *required* to do this in a unified way.
(3) Unification is an explanatory virtue – in the sense specified in (2) – in *all* cases of singular explanation.
(4) There are cases in which unification is the main aim of singular explanation (as opposed to merely an explanatory virtue).
(5) Unification is *always* the main aim of singular explanation (this is the position of Kitcher 1981 and 1989).
In Section 3, we argue that (2) is correct. A corollary of this is that (1) is refuted. Starting from (2), there are two ways to ascribe a stronger role to unification. In Section 4, we argue that the first way, generalising from *some* to *all*, does not work. In other words, in Section 4 we argue against position (3). The other way, shifting from unification as an explanatory virtue to unification as the main aim of explanation, is discussed in Section 5. There we argue that position (4) is correct, while (5) is wrong. In Section 6 we show that two different types of unification were involved in our argument: the type of unification that is considered in Section 3 is not the same as the type considered in Section 5. Before we engage in all of this we have to clarify what we mean with causation, causal explanation and unification (Section 2).

2. Causation, causal explanation and unification

2.1 We will adopt Ronald Giere's probabilistic definitions of causation. (1997, p. 204):

C is a *positive causal factor* for **E** in the population U whenever $\mathbf{P}_X(E)$ *is greater than* $\mathbf{P}_K(E)$.
C is a *negative causal factor* for **E** in the population U whenever $\mathbf{P}_X(E)$ *is less than* $\mathbf{P}_K(E)$.
C is *causally irrelevant* for **E** in the population U whenever $\mathbf{P}_X(E)$ *is equal to* $\mathbf{P}_K(E)$.

Though it can be extended to other types of variables, Giere considers only binary variables. So in his definitions, **C** is a variable with two values (C and Not-C); the same for **E** (values E and Not-E). X is the hypothetical population which is identical to U, except that each individual exhibits the value C of the causal variable **C**. K is the analogous hypothetical population in which all individuals exhibit ¬C.

An example might clarify this. If we claim that smoking (**C**) is a positive causal factor for lung cancer (**E**) in the Belgian population (U), this amounts to claiming that if every inhabitant of Belgium were forced to smoke there would be more lung cancers in Belgium than if everyone were forbidden to smoke. Conversely for the claim that smoking is a negative causal factor. Causal irrelevance is a relation between variables (represented in bold) rather than a relation between values of a variable (like the first two relations). If we claim that *smoking behaviour* (**C**) is causally irrelevant for *the occurrence or absence of lung cancer* (**E**) this means that we believe that in the two hypothetical populations the incidence of lung cancer is equally high.

2.2 A causal explanation is an explanation which somehow refers to the causes of the event to be explained. In our view not much can be said about the general structure of causal explanations, because they can have very different formats. This view is developed and defended at length in Weber, Van Bouwel & Vanderbeeken (2005) and Weber & Van Bouwel (2007). Here we confine ourselves to one point, viz. the fact that we can ask different explanation-seeking questions about facts. Suppose we have observed that John has stolen a bike and Peter has stolen a CD-player. We can ask why John stole a bike rather than something else. Or we can ask why John stole something rather than nothing. These are examples of contrastive questions. The importance of contrastive questions has been stressed by e.g. Bas van Fraassen (1980) and Peter Lipton (1990, 1993). We can also ask why John stole a bike. This is an example of a question about a plain fact. These questions are at the focus of most classical accounts of explanation, including Hempel (1965), Salmon (1984) and Kitcher (1981,1989). We can also ask why both John and Peter stole something. This is an example of a comparative question. These questions are largely neglected in the literature on explanation (we will discuss them in Section 5).

The fact that we can ask different types of questions about facts is one of the reasons why causal explanations have different formats and thus that no general account of their structure can be given. However, there are three requirements that every causal explanation must satisfy. The first is trivial: a causal explanation must refer to the causes of the explanandum, i.e. claim to provide causes of the explanandum in some way. The second requirement is that causal explanations must be accurate in the facts they use: an explanation is acceptable only if the events it claims to have happened, really took place. This second requirement can be called “factual accuracy”. The third requirement is “causal accuracy”. A causal explanation must use causal knowledge which is backed up by sufficient (scientific or other) evidence. For instance, if we explain the collapse of a bridge by referring to an earthquake, this explanation is acceptable only if the earthquake really took place and if it is established that bridges of the type considered collapse when exposed to earthquakes of a certain strength.

2.3 The term *unification* will be used in a broad sense here. Kitcher (1981) claims that besides what he calls Hempel's "official" position with respect to what understanding is (in which understanding is identified with *expectability*, see e.g. Hempel 1965, p. 337), there is an "unofficial" one:

> What scientific explanation, especially theoretical explanation, aims at is not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting the phenomena as manifestations of common underlying structures and processes that conform to specific, testable basic principles. (Hempel 1966, p. 83; quoted in Kitcher 1981, p. 508).

Kitcher ascribes to Hempel the view that, besides expectability, explanations can confer a second intellectual benefit upon us: unification. Whether or not this ascription is correct does not matter here. What we need here is the idea of unification: unifying events consists in showing that two or more different events/phenomena are manifestations of common underlying structures and processes. This broad idea can be elaborated in different ways. These will be discussed in Section 6.

3. Unification as an explanatory virtue

3.1 We start with an elaborate example. The material we use is taken from an article of Michael Taylor on revolutionary collective action (Taylor 1988) which discusses Theda Skocpol's classic *States and Social Revolutions* (Skocpol 1979). By using comparative methods, Skocpol has formulated a so-called "structural" explanation for three successful modern social revolutions in agrarian-bureaucratic monarchies (the French, Russian and Chinese revolution). The structural conditions that, in her view, make a revolution possible (the revolutions can be successfully mounted only if these structural preconditions are met), relate to the incapacitation of the central state's machineries, especially the weakening of the state's repressive capacity. This weakening is caused by external military (and economic) pressure: because of the backward agrarian economy and the power of the landed upper class in the agrarian-bureaucratic monarchy, the attempt to increase the military power leads to a fiscal crisis. Escalating international competition and humiliations particularly symbolized by unexpected defeats in wars (which inspired autocratic authorities to attempt reforms) trigger social revolutions. The foreign military and economic pressure that triggered the respective social revolutions, were:
(1) Bourbon *France* (1787-89): financially exhausted after the War for American Independence and because of the competition with England in general.
(2) Manchu *China* (1911-16): the Sino-Japanese War (1895) and the Boxer debacle (1899-1901).
(3) Romanov *Russia* (1917): massive defeats in World War I.

Skocpol's theory gives adequate answers to several contrastive questions, for instance:

> Why did the French revolution start in 1789, rather than in 1750?

The answer is that the pressure was not big enough in 1750.
However, in this section we focus on non-contrastive questions, for instance:

> Why was there a revolution in Bourbon France, Manchu China *and* in Romanov Russia?

This is a request for the explanation of a set of similar facts. If we ask questions like this, the underlying aim is a type of unification: we want to know what is *common* in the causal ancestry of the different phenomena. In the example, we want to know what the causal mechanisms that led to these three revolutions have in common. Skocpol gives a part of the answer: She endorses the following principle:

> External military/economic pressure is a necessary cause[2] of social revolutions.

According to Michael Taylor there is another causal factor which the three revolutions have in common, viz. a strong sense of community among the peasants:

> When the peasant community was sufficiently strong, then, it provided a social basis for collective action, including revolutionary collective action and rebellions and other popular mobilizations. (1988, p. 68)

Taylor shows how the participation of vast numbers of peasants in collective action could be explained by using *the logic of collective action* advanced by Mancur Olson – which implies the use of economic incentives and selective social incentives, because without incentives to motivate participation, collective action is unlikely to occur even when large groups of people with common interests exist – adding his own theory of conditional cooperation. Using this account of collective action, Taylor argues that peasant collective action in revolutions was based on community (as many historians have argued) and that this is mainly why the large numbers of people involved were able to overcome the free-rider problem familiar to students of collective action and opted for conditional cooperation.

Taylor's idea can be summarized in the following principle:

> A strong sense of community is a necessary cause for social revolutions to occur.

This does not contradict Skocpol's principle: they are different but compatible claims about factors that occur in the causal ancestry of all social revolutions (see also Van Bouwel & Weber 2008).

3.2 The question we have considered in Section 3.1 is of the following type:

(U) Which factors occur in the causal ancestry of all the events x_1, ..., x_n.

We call this U-type questions because they ask for factors that are common in the causal ancestry and therefore *unify* the mechanisms that led to the events. Note that U-type questions are non-contrastive. When answering U-type questions our aim is to provide information about the causes of the explanandum. This follows from the way we have characterised them: they ask for factors in the causal ancestry of events. That unification is a desideratum in answering U-type questions also follows from the way we have characterised them: we are looking for factors that are common in the causal

[2] A necessary cause is a positive causal factor were $\mathbf{P}_K(E)=0$

ancestry of the events involved. This means that *if* scientists regularly ask U-type questions and succeed in answering them, there are indeed contexts in which the aim of explanation is to provide information about causes. Reaching this aim requires that the explanation has unificatory power. Our example in Section 3.1 shows that the condition is satisfied: scientists ask U-type questions and try to answer them. So we have an argument for claim (2): There are cases in which the *aim* of explanation is to provide information about the causes of the explanandum, but it is *better* or even *required* to do this in a unified way.

3.3 Let us go back briefly to the overall structure of our argument. In this section we have argued that position (2): *There are cases in which unification is an explanatory virtue, while the aim of explanation is not unification* (cf. the Introduction) is correct. A corollary of this is that position (1): *There is no place for unification in the explanatory process of singular facts,* has been refuted. Starting from position (2), there are two ways to ascribe stronger role to unification. These ways are explored in Sections 4 and 5.

4. Causal explanations without unification

4.1 In this section we argue against the claim that unification is an explanatory virtue in *all* singular explanations, i.e. against the universally generalised variant of the existentially quantified claim we defended in Section 3. We try to show that there are cases in which the questions scientists are trying to answer *do* require that the explanation is causal, but *do not* require that the explanation has unificatory power.

Before we start, we have to introduce some terminology. We will refer to two types of questions in this section:

(I) Why does *x* have property P, rather than the ideal property P′?
(I′) Why does *x* have property P, while *y* has the ideal property P′?

P and P′ are mutually exclusive properties. An I-type question compares an actual fact with an ideal one (one we would like to be the case). For instance, we can try to explain why only 61 % of the Belgian population (between the age of 15 and 65) was working in 2006, while the ideal put forward by the European Union is 70%. An I′-type question does basically the same, but a different object in which the ideal situation is realised is used to emphasise that the ideal is not unrealistic. For instance, we can try to explain why only 61% of the Belgian population (between the age of 15 and 65) was working, while in the Netherlands 74,3% was working. The two types of questions we consider here are contrastive.

Contrastive questions of type (I) and (I′) are motivated by a therapeutic or preventive need: they request that we isolate causes which help us to reach an ideal state that is not realised now (therapeutic need) or to prevent the occurrence of similar events in the future (preventive need).

4.2 I- and I′- type questions as characterised in Section 4.1 can provide cases in which the questions scientists are trying to answer *do* require that the explanation is causal, but *do not* require that the explanation has unificatory power. We use a fictitious but realistic (because it is based on real causal knowledge) example. Two neighbouring cities, Koch City and Miasma City, have a history of simultaneous cholera epidemics: every ten years or so, after excessive rainfall, cholera breaks out in both cities. Suddenly, in the year X, the population of Koch City remains healthy after a summer with lots of rain, while Miasma City is hit by cholera again. Explaining the difference can help Miasma City in the future (therapeutic function).

Let us consider the following explanation of the contrast:

> There was a cholera outbreak in Miasma City because:
> (a) there was a lot of rainfall; and
> (b) Miasma City had no sewage system.
>
> There was no cholera outbreak in Koch City, despite the fact that
> (a′) there was a lot of rainfall, because
> (b′) Koch City started building a sewage system after the previous outbreak, and this system was ready now.

The explanation refers to a difference between the cities that is the result of a human intervention that was present in one case, but absent in the other. Moreover, the difference is causally relevant for the difference in the effect. These are the reasons why it can serve a therapeutic function. An answer to an I- or I′-type question is adequate only if (i) the difference that is singled out is in some way manipulable, and (ii) there is a causal relation between the difference that is singled out in the explanation and the difference that is to be explained. In the example the value of the explanation lies in the fact that Miasma City also could have built a sewage system, and that the sewage system is causally relevant for the outbreak of cholera. With respect to causal relevance, it is useful to repeat what we said in Section 2.2: a causal explanation is acceptable only if we have sufficient evidence for the causal claims it contains. In this case, the causal relevance of sewage systems should be established before giving the explanation, or be backed up with evidence after presenting it. The observed difference between the cities is certainly not enough to establish the causal relation (not all differences are causally relevant).

Manipulability and causal relevance are minimal conditions of adequacy for explanations in the context we are discussing here. However, high probability values are also important. If a sewage system is the only causally relevant factor (i.e. if cities with sewage system are never struck by cholera, cities without a sewage system always after a certain amount of rainfall), the explanation above is perfect: it describes the only possible therapy, and this therapy is 100% efficient. The value of an answer to an I- or I′-type question depends on manipulability and causal relevance, but also on the degree of efficiency and indispensability of the therapeutic measure it suggests.

One crucial question remains: is unification a desideratum in these contexts? Obviously not: the similarities between the cities (e.g. the heavy rainfall) are irrelevant in these contexts. We are interested in the difference in the causal ancestries, not in the factors they have in common.

4.3 Going back to our overall argument, we now have an argument against position (3): *Unification is an explanatory virtue in* some *but not in* all *cases of singular explanation.* In Section 5 we investigate whether unification can be the main aim of explanation.

5. Unification as the main aim of explanation?

Consider the following question:

Why do Peter and Mary both have blood group A?

Peter and Mary are two people randomly chosen in a population. There are various ways to interpret this question. The interpretation we consider here is this one:

By virtue of which law did Peter and Mary both acquire blood group A?

This question can be answered as follows:

Part A
(L_1) All humans who belong to category $I^AI^A \times I^AI^O$ have blood group A.
(P_1) Mary is a human and belongs to category $I^AI^A \times I^AI^O$.
(E_1) Mary has blood group A.
Part B
(L_2) All humans who belong to category $I^AI^A \times I^AI^O$ have blood group A.
(P_2) Peter is a human and belongs to category $I^AI^A \times I^AI^O$.
(E_2) Peter has blood group A.
Comparison
(L_1) = (L_2)

Here we show that there is a law according to which the first fact can be expected. Then we do the same for the second fact, using the same law. Finally we explicitly acknowledge that the laws are identical.

The general format of the question in our example is this:

(L) By virtue of which law did the similarity between E and E' occur?

When we are dealing with an L-type question, the aim of the explanation is to subsume the two events under the same law. If the law is a causal law, the initial conditions in the arguments that do the subsumption refer to causes of the explanandum. In that case the explanation is causal, though the aim of explanation was *not* to acquire information about the causes. Many people (e.g. Kitcher) have argued that the laws best suited for explanation of this type are always causal. In this way, they can save the intuition that all explanations are causal: some are causal because we want to have information about causes, some are causal because the best laws for subsumption are causal.

Going back to our overall argument, we now have an argument for position (4): *There are cases in which unification is the main aim of singular explanation (as opposed to merely an explanatory virtue).* The results of Sections 3 and 4 block a generalisation of this result (from *some* to *all*), so we have an argument against position (5*): Unification is* always *the main aim of singular explanation.*

6. Two types of unification

As mentioned in Section 2.3, unifying events consists in showing that two or more different events are manifestations of the same underlying structures and processes. Robert Skipper (1999) has argued that, besides Philip Kitcher's well known way to elaborate this idea, there is a second one:

> ... I have provided the foundations of an alternative to Kitcher's way of understanding explanatory unification. Kitcher claims that unification is the reduction of types of facts scientists must accept in expressing their world view, and it proceeds through derivation of large numbers of statements about scientific phenomena from economies of argument schemata. I suggest that it is very much worth exploring whether unification can be conceived as the reduction of types of mechanism scientists must accept as targets of their theories and explanations, and whether it proceeds through the delineation of pervasive causal mechanisms via mechanism schemata. (1999, pp. S207-S208)

We will call Kitcher-style unification "top-down unification", and Skipper-style unification "bottom-up unification". What is the difference? In top-down unification we show that the events to be unified are instances of the same (set of) law(s) of nature. In other words, top-down unification proceeds by subsuming different events under the same laws. Unification of this type is achieved by constructing arguments (one for each event) in which it is shown that the events could be expected (cf. Hempel's DN explanations) and in which the same laws are used. Bottom-up unification consists in showing that the mechanisms which lead to different events contain similar causal factors. This does not require subsumption under a law, so this kind of unification does not proceed by constructing arguments and showing that the events could be expected.

If we look back at the example from section 5 we see that it fits Hempel's scheme for a DN-explanation. The fact that Mary has blood type A is deductively derived from a general law and initial conditions. It is impossible for Mary to be human and belong to category I_AI_A x I_AI_O, but have another blood type than A. The same goes for Peter or any other human. If we come across a person who belongs to category I_AI_A x I_AI_O, but does not have blood type A, than there is a serious problem with the general law. The unification here works by showing that certain outcomes were to be expected, there is no room for alternative outcomes without degrading the law that is applied.

We will now take a look at the example from section 3. Skocpol's and Taylor's theory refers to certain structural conditions present that made the revolutions possible. The common factors in the explanation do not entail DN-expectability of the outcome. It is perfectly possible that there are societies who are under pressure and have a strong rural community, without going into revolution or to have revolutions that are founded on other factors. The common factors that provide the unification of the three mentioned revolutions are not sufficient for causing those revolutions. So we don't have top-down unification here. Instead we have bottom-up unification: the mechanisms leading to the events contain similar causal factors.

In general, in order to answer U-type questions such as "which factors occur in the causal ancestry of all the events x_1, ..., x_n" we need bottom-up unification. So bottom-up unification is the type of unification we encounter in cases like the ones discussed in Section 3: cases in which the aim of explanation is to provide information about the causes of the explanandum, but it is better or even required to do this in a unified way. In cases like the ones discussed in Section 5 (cases in which unification is the main goal of singular

explanation) we encounter top-down unification. This shows that the distinction between the two types is important.

7. Conclusions

In this paper we have used three distinctions. The first distinction was that between contrastive explanation seeking questions (which we used in Section 4) and comparative explanation seeking questions (which we used in Sections 3 and 5). This distinction enabled us to argue *against* position (1) and *for* (2), because unification plays a role when we deal with *comparative* questions. And it enabled us to argue *against* (3) and (5) because unification plays *no* role when we deal with *contrastive* questions.
The second distinction was that between two types of comparative questions: U-type questions as used in Section 3 and L-type questions as used in Section 5. This distinction enabled us to argue that there are cases where unification is the main aim of explanation (position (4)). Finally, we have used the distinction between bottom-up and top-down unification. This distinction allowed us to clarify what kind of unification is involved in answering U-type and L-type questions.

References

Giere, Ronald. 1997. *Understanding Scientific Reasoning* (4th edition). Fort Worth: Harcourt College Publishers.

Halonen, Ilpo and Jaakko Hintikka. 1999. Unification - It's Magnificent but is it Explanation? *Synthese* **120:**27-47.

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

Hempel, Carl. 1966. *Philosophy of Natural Science*. Englewood Cliffs: Prentice-Hall.

Kitcher, Philip. 1981. Explanatory Unification. *Philosophy of Science* **48**: 507-531.

Kitcher, Philip. 1989. Explanatory Unification and the Causal Structure of the World. In *Scientific Explanation,* eds. Kitcher & Salmon, 410-505. Minneapolis: University of Minnesota Press.

Lipton, Peter. 1990. Contrastive Explanations. In *Explanation and its Limits,* ed. D. Knowles, 247-266. Cambridge: Cambridge University Press.

Lipton, Peter. 1993. Making a Difference. *Philosophica* 51: 39-54.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*, Princeton, NJ: Princeton University Press.

Skipper, Robert A. 1999. Selection and the Extent of Explanatory Unification. *Philosophy of Science* **66**: S196-S209.

Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.

Taylor, Michael. 1988. Rationality and Revolutionary Collective Action. In *Rationality and Revolution*, ed. Michael Taylor, 63-97. Cambridge: Cambridge University Press.

Van Bouwel, Jeroen and Erik Weber. 2008. A pragmatic defence of non-relativistic explanatory pluralism in history and social science. *History and Theory* **47**: 168-182.

van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Oxford University Press.
Weber, Erik, Jeroen Van Bouwel and Robrecht Vanderbeeken. 2005. Forms of Causal Explanation. *Foundations of Science* **10**: 437-454.
Weber, Erik and Jeroen Van Bouwel. 2007. Assessing the Explanatory Power of Causal Explanations. In *Rethinking Explanation*, eds. Johannes Persson and Petri Ylikoski, 109-118. Dordrecht: Kluwer Academic Publishers.