Explanatory unification and conceptualization

Stefan Petkov
postdoctoral fellow of Tsingua University

Abstract

There are several important criticisms against the unificationist model of scientific explanation:
1. Unification is a broad and heterogeneous notion and is hard to see how a model of explanation based exclusively on unification can make distinction between genuine explanatory unification from cases of ordering or classification.
2. Unification alone could not solve the asymmetry and irrelevance problems.
3. Unification and explanation pull in different directions and should be decoupled, because for good scientific explanation an extra ad explanandum information is often required.

I'll present a possible solution by proposing to focus on an often overlooked but important element of how theoretic unification is achieved - the conceptual frameworks of theories. I'll propose that the core conceptual assumptions behind theories are decisive for discriminating between explanatory and non-explanatory unification. The conceptual framework is also flexible enough to balance the tension between informativeness and maximum systematization in constructing explanatory inferences.

Finally I'll show how my proposal could work by a short case study of orthogenetic and Darwinian explanations in paleontology.

Keywords: Scientific Explanation. Conceptualization. Conceptual Framework. Unification.

Introduction

The core assumption behind the unificationist model is very minimalistic, i.e., that there is a common ground between all forms of scientific explanation. That common ground is that scientific endeavor succeeds in providing understanding because it presents us with a coherent account of a wide range of phenomena by showing how singular facts fit together with other facts in a unified whole. This account is appealing because it could potentially satisfy both our philosophical and scientific intuitions. For one it builds up right on the pre-Socratic ideas about the cosmological union of the apparently diverse phenomena. In a more contemporary scientific spirit because unification can potentially account for explanations in biology, history, social sciences and mathematics; branches that rival models of explanation like the deductive nomological model and the strictly causal accounts cannot directly cover (Woodward, 2003; Salmon, 1984). Historically, unification has clearly played a major role in scientific advancement, because unified
theories often revealed connections between phenomena that were thought to be unrelated beforehand.

The unificationist approach, however, is challenged by two serious questions:

1. Is the notion of unification alone a sufficient condition for a good model of scientific explanation?
2. Is it possible to define unification in a manner that can distinguish explanatory unification from non-explanatory unification like ordering or classification?

I'll try to defend a position which answers affirmatively both questions. Typically the proponents of the unificationist model have analyzed scientific theories in purely syntactic terms, as sets of argumentative patterns (Kitcher, 1989), or models with a different degree of generality (Barenbolt, 2002; (Scurtz, 1999). However, if explanatory unification is analyzed solely as structural properties of explanatory arguments, two further problems arise:

i. Analyzing explanations as simple derivations from only a few argumentative patterns or instantiations of similar models does not automatically rule out explanatory asymmetry. There are just as many inferences from causes to effects as from effects to causes and both inferences are unifying (Barns, 1992).

ii. The unificationist model should satisfy two conflicting conditions for good explanations: (a) maximum systematization (which requires a sufficient level of abstraction), and (b) informativeness (which requires extra ad explanandum premises). This is apparent if we focus on explanatory patterns, an explanatory pattern riches maximal unification if it has as few as possible premises and if it's scope is as general as possible. However, such an argument is rarely considered scientifically interesting and informative.

To solve these problems I'll try to present a view in which explanatory unification is a multi-layered process which depends equally on the structure of argumentative patterns and on the basic theoretic elements (concepts).

I build on the view, proposed by Berght Hansson (2007) that the way we conceptualize the explanandum is essential in reaching a satisfying explanation. Thus we should not expect that the structure of the argumentative patterns or models can account completely for explanatory and non-explanatory inferences. This structure should only account for explicitly illuminating the relational links between the already conceptualized explanandum. The conceptual framework on the other hand serves a broader unificatory function and it is plastic enough to account for the systematization-informativeness tension. I will also suggest that the conceptual framework gives a preliminary suggestion which inferences should be counted as explanatory and which do not provide direct explanations.

In section one, I'll present the evolution of the unificationist model. I'll show that it focuses mainly on the structure of explanatory patterns and it defines good scientific explanations as systematization of beliefs, due to application of few broadly applicable and similar argumentative patterns.

In sections two to four, I'll present the main critical claims against the unificationist model. I'll propose that the issue they raise could be solved if we shift our focus from structural properties of theories, to conceptualization. I'll propose that conceptualization
is the main ingredient of explanatory unification, because the scope of application of explanatory patterns and the choice which inferences should be considered explanatory and which not, depends on how concepts are defined and employed within theories.

In section five, I'll present further requirements for the conceptual framework of theories. I'll propose that explanatory unification works, because conceptual frameworks present stable conceptual links that are made explicit in the explanans. More the general theoretic concepts themselves permit the constructions of potentially informatively extendable explanatory arguments. This extension is achieved by giving further information about the employed concepts and how they capture the salient features of the explanandum.

Finally in section six, I'll present a more substantial discussion by a short case study of explanatory unification from Paleontology.

1. The Explanatory unification model

The first account of unification as the basis of explanation has been made by Friedman (1974). He focuses on the question in virtue of what does scientific explanation provide understanding? The immediate answer given by the classical deductive nomological model is that explanation provides understanding because it subsumes the explanandum under a more general regularity given in the form of a scientific law. This however proved to be somewhat problematic, because it left the general laws themselves unexplained. This has lead Friedman to accept the idea that explanation is on one hand a reduction of the number of independent phenomena, by subsuming them under laws and on the other hand the reduction of the number of independent laws to a smaller set from which the old laws follow. As a result explanation provides understanding in virtue of unifying phenomena under progressively smaller number of more general regularities.

Two years later, Philip Kitcher, analyzed Friedman's work and found his conclusion too restrictive. The problem was that some forms of conjunctions of independent generalizations could potentially lead to a complete explanation of a system (Kitcher, 1976).

Kitcher, however, kept the main idea of Friedman and proposed that explanatory unification is achieved not by replacement of the large number of independent laws by a smaller number of more general laws, but by the repeated use of a smaller number of “type” laws which relate a large class of apparently diverse phenomena to a few fundamental magnitudes and properties. So, explanation is subsuming of a large set of diverse phenomena under a small number of similar general argumentative patterns (Kitcher, 1976). Here we might note that even in this rudimentary form the unificationist approach requires “unification” as a notion working on at least two levels: at the level of conceptualizing of “magnitudes and properties” and at the level of the general argumentative patterns.
About a decade later, Kitcher developed a more elaborate variant of the unificationist model (Kitcher, 1989). However, there he completely disregarded conceptualization and focused exclusively on the structure of argumentative patterns.

Kitcher adopted a strong inferentialist approach, defending the position that all scientific explanations are forms of deductive derivations. His main idea was that unificatory power of the deductive argument does not have to follow from the generality of the premise containing the law like sentence, but is a characteristic exhibited by the whole argumentative structure, which Kitcher coins as an argumentative pattern. This step permits an extension of the explanatory model beyond physics based explanations to non-law based scientific disciplines such as biology, history, social sciences etc.

To see how “unificatory power” is distributed to the whole structure of the argumentative pattern we have to go over the technical details of Kitcher's model. The other reason for scrutinizing his model more closely is that it's main features are inherited by later attempts at explanatory unification.

An argumentative pattern contains the following elements:

1. **A schematic sentence** which is proposition obtained by replacing some of the non logical expressions with dummy letters.
   *For instance:*
   “Organisms homozygous with non-functional melanin coding genes develop albinism”, is schematized as: “Organism homozygous for A alleles develop P”.
   “The mouse Robert has non-functional melanin coding genes.” as “M is an organism homozygous for A”
   “The mouse Robert develops albinism” as
   “M develops P”

2. **A set of filling instructions** for the schematic sentence which give directions for replacing the dummy letters with concepts within the scope of the theory.
   *In the discussed example the filling instructions might specify that A can be replaced by a name of an allele, P by the name of the phenotypic trait and M as an organism.*

3. **A schematic argument** is a sequence of schematic sentences.
   A1. “Organism homozygous for A alleles develop P”
   A2. “M is an organism homozygous for A”
   A3. “M develops P”

4. **A classification for a schematic argument** is an extra set of statements which describe the inferential characteristics of the schematic argument. It gives information which propositions are to be regarded as premises, which are inferred from which etc.
   A3 is to be inferred from A1 and A2.

Finally:
5. A general argument pattern is a triple consisting of a schematic argument, a set of filling instructions, one for each propositions of the schematic argument and a classification for the schematic argument.

The term Kitcher uses for the scope of an argumentative pattern is “stringency”. Naturally the stringency depends on how many restrictions the pattern imposes on the arguments it instantiates. Thus if we completely relax the stringency, the argumentative pattern degenerates to the general structure of any deductive argument. If we strengthen the stringency to the limit, the argumentative pattern will have just one instance – equivalent to itself. It turns out that stringency is essential in distinguishing explanatory from non-explanatory patterns and is also directly related with unificatory power.

Stringency, however, is not sufficient to understand how the unificationist model provides understanding. Another important notion is that of systematization. Kitcher defines systematization for a belief system K, as a set E(K) which derive some members of K from others. In order for E(K) to count as unificatory it should systematize K by using as few as possible and as stringent as possible argumentative patterns, which should provide as their instantiations all singular statements of K.

Thus explanatory unification has two requirements for explanatory inferences: 1.) They should systematize singular beliefs 2.) in an unificatory manner by using only a few similar and as stringent as possible argumentative patterns.

We can clearly see that explanatory unification, as defined by Kitcher, crucially depends on the stringency condition which determines the scope of application of the argumentative pattern. Stringency itself however depends partially on the structure of the argumentative pattern (as the number of premises it has) and on the way the non-logical vocabulary or concepts are positioned and defined. However focusing exclusively on the scope of explanatory patterns without taking note on the conceptual framework that instantiates them opens up the problem of asymmetry, because as we shall see derivations from cause to effect and from effect to cause in systems with symmetric laws, are made by equally stringent argumentative patterns. This problem could be avoided if we take note that the conceptual framework of theories determines the direction of application and the relational links between the concepts employed in the explanans.

In the following years however efforts in developing the explanatory unification model again focus exclusively on syntactic structure of argumentative patterns. Thus for instance Schurtz (1999) analyzes explanation in four stages:

1.) Explanatory question within the general form of “Why P?”; 2.) The cognitive state of the questioner C. 3.) The answer A. 4.) The expanded cognitive state C+A.

Explanations should satisfy four further constrains: i.) A must satisfy the form “P because of the reasons Prem.” ii.) Prem must be true. iii.) The inference Prem⇒P must be deductive or probabilistic. iv.) State C+A should be more coherent corpus of belief than C. v.) Prem must be less in need of explanation in C+A than the explanandum P in C.

Requirement i. to iii. have to do with soundness, validity and relevance of the explanatory argument and iv. is actually similar to Kitcher’s requirement of systematization of singular beliefs. The last fifth requirement, however remains problematic, since some explanations require mastering concepts by learning new facts
about them, or introduction of completely novel concepts (Hansson, 2007). For instance the introduction of the concept of an “allele” as variation of a gene was introduced in the simple Mendelian argumentative pattern. But the history of genetics itself makes it is fairly obvious that specifying the meaning of “genes” and “alleles” actually opened a giant field for future research.

I will propose that the explanans should not only unify the explanandum within a background theory, by subsuming it under an argumentative pattern. If we have a good explanation the explanatory pattern should potentially permit a explanatory regress or why-regress (Lipton, 2004) extending the premises containing the theoretic concepts. Thus the pattern should be potentially extendable in a non-contradictory and coherent way, by providing further information about the employed concepts and how they capture the salient features of the pheneomena.

Bartelborth (2002) notes en passing that explanation starts at the level of conceptualization but abandons an analysis in that direction in favor of a more traditional, syntactic account of explanatory arguments. He proposes that explanation is related to the general structure of theories. Thus he interprets unificatory explanation as embedding the explanandum in several hierarchically linked models. The models in turn exemplify a net of explanatory patterns with varying degrees of generality. However it again remains unclear how the general argumentative pattern or model should be extended toward generalizations of a limited scope or how new information should be handled by the argumentative pattern.

In the following sections I'll present the main critical claims against the unificationist model. I'll try to address them by defining explanatory unification in terms of stable conceptual links that are potentially coherently extendable to form progressively more informative explanatory arguments.

2. The stringency-systematization problem

The most obvious problem of the unificationist model as it stands is that the notion of systematization and stringency create tension. Broader systematization requires more abstract argumentative patterns, stringency, on the other hand, by adding extra requirements for valid derivations provides more information which is also an essential element for explanatory inferences. The unificationist model is easily vulnerable on account of this tension supplied by the fact that singular explanations often provide extra information which is not directly derivable by the abstract argumentative pattern of the theory.

This criticism has been raised by Margaret Morrison (2000) and Halonen & Hintikka (1999). Morrison raises the point, that unification provides only a general idea about regularities and explanation is more about providing particular casual information. Thus explanation and unification pull in different directions and should be de-coupled and analyzed as separate notions.

Halonen & Hintikka follow similar vein of argument. For them unification is related to theory choice and theory formation and is a part of the background theory operating on the explanandum. Explanation itself is concerned more with supplying particular
information about the explanandum in order to subsume it to the background theory. Reaching a satisfactory explanation is a two step process requiring more than just logical manipulation with an already existing general theory.

The first step is interrogative, related to experiments and observations. It is contextually dependent on the explanandum and as such it supplies information that is not directly deductively derivable from the background theory.

The second step is inferential. It is ordering the background theory and ad explanandum premises in an argumentative structure that will show how the explanandum follows from the theory. As a consequence the explanans involves acquiring new beliefs and it should not be considered as a serial exercise in applying a ready made deductive inference.

I will try to address the information-systematization tension problem of the unificationist model in two steps. The first will show that unification is not irrelevant element of explanations, the second will set a general requirement for the argumentative patterns that will satisfy the tension. In my response I'll build up on the work of Weber & Van Dyck (2002) and Bartelborth (2002).

Weber & Van Dyck, argue that there are particular why-questions whose context requires that the answer should be an unificatory act which subsumes the separate facts under a general theory. If a theory can't address the question, it is precisely due to a lack of unificatory power. For instance a question of the type: “Why do Mary and Peter both have blood type A?” requires subsuming Mary and Peter under the same category, which is part of the conceptual framework of hematology and genetics.

However showing that unification and explanation are not completely separate notions is only a partial solution, we still have to understand how informativeness and systematization could fit together.

We can build up on Weber & van Dyck's contra argument by noting that the element which brings new information in the answer is also the one that does the explanatory unification; the categorization that both Mary and Peter belong to the same group of people who have a particular antigen on the surface of their red blood cells. I believe we can continue to develop the idea that the way the explanadum is conceptualized can also account for the tension between systematization and informativeness.

We can do this by first noting that the informativeness of explanation is usually related to a potential explanatory regress. In the example we can further our inquiry: “Why does Mary and Peter have the similar antigens on the surface of their red blood cells?” by giving more clarifications from genetics and hematology, thus extending our understanding of “blood types”. The conceptualization “blood types” is potentially extendable in supplying a further information about the different categories extending them to more and more contrastive sub-classes. It is now known that there is an extensive heterogeneity underlying the molecular basis of the alleles producing subgroups within ABO (Yip SP, 2002).

Thus the explanans should not only unify the explanandum within a general explanatory pattern. If we have a good explanation the explanatory pattern should potentially permit an explanatory regress in a non-contradictory and coherent way.
Similar suggestion has been made also by Bartelborth (2002). He has proposed that the tension could be resolved by the structure of theories. According to him scientific theories could be represented as a hierarchically ordered tree like structure. In the structure the center is occupied by the general argumentative schema which has the greater unificatory power. At the periphery we have models with greater empirical content and only partial generality. Thus a theory-relevant why-question could be answered within the theory with a varying degree of precision and depth.

However the point H&H make is more elaborate. The idea that explanation and unification should be decoupled is related to the fact that usually theories do not present a complete explanatory store for the phenomena in their scope. Scientific development, related to experimental and observational research, bring new data and new concepts are introduced to the explanatory store of theories.

However the unificationist model originally analyzes explanation essentially as a subsumption of the explanandum under the ready-made argumentative pattern of the theory; it does not include prescriptions how the pattern should be extended to include new information and more detailed contrastive explanations.

My suggestion is that this requirement could be handled by the conceptual framework of the theory. The conceptual framework presupposes stable relations between the central theoretical notions, which are made explicit in the general explanatory patterns. The argumentative schemas themselves should be potentially extendable, by providing extra information on how the general concepts apply to the explanandum.

The additions to the explanatory pattern should be handled by further requirements for stringency. The additions should not be irrelevant and should create a coherent structure evolving from more general to progressively more contrastive explanations. They should also form well defined links between the non-logical elements (or the concepts) included in the schema, by giving further information about the employed concepts and how they capture the salient features of the explanandum.

3. Explanatory asymmetry, unification and causality

The other important objection is that Kitcher's model relies perhaps too much on unification alone. If this is the case, it should provide a more clear definition of what exactly counts as explanatory unification and what is a mere case of ordering and classification (Woodward, 2003). More it seems that unification as it stands does not provide us with a sufficient solution to explanatory asymmetry (Barns, 1992). Both Barns and Woodward objections also aim at showing that causality is an essential element in explanation, that can't be derived from unification alone.

1 By the requirement for potential extension of the explanatory pattern, in terms of providing further information about the central explanatory concepts, the unificationist approach could be made to account for one of the key virtues of the causal mechanistic model. For instance Thagard (2007) has argued that good scientific theories display a constant increase of explanatory coherence. Explanatory coherence itself is defined as: “broadening” the explanatory store of theories by accounting for new phenomena and “deepening” the understanding of the elements of the explanatory mechanisms in terms of more fundamental mechanistic description of their constituent parts.
I'll discuss these objections in two separate paragraphs. First, I'll try to show how asymmetry and causation could be handled by the unificationist model. Then I'll try to define explanatory unification more precisely in order to distinguish it from other non-explanatory forms of scientific unification.

My argument will be based on the idea that considerations about unification and causation are intrinsically interwoven. We evaluate claims about causal links by their adequate application to a broad range of situations or due to their adequacy to unify a large portion of the available evidence. In both cases causal claims depend on conceptualizing the situation within a tangible argumentative structure, that remains stable in face of new information.

I'll defend the view that the key element in a good explanation is bringing the explanandum under such a stable well understood conceptual framework. Thus for a good explanation it is not always necessary to describe a full blown and as intricate as possible causal story, a requirement which a causal purist must adhere.2

This of course has the main virtue of covering the wide array of not strictly causal explanations such as equilibrium explanations (Sober, 1983), distinctively mathematical explanations such as inability to fit a square peg into a round hole (Putnam 1975, 295–297) and functional explanations which employ the effect a system produces to explain its existence. I will not argue that those cases are non-causal or that they can't be reduced in principle to a causal explanations. It is however that they don't require causation to be effective (mathematical explanations) or employ causation in a fuzzy or sketchy way (functional explanations and equilibrium explanations), instead depending on a type of conceptual stability.

The problem of asymmetry appears in cases of systems obeying symmetric laws in which inferences from cause to effect and inference from effect to cause are equally stringent. The only difference between both argumentative patterns lies in the way the non-logical vocabulary is positioned. If we use the classic shadow and pole example as illustration: In one inference we have the length of the pole as a conclusion, in the other the length of the shadow as the conclusion; leaving the rest of the argumentative structure intact. Both derivations are due to valid instantiations of the geometrical optic model. Thus if we stick to the idea that explanatory and non explanatory patterns are discriminated only by unificatory power we have to face the decision whether we should call both derivations explanatory or we should admit that since causation and explanation run in parallel direction, only derivations that correctly describe causal links should count as explanatory.

However, it is immediately visible that the asymmetric discrimination relies on the way the explanandum is conceptualized. If we step back and look closely where exactly the causal element enters into the symmetric case, we'll notice that in the simple optical example the explanation is constituted of two elements: a geometrical model and the conceptual apparatus of optics. The geometrical model as it stands is causally insensitive since it represents the situation within Euclidian geometry and makes derivations about lengths via inferences from the pythagorean theorem. Thus causal considerations are part

---

2 Take for instance the purist position in which all causal influences are relevant to explanation to which (Railton, 1981) and (Salmon, 1984) come close.
of the second element and are related to how concepts like shadow, length and light are defined. Clearly optics is conceptually limited to statements about light propagation and can't generate conclusions about the properties of opaque objects without further clarifications. Causality within optics is handled by the way the concepts are linked together. If we want to follow a further explanatory regress as a result of accepting the shadow-to-pole-pattern, by asking for instance “why do shadows define lengths of objects?”, we’ll reach an explanatory dead end, quite simply because optics has nothing to say about the subject.

The geometrical model alone, however, is not limited so, and can generate empirically correct conclusions with broader scope. Van Fraaseen (1980) has shown that conclusions from shadow to length are not completely unwanted characteristic and could potentially be a part of an acceptable explanation in some situations. For instance an architect can determine the length of a tower and its position because of its disposition to cast a shadow of particular length. In this case it turns out that the shadow-to-pole derivation is essential to the correct explanation for the towers' length. Notice that if we stick to causal purism about the van Franseen case we have to include the mental state of the architect which anticipates the disposition of the tower to cast a particular shadow and this unnecessarily over-encumbers an otherwise more simple and straight-forward explanation.

It is important to note that Van Fraassen's example is not a special case. Typically explanations where the dispositions of a system are employed to explain its existence are classified as functional explanations. Functional explanations form the backbone of biology and perhaps the main reason for their wide application is that analyzing a complex systems such as biological organisms by ascribing functions, reduces and organizes the complexity of the causal mechanisms that we'll strive to include in our explanation. I will return to functional explanations by going over a more detailed example in the last section of the paper, but the important notion is that conceptualization about functions brings the explanandum in a tangible perspective by minimizing the causal data that we have to include in our explanation.

Another reason why bringing the explanandum within stable conceptual boundaries is more essential than simply exhaustively describing its causal structure, are distinctively mathematical explanations. For instance answers of questions such as: “Why can't a man find a path in Königsberg that would cross each bridge only once?”, does not require causes. In that case a more elegant and simple explanation is achieved by subsuming the structure of the situation to a mathematical fact and deriving the conclusion from the features of the mathematical fact. This type of mathematical explanations create a stable explanatory pattern that can be extended to a variety of cases describable under the mathematical concepts, regardless of their causal history. More, Lange (2012) has suggested that this type of explanations are modally stronger than the causal explanations.

However pushing causal considerations to the conceptual background does not solve our initial problem completely. We still would like to know how the causal and the unificationist accounts of explanation could fit together.
Recently Woodward (2003) has defended the position that arguments that cite causes, display invariant relations between dependent variables. We can compare this invariance between the conceptual links within the explanatory argument with the same type of relational stability between the concepts of mathematical explanations. In both cases events are explained because they are brought under an explanatory schema which displays stable relational links between the non-logical elements. Both types of conceptualizations are further extendable to a potential explanatory regress without violating the stability of the conceptual links. Thus it seems that the features of an explanatory argument are its stability between the conceptual links and possible extendability of the concepts as a result of an explanatory regress.

However it is also trivially true that the conceptual apparatus have to be empirically accurate. It should link to particular features of the phenomena and it serves not only to be descriptive but also serves a pragmatic purposes. In that sense, Woodward is correct that conclusions about invariant relation between dependent variables are very often made on the basis of manipulation and the ability for a potential control over the situation.

Finally it seems to me the notion of causation is intrinsically interwoven with that of unification. Even if we define causation in Woodward's terms as invariable relation between dependent variables, we still end up with a potentially inexhaustible amount of relevant causal information that we have to cite if we stick to causal purism. Or we end up with different partial generalizations within a single domain that still have to be reconciliated into a single one.

Instead in explaining phenomena we don't go too further backwards or broaden too much the context, we strive to keep it within certain conceptual boundaries relaying on conceptual links that proved to be stable in covering a broad range of cases, even in the

---

3 Woodward has proposed that causation should be expressed as relational invariance under some interventions or within certain interval for the dependent variables. The scope of invariance could in turn be defined experimentally or counterfactually.

Hintikka (1988, 1999) has shown that this type of relational invariance is usually expressed via mathematical equations that represent the relation between the variables. This equations ideally should represent in an abstract form the results of interrogative process, precisely of the form Woodard suggests; via controlled experiments and observations. This process however results in partial generalizations which confine the variables within a certain intervals of $x_1 < x < x_2$. Other experiments might define different equations for different intervals $x'_1 < x < x'_2$. Here if we stick to Woodward's definition we have obtained different causal relations which naturally are represented by different equations, since causality is defined as “invariance under some interventions”. This however is in par with actual scientific practice which strives to unify partial generalizations under more general ones.

A way around could be to define the scope of invariance, not by actual experiments which are limited in range, but by some form of counterfactual. This however is not so straightforward. To measure the scope of invariance in this way one again requires some form of quantifying or counting the interventions. Here however we're interested in conceptually possible interventions not on actual experimental ones. But this leads to infinity of similar interventions because possible worlds in which the interventions are carried out are defined within a logical space which is characterized by “an abundance” of possible worlds. To quote Lewis (1986, 86):

“1.) absolutely every way that a world could possibly is a way that some world is and
2.) absolutely every way that a part of a world could possibly be is a way that some part of some world is.”
cost of initially having a somewhat distorted or simplified account by omission of some causal data.

4. Explanatory and non explanatory unification. Classification schemas

The third important objection against the unificationist approach is due to Woodward (2003). He raises the issue that explanatory unification isn't defined precisely. The notion itself is quite heterogeneous and covers a broad range of scientific activity like classification, using a common descriptive vocabulary or similar mathematical framework. Of all this cases Woodward claims, only genuine physical unification, where heterogeneous phenomena are unified under a common set of mechanisms or causal relationships, could be considered related to explanation. The unificationsit model however can't distinguish between those cases because it rates explanations only in terms of descriptive economy, information compression or deriving as much as possible from as less, as possible inferential patterns.

More Woodward claims that the criterion for evaluating explanations is arbitrary, because it remains unclear how to decide between a more stringent pattern and a pattern with wider scope of application. For instance, in the case of a common set of mathematical equations describing the behavior of different systems, we have a common argumentative schema with a wide scope. The scope is made more stringent only by the non-logical elements that determine how the variables are defined within the different domains of application of the equations. It is however exactly because of the extra concepts, which define the scope of the mathematical equations within the theory, that the equations become meaningful and explanatory. But the unificationist model as it stands, does not provide us with a reason why we should consider explanatory the more stringent explanatory schema over the one which covers all range of application of the mathematical equation.

A natural response to this problem is to look for the information-systematization requirement.

It is clear that the extension of the mathematical pattern to several particular domains of application involves adding unrelated concepts to the pattern. Thouse “in theoretic” concepts define the shared variables and the exact meaning of the equations within the theory. Thus the extension of the mathematical pattern within a particular theory introduces extra concepts in a well defined manner creating a coherent explanatory structure. On the contrary taking the whole scope of different applications of the equations leads to a non-coherent extension, because the added different theoretic concepts remain unrelated.

The case is similar concerning classificatory schemas. First it is worthy to note that classification is not always a case of non-explanatory unification. Two examples readily come to mind: Evolutionary Threes and the Periodic Table.

Phylogenetic threes classify biologic diversity in terms of inferred evolutionary relationships or common ancestry. Thus at least some of the traits of the species could be derived by the position of the species within the three. More the construction of this classificatory schema relies on the general explanatory pattern of evolutionary biology.
This point is more obvious in the case of the Periodic Table. Here the basis for the classification is the atomic number and electron configuration, this in turn leads to periods of elements having similar properties. As a result one can derive the properties of an element from its position on the table. Thus some forms of classification should be distinguished from cases of ordering, where the ordering procedure is arbitrary, serves perhaps a purely pragmatic purpose and does not link the ordering procedure to the theoretic background in a meaningful way.

Explanatory classification schemas on the other hand, don't just order the data, they link the systematization procedure to the general argumentative pattern of the theory. As a result explanatory information can be derived directly by the position of the data quantors in the classification. Notice that in the case of the periodic table the derived information is not circular, the properties of the elements are not explained directly by their position, but the position of the element leads to a conclusion about their atomic structure which in turn serves to explain the properties. More the argumentative schema displays both stability of the conceptual links and is potentially extendable since the Periodic Table kept the integrity of its structure in face of new information from particle physics. Thus I believe some forms of classification could actually be exemplified as a supportive case for explanatory unification.

To recapitulate, Woodwards argument against unificationist model depends on the fact that explanatory unification is defined only in terms of the scope of application of an argumentative pattern. I however defined explanatory unification in terms of stable conceptual links that are potentially coherently extendable to form progressively more informative explanatory arguments. In that sense, forms of classification which link their systematization procedure to the general conceptual framework of the theory are actually explanatory. More the stability of a classification schema in face of new information is an actual evidence in favor of explanatory unification.

In the next section I'll go into some more detail on what kind of conceptualization is suitable for constructing good explanatory arguments.

5. Requirements for the conceptual framework.

From what we have said so far one can conclude that a lot of complications could be avoided if we just focus on conceptualization as the main ingredient of explanation. Such claim has been defended by Bengt Hansson (2007), according to whom explanation is not a logical, but an epistemic structure which primary function is to organize conceptual content. I partially agree with this view but strictly speaking, conceptualization alone is not sufficient for good explanations and there are several problems with Hansson's approach. For one he's well aware that: “...placing explanations in the realm of concepts means that one is exposed to the danger of individual or collective ad hoc constructions.”(Hansson, 2007, 10). More the conceptual links that he relies to bring understanding in an explanation could create a circular argument which as a whole could not be considered explanatory. Finally, in some cases, even if a conceptualization is of sufficient level of generality and is part of an intricate belief system, it still can't be considered a basis for constructing a good explanatory arguments.
In what follows I'll give some examples of these difficulties and I'll propose that they can be avoided if some further requirements are added for the explanatory conceptual structure. The extra requirements will serve to limit the types of conceptual frameworks suitable for constructing good explanatory arguments.

The first problem as Hansson himself notes are the *ad hoc* constructions. *Ad hoc* explanations could arise in 2 guises. One could be a list forming a sufficiently large set of concepts and definitions which will link them to phenomena forming very detailed descriptions of particular events. We, however, will not consider such *ad hoc* narratives explanatory quite simply because they don't present any stable conceptual links or present us with any coherent view of phenomena.

The other form of *ad hoc* constructions is the classical case where we have a sufficiently general conceptualization which is non-discriminative and will link to the explananadum by adding premises containing arbitrary concepts. The old joke about psychoanalysis is a good example: “A man is diagnosed with Oedipus complex. To which he protests: “But I love my father.”; “Oh well, that's because you're overcompensating.” replies the doctor.” The extra concept of overcompensation is part of the structure of psychoanalytic theory however its addition does not lead to a progressively more contrastive explanations. The addition actually blurs further distinctions. As a result the requirement for the construction of the explanatory argument should be that the addition of concepts and extra premises should define progressively more specific classes within the conceptual boundaries of the theory.

Circularity could be avoided by the requirement that the conceptual framework should be potentially extendable in a stable way, as a result of an explanatory regress. Think of example from genetics. The concept of gene although initially very broadly pointing to a yet unspecified mechanism of heritability, proved to be further specifiable into more detailed concepts as the DNA was discovered. In turn DNA was further specified by detailed description of the structure of the molecule and the process of replication. This traced progressively more contrastive descriptions of heritability. As a result the general Mendelian conceptual framework was a stable basis for progressively more extendable and detailed explanatory patterns with the addition of new information about particular genetic mechanisms.

The final problem is when the explanatory conceptualization is an element of well connected global belief system. The practice of bloodletting in pre-modern medicine, presents an intriguing example. In that case the explanation why pre-modern doctors thought the practice was a legitimate way of healing certain diseases was due to a completely unified view about nature. It presented the human organism in terms of four humors – black bile, yellow bile, phlegm and blood. They in turn were a basis for classifying diseases as caused by a dis-balance between the four humors (Porter, 1999). Finally the four humors were linked to the four classical elements which were a part of a metaphysical system which could represent all phenomena as dynamic mixtures between them. The problem in this type of explanatory conceptual structure is not at the level of systematization nor at the amount of particular information that is subsumed under the general conceptual framework. The flaw is that the conceptual framework as a whole
does not permit construction of refutable arguments that could potentially limit the scope of key explanatory concepts.

From what we have said so far we can draw a set of requirements for a conceptual framework that could serve as a basis for developing good explanatory arguments:

1. It should permit a construction of refutable, non-circular argumentative patterns.

2. The argumentative patterns should exemplify stable conceptual links.

3. The argumentative patterns should move from more general concepts to concepts with progressively limited scope. The extension of the argumentative pattern or the addition of concepts and extra premises should be handled in well defined and coherent manner creating progressively less overlapping classes within the conceptual boundaries of the theory.

6. Explanatory unification: A case from Paleontology

So far the analysis of explanatory unification has been focused mostly on general concerns. In the last section of this paper I would like to give the discussion more substance by going over two examples of explanatory unification within evolutionary biology. Both cases will be from paleontology, they will cover two episodes from the dispute between orthogenesis and Darwinian evolutionary theory. The first case will focus more on the general conceptual framework of orthogenesis and on the question why the theory was unsuccessful in spite the fact that it provided a relatively unified view of evolution. To anticipate I will argue, that the conceptual framework it presented was not a good foundation for argumentative patterns which could assimilate new data, because the explanatory schema was circular and failed to satisfy explanatory regress.

The second case will be a particular example of Darwinian explanation – the evolution of the Irish elk (Megaloceros giganteus). Here Gould’s argument presented a better unified explanation which systematized a large amount of data, better than the rival orthogenetic explanations (Gould, 1977). Again I shall argue that the success was due to extendable argumentative pattern which remained conceptually stable as new evidential data has been presented.

Orthogenesis or progressive evolution was one of the primary rivals of Darwinian evolutionary theory before the advent of the evolutionary synthesis. Very briefly orthogenetic theories hold that evolution is unilinear, due to some internal driving force. This concept was in brisk contradiction with natural selection. The orthogenetic argument went that, if evolution is a result of a fixed pattern, natural selection could do little to change the direction of evolutionary change.

Orthogenesis primary support came from paleontology. Paleontologists who supported the theory, tried to systematize the fossil record using a classification schema attuned to the orthogenetic conceptual framework. A particularly nice example is the work of Alpheus Haytt on the evolution of planorbid snails in Steinheim meteorite crater lake in Germany. In 1880 Haytt published a monograph on the snails in which he structured their evolutionary history differently than the existing Darwinian phylogenetic three of Hilgendorf (Gould, 2002; Rasser, 2013).
Fig.1. Theoretical considerations about the concepts applied to the Miocene Steinheim snails. Below: Field observation reveals the distribution of six hypothetic species in three consecutive, hypothetic sedimentary beds. Upper left: Hilgendorf used the vertical distribution combined with morphological similarities and transitions in order to construct a genealogic tree. Upper right: Hyatt refused the branching of one species into two new species and ignored the stratigraphic position of morphs. Re-drawn from (Rasser, 2013).

Haytt ordered the fossil record in four separate lineages evolving in parallel (see fig.1). This systematization was a direct consequence of his orthogenetic theory according to which evolution follows a pattern similar to the stages of ontogeny (Gould, 2002).

Hyatt ordered the four lineages on the basis of supposedly progressive and retrogressive characters. Progressive characters included increase in size, shell thickness and strength of ornamentation. On the contrary regressive characters included decrease in size, loss of ornamentation, thinning of the shell and tendency for irregular growth by uncoiling. This choice of regressive and progressive characters was based on presupposed literal similarity between the developmental stages of individuals and the general orthogenetic pattern.

What's striking about his classification is that its vertical gradient is not temporal. Snail shells on the same level were supposedly on a common stage of the orthogenetic pattern but did not necessarily live at the same time period.

Notice that even thou Haytt systematization of the fossils is unifying and is explanatory for the different characteristics of the shells it does not create a good explanatory argument. There are several flaws which are easily visible if we apply the
criteria we've set for suitable conceptual framework on which argumentative patterns could be based.

First, Haytts' explanatory schema is obviously circular. The ordering of the shells is based on his conceptual framework which classify the fossils in terms of orthogenetic patterns displaying phyletic youth, maturity and senility. Thus a particular shell form is explained because of if it's position on the orthogenetic pattern and the orthogenetic pattern is explained by the ordering itself. Second, Hayyts' explanatory schema is not stably extendable as a result of an explanatory regress, because it fails to present any mechanism which would account for the goal oriented evolutionary process. Finally the argumentative schema failed to assimilate new data found by paleontologist, which displayed clearly non-linear branching.

On the contrary, the phylogenetic systematization of Hilgendorf (perhaps one of the first attempts at a phylogenetic three) was based on the Darwinian conceptual framework. The Darwinian conceptual framework, proved to be a stable systematization for paleontology; the branching representing specification could potentially be successfully linked to unified historical explanations of particular species formation.

To show how the Darwinian conceptual framework was applied successfully to systematize and explain particular data from the fossil record, I'll do a short survey on the famous case of the Irish Elk (Gould, 1977).

The problem how to explain the animal's large antlers, was a puzzle addressed by both Darwinian evolutionary theory and orthogenesis. Darwinian inclined paleontologists conceptualized the problem in terms of “What kind of adaptive function did the antlers had?”, on the contrary orthogenesis proponents conceptualized it as a case of unwinding orthogenetic pattern that developed regardless of selective pressures.

Gould presented what I believe is the best unifying explanation. He measured a disproportion between the antlers and body size. Clearly a possible Darwinian explanation will be that the disproportion was due to a selective pressure acting on the antlers size. Thus Gould systematized the fossil measurements in terms of a functional analysis. His solution to the problem was that the antlers played a sexual selection function. The specific structure of the trait was in accord to this function and known behavioral patterns of non-extinct relative species supported the explanation.

It is clear that the explanation Gould provided, largely relies on unification. As a historical explanation, its success depends on organizing a large body of data under the conceptualization about the function of the antlers for sexual selection. However the available information is not just systematized under the function. The proposed function itself is subsumed under the general conceptual framework of evolution via the concept of natural selection. As a result, the successful explanatory argument depends on the extension of the Darwinian explanatory pattern to Megaloceros giganteus and the stability of the pattern which permits systematization of large body of actual observational data.

Explaining functions, is a standard affair for evolutionary biology and as far as we steer clear from the problem of holism V.S. reductionism, explaining functions of traits does not violate the temporal structure of causality, because effects that serve an adaptive
purpose form a feedback loop which sustains the persistence and reproduction of the biological system.

The final problem I would like to draw attention to, is that complex systems such as organisms can in principle produce an inexhaustible amount of effects, but we never strive to construct a complete causal account of the mechanisms that brought them about. It is clear that causality alone can't be used to construct biological explanations. The more so, because in evolutionary explanations we're bound to link observational and experimental results with historical inferences, that can't in principle be tested directly.

Conceptualizing about functions on the other hand minimizes the amount of observations that we would like to potentially explain. Linking functional ascription to the evolutionary background by the concept of optimization and adaptation still narrows down the data.

Returning to our example about the length of the antlers the evolutionary explanation before Gould was that the antlers were not a decorative element but served as weapons for actual battle. Gould measurements proved that this presumption can't organize the observational data in a meaningful way in spite of the consistence of the hypothesis with the Darwinian background. Instead his own functional analysis proved to draw a more unified picture. Thus we can note that conceptualizing about biological functions strives to achieve unification on two sides:

1. It should link the explanation to the Darwinian background and should be maximally coherent with the theoretic background, extending the explanatory store of Evolutionary Biology.

2. The functional ascription itself should agree with as large as possible body of evidence and contradict as little as possible with known facts.

It seems to me that considerations about unification bring the main reason for the acceptance of this type of evolutionary explanations.

Conclusion

I've tried to solve some of the difficulties of the unificationist model, by developing the view that explanatory unification is a multi-leveled process which goes beyond the syntactic structure of argumentative patterns. I've shown that the explanatory power of argumentative patterns crucially depends on the stringency condition. The stringency itself relies on the application of the non-logical elements of the argument or the employed concepts. As a result I've suggested that the conceptual framework of theories plays a crucial role in unificatory explanations. It presents stable conceptual links that should be made explicit in the argumentative patterns. The general argumentative patterns themselves should be potentially extendable to cover particular explanatory cases by adding extra premises and clarifying concepts. The addition itself should strive to form a coherent structure, as a result of adding extra information, on how the employed general concepts apply to the salient features of the explanandum.

REFERENCES


