Chapter 8
Unification through Confirmation

Ioannis Votsis

Abstract The connection between unification and confirmation has been underappreciated. Although seminal works in the literature allude to this connection, they typically fail to provide critical details. Moreover, in the same works the burden of analysing the concept of unification falls on the concepts of understanding and explanation. I argue that the prospects of this approach appear bleak as the latter concepts, at least as they are traditionally construed, are opaque and not readily amenable to an objective treatment. As an alternative, I shift the entire burden of the analysis to confirmational concepts, offering not just a novel account of unification but, more importantly, something that has been virtually missing from the literature, namely a quantitative measure.

Keywords Unification measure • Confirmation • Relevant deduction • Scientific methodology • Informational relevance measure

8.1 Introduction

The connection between unification and confirmation has been underappreciated. Although seminal works in the literature allude to this connection, they typically fail to provide critical details. Moreover, in the same works the burden of analysing the concept of unification falls on the concepts of understanding and explanation. I argue that the prospects of this approach appear bleak as the latter concepts, at least as they are traditionally construed, are opaque and not readily amenable to an objective treatment. As an alternative, I shift the entire burden of the analysis to confirmational concepts, offering not just a novel account of unification but, more importantly, something that has been virtually missing from the literature, namely a quantitative measure.

I. Votsis (✉)
Philosophy Faculty, New College of the Humanities, 19 Bedford Square, London, WC1B 3HH, UK

Department of Philosophy, Logic & Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK
e-mail: ioannis.votsis@nchlondon.ac.uk; i.votsis@lse.ac.uk

M. Massimi et al. (eds.), *EPSA15 Selected Papers*, European Studies in Philosophy of Science 5, DOI 10.1007/978-3-319-53730-6_8

8.2 Why Unify?

Unification is a guiding force in theory choice, that is, in the construction, modification, acceptance and refutation of theories. Some of the most famous examples of theories in the history of the natural sciences form part of a series that exhibits an unmistakeable unificational trend. Galileo's physics unifies the motions of free falling, rolling and projectile bodies. Newton's physics merges Galileo's terrestrial and Kepler's celestial laws of motion and much else besides. Maxwell's theory of electromagnetism integrates electrical, magnetic and optical phenomena into a unified whole. Einstein's special theory of relativity makes possible the marriage of Maxwell's electromagnetism with the principle of light speed invariance and Galileo's principle of relativity. Finally, the Glashow-Salam-Weinberg theory unites electromagnetism and the weak nuclear force.

Scientists, especially those in the natural sciences, are quick to sing the praises of simpler and more unified theories.

> Nature is after all simple, and is normally self-consistent throughout an immense variety of effects, by maintaining the same mode of operation (Newton 1959, p. 418).

> Our job in physics is to see things simply, to understand a great many complicated phenomena in a unified way, in terms of a few simple principles (Weinberg 1979, p. 543).

> Biology is a science of three dimensions. The first is the study of each species across all levels of biological organization, molecule to cell to organism to population to ecosystem. The second dimension is the diversity of all species in the biosphere. The third dimension is the history of each species in turn, comprising both its genetic evolution and the environmental change that drove the evolution. Biology, by growing in all three dimensions, is progressing toward unification and will continue to do so (E. O. Wilson 2005, p. 1).

The first and second quotes are informative in their suggestion of how simplicity and unification link up. If to unify is, roughly speaking, to demonstrate how two or more seemingly distinct domains of phenomena can be accounted for together, not by the same number of principles as they once were but by fewer, then to unify is also to simplify. A similar story can be run backwards, i.e. from simplicity to unity. If to simplify is, roughly speaking, to show how the same phenomena can be accounted for by fewer principles and by doing so show that those principles are enough to treat phenomena that were once treated separately, then to simplify is also to unify.

Why are scientists, and why should they be, on the lookout for unified theories? There are two main answers to such questions. Some, e.g. Myrvold (2003), suggest that unification leads to the truth. Others, e.g. van Fraassen (1980), suggest that unification is merely a pragmatic or, worse, aesthetic marker. This dispute cannot be resolved unless we first have a good grasp of the notion of unification. And, if a theory can be more or less unified, a good grasp of that notion would have to go hand-in-hand with a quantitative measure. We thus seem to have hit upon three crucial questions in our quest to fathom unification.

1. What does it mean for a theory to unify phenomena?
2. How, if at all possible, can we measure the degree of a theory's unifying power?
3. Is unification truth-conducive?

This essay is primarily concerned with the first two questions. As we will see in the next section, existing answers to the first question have pretty much stalled in their attempt to make considerable progress. More damningly, the second question has been largely ignored.

8.3 Unification through Explanation and Understanding

By far the most influential accounts of unification are those of Friedman (1974) and Kitcher (1989). According to the first, an explanation is more unified when there is a reduction in the number of independently acceptable law-like assumptions that are employed in the derivation of the explanandum. According to the second, our explanations unify by, on the one hand, increasing the number of conclusions derivable from argument schemas, while, on the other, decreasing the number of such schemas. These two accounts have a number of virtues, not least of which is the use of deductive logic as a formal backbone to their conceptions. They also have well-documented failings – see, for example, Schurz and Lambert (1994). As space is at a premium, I will not rehearse these failings here. Instead, I will voice a general grievance I have with these and other related approaches.

A good analysis, or explication (in the Carnapian sense), endeavours to articulate the concept under study with a handful of other concepts that are presumably better understood. In the case at hand, that honour often falls on the concepts of explanation and understanding. The trouble with this approach is that it seeks to articulate one hard-to-pin-down concept with two other concepts that are, arguably, even harder-to-pin-down. Part of the problem is that the proposed analyses attempt to do justice to an unwieldy mix of intuitions about scientific and everyday uses of the concepts of explanation and understanding. The expectation that a single analysis can accommodate both types of uses goes back to the early critiques of the deductive-nomological model of explanation. Scriven (1962), one of the model's main critics, reasoned that we must square our view of explanation not only to scientific, but also to everyday, cases. Indeed, many of the counterexamples to the deductive-nomological model – think of the inkwell counterexample – were inspired by intuitions elicited in everyday cases. Since then, the view has become highly influential. As Woodward notes, many treatments of scientific explanation focus "on 'ordinary life' singular causal explanations . . . the tacit assumption being that conclusions about the structure of such explanations have fairly direct implications for understanding explanation in science" (2014).

Using everyday cases to inform the analysis of a concept is not in-and-of-itself a bad thing. Having said this, those who do so run the risk of ignoring the possibility that the intuitions underlying such ‘ordinary life’ cases do not form a coherent whole, especially when taken together with intuitions underlying ‘scientific life’ cases. Indeed, more than being a mere logical possibility, it is quite obvious that some intuitions concerning the goodness of explanations and the genuineness of understanding are downright contradictory. For example, the intuition that listing causally relevant factors is sufficient for an explanation contradicts the intuition that sufficiency can be secured by making the *explanandum* highly expectable. After all, and as the well-known paresis case illustrates, factors can be deemed causally relevant even when such expectability is low.

8.4 Unification through Confirmation

That light can be shone on unification through considerations arising in the study of confirmation is not a novel idea. In fact, this idea is present in all of the abovementioned works. Even so, it is not properly appreciated and gets lost in a sea of other issues, which as we have already seen include what is a good explanation, what provides understanding and how the two relate. In this section, I offer a conception of unification and an associated quantitative measure, both of whose only recourse is to confirmational concepts.

The account and measure to be outlined below conceives of unification in terms of the notion of confirmational connectedness. Put crudely, confirmational connectedness is a relation that may or may not hold between two propositions. If the said propositions reflect aspects of nature that are connected, then, the argument goes, support for the one proposition should spread to the other and vice-versa. This spread amounts to an increase in the overall confirmation of the two propositions. As such, the ability of a hypothesis to reflect unities in nature turns out to be an empirical virtue. An earlier version of this account and measure can be found in Votsis (2015). Given limitations of space, not every detail merits repeating here but I will do my best to communicate the central ideas.

For illustration purposes, consider the following conjunctive hypothesis: ‘JFK was shot dead on November 22 1963 *and* the W boson has a mass of approximately 80.385 GeV/c^2’. Arguably, this hypothesis is highly disunified because its conjuncts are unrelated. One way to demonstrate their unrelatedness is through confirmational considerations. The evidence required to support the one conjunct is clearly unrelated to the evidence required to support the other. In the language of my account, the two conjuncts are *confirmationally disconnected*. Compare this to the arguably highly unified conjunctive hypothesis that ‘JFK was shot dead on November 22 1963 *and* Lee Harvey Oswald was the only person to shoot at JFK on that date’. The evidence required to support the first conjunct is clearly related to the evidence required to support the second. The two are thus, again in the language of my account, *confirmationally connected*.

Having laid out a rough understanding of what the notion of confirmational connectedness means, it is now time for a more rigorous look.

Confirmational Connectedness Any two content parts of a non-self-contradictory proposition Γ expressed as propositions A, B are *confirmationally connected* if and only if for some pair of internally non-superfluous and non-sub-atomic propositions α, β where α is a relevant deductive consequence of A and β is a relevant deductive consequence of B: either (1) $P(\alpha|\beta) \neq P(\alpha)$ where $0 < P(\alpha), P(\beta) < 1$ or (2) there is at least one atomic proposition δ that is a relevant deductive consequence of $\alpha \wedge \beta$ and is not a relevant deductive consequence of either α or β on their own.

Confirmational Disconnectedness Any two content parts of a non-self-contradictory proposition Γ expressed as propositions A, B are *confirmationally disconnected* if and only if for all pairs of internally non-superfluous and non-sub-atomic propositions α, β where α is a relevant deductive consequence of A and β is a relevant deductive consequence of B: (i) $P(\alpha|\beta) = P(\alpha)$ where $0 < P(\alpha), P(\beta) < 1$ and (ii) there is no atomic proposition δ that is a relevant deductive consequence of $\alpha \wedge \beta$ and is not a relevant deductive consequence of either α or β on their own.

We need only take a closer look at one of these two notions to understand the other. Take disconnectedness. The first question that pops to mind is what does it mean to be a content part? A content part c of a proposition Γ is a non-trivial consequence of Γ from which Γ cannot be derived. That is, the content of c is strictly smaller than the content of Γ. The non-superfluousness requirement is there to, among other things, reduce the evaluation's complexity.

Now consider clause (i). Note, first, that the probabilities are meant to be objective. That is, they are meant to indicate true propensities and/or true relative frequencies of events, states-of-affairs, etc., expressed by propositions. This is an important point, as the aim is to de-subjectivise the notion of disconnectedness (and hence connectedness) by requiring that probabilities are fixed by worldly facts. Note, moreover, that the notion of probabilistic independence lets us express one important way through which two content parts are confirmationally disconnected. For if α, β are probabilistically independent the probability of the one proposition is not changed if we assume something about the truth-value of the other. This is in agreement with the Bayesian relevance criterion of confirmation according to which a piece of evidence e confirms or disconfirms a hypothesis h if and only if the two are probabilistically dependent.

More needs to be said to account for the confirmational disconnectedness of two propositions. We must inspect not just their content in total but also the content of their parts. After all, two propositions may be probabilistically independent although certain of their parts are not. That's why we need the notion of *deductive consequence*. By checking whether each and every deductive consequence of the one proposition is probabilistically independent from (or dependent on) each and every deductive consequence of the other, we seek to avoid missing out on any further confirmation relations between the two propositions.

A further qualification is in order. We do not want to consider *all* deductive consequences. After all, some such consequences are trivial. In fact, were we to consider these we would make the concept of disconnectedness unsatisfiable. We can demonstrate this with a simple example. Regardless of their content, A and B each have validly derivable but trivial consequences that they share. One such trivial common consequence is $A \vee B$. Thus, if we take both α_i and β_i to be $A \vee B$, $P(\alpha_i | \beta_i) \neq P(\alpha_i)$ provided $0 < P(\alpha_i) < 1$. Otherwise put, there is a guarantee that A, B are connected. To rule out cases of this sort we impose the restriction of *relevant* deductive consequences. This notion of relevance originates in Schurz (1991) where it is explained as follows: "the conclusion of a given deduction is *irrelevant* iff the conclusion contains a component [i.e. a formula] which may be replaced by any other formula, salva validitate of the deduction" (pp. 400–401). In the case at hand, $A \vee B$ counts as an irrelevant consequence of A precisely because if we replace B with any other proposition, including $\neg A$, the validity of the conclusion is assured.

Although two propositions A, B may be probabilistically independent through and through, it may still be the case that they are confirmationally related by jointly having a relevant deductive consequence δ that neither relevantly entails on its own and whose truth would provide support to both. Clause (ii) is there to guarantee that there is no 'indirect' relation of support between A and B via δ. The presence of such a consequence would mean that A, B are not disconnected. More precisely, unless we place additional restrictions on δ, the concept of disconnectedness will once again become unsatisfiable. If the restriction was merely that there is no δ that any α, β possess as a joint (but not separate) relevant deductive consequence, then there would always be such a δ. For example, δ could be $\alpha \wedge \beta$ where α is logically inequivalent to β. More generally, a proposition that is jointly (but not separately) entailed by two propositions α, β cannot function as a δ if it is logically equivalent to $\varepsilon \wedge \zeta$ where ε is a relevant consequence of α and ζ is a relevant consequence of β. Such propositions are trivial consequences for our purposes and in that respect incapable of assisting our quest to find confirmational connections between A, B.

To solve this problem we need the notion of an atomic proposition. In this context, this notion does not mean the same as the familiar one from logic. Roughly speaking, an atomic proposition in the new sense involves containing the least amount of content that can be used for the purposes of confirming a hypothesis. To be more exact, a proposition φ is atomic if, and only if, φ is nonsuperfluous and truthfully represents all and only the content of an atomic state of the world. The important point here is that the content of an atomic proposition cannot be decomposed into distinct atomic (or indeed molecular) content parts. How does this restriction rule out guaranteed joint relevant consequences like $\alpha \wedge \beta$? If δ satisfies those restrictions, then it cannot be decomposed into a logically equivalent conjunction $\alpha \wedge \beta$. For suppose δ is indeed logically equivalent to $\alpha \wedge \beta$. Recall that α, β are definitionally required to be not sub-atomic. So either one or both are atomic or molecular and hence δ is not atomic. Contradiction! Requiring atomicity disallows such decompositions. Thus, consequences like $\alpha \wedge \beta$ are banned.

ioannis.votsis@nchlondon.ac.uk

Now that we have an exposition of the notions of connectedness and disconnectedness we can define a measure of the level of a hypothesis' connectedness and hence of its unification. The degree of unification u of a proposition Δ is given by the following function:

$$u(\Delta) = 1 - \sum_{i=1}^{n} d_i^{\alpha,\beta} / \sum_{i=1}^{n} t_i^{\alpha,\beta}$$

The term $d_i^{\alpha,\beta}$ denotes the number of disconnected pairs α, β in a given content distribution i, the term $t_i^{\alpha,\beta}$ denotes the total number of connected plus disconnected pairs α, β in a given distribution i, and the term n denotes the total number of content distributions. The number of disconnected pairs in a given content distribution is fixed by computing the number of times a different pair of relevant deductive consequences α, β satisfies clauses (i) and (ii) concurrently. Any pair that is not disconnected counts as connected. The higher (/lower) the value of $u(\Delta)$, the more (/less) unified the content expressed by Δ.

Consider a toy case first. A hypothesis that conjoins propositions that express unrelated facts, e.g. f_1 is a white swan $\wedge f_2$ is a red dwarf, gets a low unification score because the two corresponding conjuncts are neither probabilistically related nor do they jointly and relevantly entail an atomic consequence. By contrast, a hypothesis that conjoins propositions that express related facts, e.g. g_1 is a white swan $\wedge$ g_2 is a white swan $\wedge$ ... $\wedge$ g_n is a white swan, gets a high unification score. That's because the stated facts are systematically related both via probabilistic dependence, e.g. the offspring of white swans are likely to be white, and via jointly and relevantly entailing atomic consequences, e.g. the proposition (or, if it's not itself atomic, a content part thereof) that there are at least two white swans.

Now consider a real case. One of Newton's main accomplishments was to show that the motion of planets and the motion of free falling objects on Earth are instances of the same force, i.e. gravity. To each instance of this regularity there corresponds a unique proposition, e.g. p_1 : '$\{b_1, b_2\}$ is a pair of objects with mass satisfying Newton's inverse square law', p_2 : '$\{b_3, b_4\}$ is a pair of objects with mass satisfying Newton's inverse square law', etc. When we take the conjunction of any two such propositions they relevantly entail a third, viz. q: 'There are at least two pairs of objects with mass satisfying Newton's inverse square law'. Proposition q cannot be relevantly entailed by either of the original propositions on its own. Even if we assume that q is not itself atomic but molecular, at least one of its content parts, call it q_a, is atomic. Thus, the second condition of our definition of connectedness is satisfied. Moreover, since *all* conjunctions made up of p-like distinct propositions, e.g. $p_3 \wedge p_4$, relevantly entail q and hence q_a, there is a systematic connection between the content parts of Newton's inverse square law. But that just means that the law qualifies as a highly unified hypothesis.

The proposed measure is not without its faults but it has at least two merits in addition to its objective character. The first concerns the requirement that all content distributions be taken into account. Take two propositions A_1, B_1 where $A_1 : c_1 \wedge c_2 \wedge c_3$ and $B_1 : c_4$. Take also two propositions $A_1{}'$, $B_1{}'$ where $A_1{}' : c_1 \wedge c_2$ and $B_1{}' : c_3 \wedge c_4$. Suppose we use some measure on the two pairs that outputs score s_1 for A_1, B_1 and score s_2 for pair $A_1{}'$, $B_1{}'$. Since $A_1 \wedge B_1$ has the same content as $A_1{}' \wedge B_1{}'$ any proposed measure should ensure that $s_1 = s_2$. What can we do to guarantee this? The above solution calculates scores by considering *all* distinct ways of distributing the same content over two different propositions. Thus, $u(\Delta)$' s value is the same no matter how we cut Δ, e.g. into A_1 vs. B_1 or into $A_1{}'$ vs. $B_1{}'$, since we consider *all* other ways the same content, i.e. $A_1 \wedge B_1$ or equivalently $A_1{}' \wedge B_1{}'$, can be distributed over two propositions. Such cuts have *all and only* the same content distributions.

The second advantage is that the measure is quite flexible in its applicability. This is because the notions of connectedness and disconnectedness don't place any restrictions on the compared propositions other than consistency. The propositions can thus be chosen quite liberally and include central and peripheral hypotheses, laws, accidental generalisations and assertions of background conditions. This not only allows us to gauge the degree of unification between a central hypothesis and a peripheral hypothesis, but also between any other consistent pair of propositions, e.g. two peripheral hypotheses.

8.5 Myrvold's Measure: A Brief Comparison

In this last section, I would like to consider one of a handful of other unification measures to have been proposed and the only other, besides the current one, where the relation between confirmation and unification takes centre stage.[1] Myrvold (2003) analyses unification as the ability of a hypothesis to increase the degree of informational relevance of a body of phenomena. In other words, a hypothesis that unifies makes prima facie disparate phenomena yield information about each other. Crucially, such information provides an additional confirmational boost to the hypothesis in question. In this short section, I'd like to briefly contrast Myrvold's account with my own.

The first thing to note is that the notion of informational relevance is similar to the notion of confirmational connectedness. Both underwrite the idea that unity

[1] As one referee rightly indicated, there is an abundance of proposed measures of coherence that may be profitable in this context. For a cursory discussion of the similarities and differences between coherence and unification the reader may consult Votsis (2015, §8).

can be found where phenomena are probabilistically linked. To be precise, Myrvold defines his measure of informational relevance $I(\cdot, \cdot|\cdot)$ as follows:

$$I(p_1, p_2 | b) = \log_2(P(p_1 | p_2 \wedge b)/P(p_1 | b))$$

Here p_1 and p_2 are propositions expressing phenomena and b is background knowledge. The informational relevance of two propositions is determined by whether or not, and, if so, to what extent, they are probabilistically dependent. Clearly, if p_1 and p_2 are probabilistically independent then $I(p_1, p_2 | b) = 0$ but otherwise $I(p_1, p_2 | b) > 0$. Myrvold then defines his unification measure $U(\cdot, \cdot; \cdot | \cdot)$ in terms of the relative increase in the informational relevance between two propositions that a hypothesis h may provide.

$$U(p_1, p_2; h|b) = I(p_1, p_2|h \wedge b) - I(p_1, p_2|b)$$

Thus, by comparing the informational relevance of two propositions both in the presence and in the absence of a hypothesis we gauge how much informational unity is accrued in light of h.

One advantage Myrvold's account has over my own is that it checks not merely for probabilistic dependence but also for the strength of that dependence. This is something that I plan to incorporate in future versions of my account but that, as of yet, is not fully worked out.[2] To give the reader a little foretaste, instead of using $u(\Delta)$ to simply count the number of connected pairs of relevant deductive consequences α, β in a given content distribution (and ultimately the sum when all distributions are considered), we can use it to count both the number and the strength of those connections. For example, each pair connected through the satisfaction of condition (1), i.e. because of a probabilistic dependence between α, β, may be assigned a weight that reflects the strength of that dependence. The stronger the dependence, the higher the weight assigned. One open question concerns the weight assigned to non-probabilistic cases, i.e. those that satisfy condition (2). Since the inferential connection between α, β via an atomic proposition δ is deductive and hence the strongest form of inference, there is a case to be made about assigning the highest possible weight to such connections.[3] Implicit in this move is the idea that non-probabilistic cases count more than probabilistic ones. This idea could be implemented by assigning to each non-probabilistic case the full weight of one connection and to each probabilistic case a real-valued fraction of one connection; the latter, of course, depending on the strength of the given probabilistic dependence.

[2] I have alluded to such a modification in Votsis (2016, p. 309, f14).

[3] Though there may be something to be said about judging the strength of these connections also via the number of atomic propositions that satisfy condition (2) for each specific pair.

Needless to say, these thoughts are rather tentative. But suppose, for argument's sake, that my account were indeed suitably modified so as to be sensitive to the strength of probabilistic dependence between propositions. Would the two accounts, mine and Myrvold's, be rendered identical? The answer to this question is negative. Indeed, it would remain negative even if the two accounts employed the same way to measure the strength of probabilistic dependence. The reasons are many. There are general differences. For example, Myrvold embraces Bayesianism. In attempting to model deductive relations separately, I implicitly rule it out. Myrvold opts for a subjective construal of confirmation. I insist on an objective one. But there are also particular differences. As I have just reminded the reader in the previous paragraph, my account's second condition attempts to capture an important type of confirmational connection. This is the connection that holds between two propositions when both are required to (relevantly) deduce a third. It is unclear whether Myrvold's account is capable of giving such cases their due. Moreover, my account embodies several vital restrictions, e.g. limiting deductive consequences through a criterion of relevance, whose purpose is to prevent the trivialisation of the notion of confirmational connection. Myrvold's account does not impose such restrictions. That leaves the door open to all sorts of vulnerabilities, no least of which confirmational paradoxes.[4] Alas, there is no space to explore such vulnerabilities here.

One aspect of Myrvold's view that perhaps mitigates against the risks outlined above is its scope. He is quite upfront about that scope not extending to everything the concept of unification is meant to cover. In his own words, "[n]o claim is being made that every case that one can reasonably regard exhibiting unification will be captured by this account" (p. 400). This humility is not only admirable but also pragmatically sound. Implicit in my earlier admission that my account has its faults was a similar kind of humility. The only difference is that I have not given up on the idea that apt modifications to my account may yet lead to a universally valid conception of unification. That is the hope, at any rate.

Allow me to bring this essay to a close by considering one more element that unites the two accounts. Both recognise the importance of the link between confirmation and unification. Moreover, and despite their aforementioned differences, the two accounts agree that hypotheses with unifying prowess earn extra confirmational boosts. On this view, unification is not a super-empirical virtue as it is customarily claimed, but an empirical one. If anything else, I hope that is the one take-home message imparted upon the reader.

[4]Myrvold briefly discusses and dismisses the danger posed by 'irrelevant conjunction' type scenarios – see (p. 410, f6). For a sustained critique see Lange (2004), where it is claimed that Myrvold's account is in some cases too easy and in others too difficult to satisfy.

ioannis.votsis@nchlondon.ac.uk

References

Friedman, M. 1974. Explanation and scientific understanding. *Journal of Philosophy* 71: 5–19.

Kitcher, P. 1989. Explanatory unification and the causal structure of the world. In *Scientific explanation: Minnesota studies in the philosophy of science*, vol. 13, ed. P. Kitcher and W. Salmon, 410–505. Minneapolis: University of Minnesota Press.

Lange, M. 2004. Bayesianism and unification: A reply to Wayne Myrvold. *Philosophy of Science* 71 (2): 205–215.

Myrvold, W. 2003. A Bayesian account of the virtue of unification. *Philosophy of Science* 70 (2): 399–423.

Newton, I. 1959. Newton to Briggs, 25 April 1685. In *The correspondence of Isaac Newton*, vol. II, ed. H.W. Turnbull, J.F. Scott, A.R. Hall, and L. Tilling. Cambridge: Cambridge University Press.

Schurz, G. 1991. Relevant deduction: From solving paradoxes towards a general theory. *Erkenntnis* 35: 391–437.

Schurz, G., and K. Lambert. 1994. Outline of a theory of scientific understanding. *Synthese* 101: 65–120.

Scriven, M. 1962. Explanations, predictions and laws. In *Scientific explanation, space and time*, Minnesota Studies in the Philosophy of Science: Vol. 3, ed. H. Feigl and G. Maxwell, 170–230. Minneapolis: University of Minnesota Press.

Van Fraassen, B.C. 1980. *The scientific image*. Oxford: Clarendon Press.

Votsis, I. 2015. Unification: Not just a thing of beauty. *Theoria* 30 (1): 97–113.

———. 2016. Ad hoc hypotheses and the monsters within. In *Fundamental issues of artificial intelligence*, ed. V.C. Müller, 299–313. Berlin: Springer.

Weinberg, S. 1979. Conceptual foundations of the unified theory of weak and electromagnetic interactions. Nobel Prize Award Lecture, http://www.nobelprize.org/nobel_prizes/physics/laureates/1979/weinberg-lecture.pdf.

Wilson, E.O. 2005. Systematics and the future of biology. In *Systematics and the origin of the species: On Ernst Mayr's 100th anniversary*, ed. J. Hey, W.M. Fitch, and F.J. Ayala. Washington, DC: The National Academies Press.

Woodward, J. 2014. Scientific explanation. *The Stanford Encyclopedia of Philosophy*, Winter 2014 Edition, ed. Edward N. Zalta. URL=http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/.

ioannis.votsis@nchlondon.ac.uk