



HAL
open science

Relevance, not Invariance, Explanatoriness, not Manipulability: Discussion of Woodward on Explanatory Relevance

Cyrille Imbert

► **To cite this version:**

Cyrille Imbert. Relevance, not Invariance, Explanatoriness, not Manipulability: Discussion of Woodward on Explanatory Relevance. *Philosophy of Science*, 2013, 80, pp.625 - 636. 10.1086/674002 . hal-01098487

HAL Id: hal-01098487

<https://hal.science/hal-01098487>

Submitted on 25 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Please use the published version when quoting or citing

Philosophy of Science, Vol. 80, No. 5, December 2013, pp. 625-636.

**Relevance, not Invariance, Explanatoriness, not Manipulability:
Discussion of Woodward's Views on Explanatory Relevance.**

Cyrille Imbert

<version before final corrections>

<please use the published version when quoting or citing>

cyrille.imbert@univ-lorraine.fr
Laboratoire d'Histoire des Sciences et de Philosophie – Archives Poincaré
(Université de Lorraine / UMR 7117 CNRS)

shorter version:

cyrille.imbert@univ-lorraine.fr
Archives Poincaré
(Université de Lorraine / CNRS)

Imbert, Cyrille. « Relevance, Not Invariance, Explanatoriness, Not Manipulability: Discussion of Woodward's Views on Explanatory Relevance ». *Philosophy of Science* 80, n° 5 (décembre 2013): 625-636. doi:10.1086/674002.

Abstract

According to Woodward's causal model of explanation, explanatory information is relevant for manipulation purposes and indicates by means of invariant causal relations how to change the value of certain target explanandum variables by intervening on others. Therefore, the depth of an explanation is evaluated through the size of the domain of invariance of the generalization involved.

In this paper, I argue that Woodward's account of explanatory relevance is still unsatisfactory and claim that the depth of an explanation should be explicated in terms of the size of the domain of circumstances which it designates as leaving the explanandum unchanged.

Acknowledgments. I want to thank Andrea Woody for past discussions about the issue of scientific explanation and for her help and patience as an editor. All remaining shortcomings are mine.

Relevance, not Invariance, Explanatoriness, not Manipulability: Discussion of Woodward's views on Explanatory Relevance.

1. Introduction

The question of explanatory relevance has long been a challenge for explanation theorists. It is well-known for example that Hempel's DN model, Salmon's SR model or Salmon's causal models fail to philosophically characterize what type of information is relevant to the explanation of a given fact and should therefore figure in its explanation.

Over the last two decades, James Woodward has developed a manipulationist model of explanation, which seems to fare better than previous accounts on the issue of explanatory relevance and how to solve the usual tricky cases. In this model, explanatory information is information which is relevant to manipulation or control and indicates how to change the value of certain target explanandum variables by intervening on others. Accordingly, the depth of an explanation is evaluated through the size of the domain of invariance of the generalization involved.

In this paper, I argue that Woodward's treatment of relevance in terms of invariant causal relations is still subtly but unavoidably wanting because it forces one to include within the explanation of a fact F a great deal of information which may be relevant to account for other facts of a same physical type but still irrelevant to F. I further claim that the depth of an explanation can be evaluated through the size of the domain of circumstances which it describes as leaving the explanandum phenomenon unchanged.

In section 2, I briefly present Woodward's account of explanation and his notion of explanatory depth. In section 3, I develop at length a test case example dealing with the explanation of the law of Areas and describe two ways to explain this physical regularity. In section 4, I show that, while the first explanation includes clearly irrelevant facts, according to Woodward's account, it cannot be said to be less explanatory than the second. I further analyze why satisfying the manipulability requirement may imply including irrelevant facts in explanations in order to make them deeper (in Woodward's sense). I further describe in section 5 a new criterion for judging explanatory depth and argue that this criterion is incompatible with Woodward's criterion. I finally emphasize in section 6 that manipulability is still a virtue, even if not an essential virtue of explanations and that, depending on the circumstances, it may be of interest to develop explanations which are less explanatory (because they contain irrelevant facts) but still enable a degree of control of physical systems.

2. Woodward's Manipulationist Account of Explanation

It may seem strange to challenge Woodward (and Hitchcock) on the question of explanatory relevance for they have themselves showed much acumen in diagnosing where

existing accounts fail and offered new answers to the problem. Indeed, in his 1995 article, Hitchcock elegantly shows that the problem of explanatory relevance is still a problem for Salmon's causal model because identifying all the intermingled spatio-temporal causal processes at work in some physical circumstances still fails to indicate exactly why the target phenomenon occurs in those circumstances. As Woodward further notes, even if the right causal processes are identified, "features of a process P in virtue of which it qualifies as a causal process (ability to transmit mark M) may not be the features of P that are causally or explanatorily relevant to the outcome E that we want to explain" (Woodward 2003, 353).

In this context, it comes as no surprise that Woodward tries to answer the above concerns by means of his causal model. This article is too short to do full justice to all aspects of Woodward's rich treatment of explanatory relevance and explanation and therefore the next paragraphs are merely intended to remind the reader some important aspects of Woodward's account to clarify the ideas prior to the subsequent analysis of the example.

For Woodward, "explanation is a matter of exhibiting systematic patterns of counterfactual dependence" (2003, 191). The explanatory generalizations used in an explanation must indicate that the explanandum was to be expected *and* how it would change, were the circumstances that obtained different; said differently, good explanations "are such that they can be used to answer a range of counterfactual questions about the conditions under which their explananda would have been different" (*ibidem*).

In this perspective, "explanatory relevant information is information that is potentially relevant to manipulation and control" (2003, 10). In other words, something is relevant information if it essentially figures in an explanation making clear that the explanandum was to happen and describing how it would change, if the properties described in the explanans were modified. This requirement also discards irrelevant circumstances through the identification of irrelevant variables: "an explanans variable S is explanatorily irrelevant to the value of an explanandum variable M if M would have this value for any value of S produced by an intervention" (2003, 200).

Woodward further defines the notion of invariance of a generalization. A generalization may remain stable under various changes of conditions which are not mentioned in this generalization. For example, Coulomb's law holds under changes in the weather. By contrast, a generalization that "continues to hold or is stable in this way under some class of interventions that change the conditions described in its antecedent and that tells us how the conditions described in its consequent would change in response to these interventions is invariant under such interventions" (1997, S.31).

It is clear then that invariance is a gradual notion because a generalization can hold under more or less interventions. Accordingly, depending on the degree of invariance of the generalization they rely upon, explanations provide patterns for answering more or less what-if explanatory requests about these counterfactual circumstances and therefore for controlling the corresponding systems.

Woodward further claims that the concept of invariance provides a means for evaluating how good explanations are – what he calls "explanatory depth": "We can thus make comparative judgments about the size of domains of invariance and this is all that is

required to motivate comparative judgments of explanatory depth of the sort we have been making” (1997, S.39). More briefly, the more invariant an explanation, the more explanatory it is, or to use Woodward’s own words: “generalizations that are invariant under a larger and more important set of changes often can be used to provide better explanations and are valued in science for just this reason” (2003, 257).

At this point, my claim can be precisely formulated thus: even if they are valued in science, more invariant explanations are not always more explanatory because the request for invariance may be contrary to the fundamental request for relevance that explanations should primarily satisfy.

3. The Law of Areas and its Explanations

The test case I now want to investigate is the explanation of the law of Areas (also called "Kepler's second law"), which states that, “for planets in our solar system, a line joining a planet and the sun sweeps out equal areas during equal intervals of time”. I shall describe two explanations of it and compare them with respects to invariance and relevance.

As we shall see, the first explanation (hereafter explanation 1) relies upon the general angular momentum theorem. Let us assume a Galilean reference frame, a fixed axis M' with position given by vector \mathbf{r}' and a moving material point with position given by vector \mathbf{r} , having mass m and momentum \mathbf{p} (bold characters denote vectors). The angular momentum of M about M' is defined by: $\mathbf{L}_{r'} = (\mathbf{r}' - \mathbf{r}) \times \mathbf{p} = m(\mathbf{r}' - \mathbf{r}) \times \mathbf{v}$, where the symbol “ \times ” stands for the usual external product. Let \mathbf{F} denotes the sum of forces applied to M . The momentum of \mathbf{F} about axis M' or torque is defined as $\mu_{F/M'} = (\mathbf{r}' - \mathbf{r}) \times \mathbf{F}$. Then, deriving the angular momentum yields

$$\frac{d\mathbf{L}_{r'}}{dt} = \frac{d((\mathbf{r}' - \mathbf{r}) \times \mathbf{p})}{dt} = (\mathbf{r}' - \mathbf{r}) \times \frac{d\mathbf{p}}{dt} + \frac{d(\mathbf{r}' - \mathbf{r})}{dt} \times \mathbf{p}$$

Because the momentum \mathbf{p} is collinear to the speed of M , the second term in the right-hand part of the equation is null. So far no physics has been used. Newton’s second law says that $d\mathbf{p}/dt = d(m\mathbf{v}) / dt = m\mathbf{a} = \mathbf{F}$. So finally, this gives

$$(1) \frac{d\mathbf{L}_{r'}}{dt} = (\mathbf{r}' - \mathbf{r}) \times \mathbf{F} = \mu_{F/M'}$$

For a collection of particles, one can also define the total torque $\mu = \sum \mu_i$ (the sum of the torques on each particle) as well as the total angular momentum \mathbf{L} (the sum of momentum of each particle) and this gives (1.5) $\mu = \sum \mu_i = d\mathbf{L}/dt$. The total torque is the sum of the momentum of all internal and external forces. However, Newton’s law of action and reaction means that the torques on two reacting objects compensate and therefore, the internal torques balance out pair by pair. In conclusion, “the rate of change of the total angular momentum about any axis is equal to the external torque about that axis”. This is the general angular momentum theorem, which is true for any collection of objects, whether they form a rigid body or not.

If one wants to explain the law of Areas, one should finally note that, in the case of the Earth/Sun two-body system, if \mathbf{v}_E denotes the speed of the Earth, \mathbf{r}_E its position, \mathbf{F}_G the gravitational force, \mathbf{L}_E the Earth momentum about the Sun, α the angle between \mathbf{r}_E and \mathbf{v}_E ,

and $A_E(t)$ the swept area in function of time, in virtue of the definition of the outer product,

$$(2) \frac{\|\mathbf{L}_E\|}{m_E} = \frac{\|\mathbf{r}_E \times \mathbf{v}_E\|}{m_E} = \|\mathbf{r}_E\| \|\mathbf{v}_E\| \sin(\alpha) = 2 \cdot \frac{dA_E(t)}{dt}$$

Because this relation holds for each mass point, the relation $\mu = \sum \mu_i = dL/dt$ can now be seen as describing the variation of the variation of the sum of the mass weighted areas swept by each mass point of a system about an axis, be it a rigid body or a set of independent mass points.

In the case of the Earth-Sun system, it should further be noted that the momentum of the gravitational force \mathbf{F}_G about the Sun is zero (because the force and the vector \mathbf{r} are collinear). Therefore, because of (1.5), the angular momentum of the Earth about the Sun is constant and because of (2), $A(t)$ grows linearly with time, which demonstrates that the law of Areas holds true in such cases.

This explanation perfectly fits Woodward's account of explanation and, *mutatis mutandis*, one may simply repeat here Woodward's analyses of the paradigmatic explanation in terms of Coulomb's law of the electrostatic relation $E = \lambda / (2\pi\epsilon_0 r)$ (203, 196-204). The explanation exhibits the features emphasized by DN theorists: it is a deductively valid argument in terms of Newton's second law and the description of the system (positions, speeds and masses of the points, forces). But in addition, it also exhibits a systematic pattern of counterfactual dependence which can be summarized by combining (1.5) and (2) into the general relation

$$(3) \mu = dL/dt = 2 \sum (m_i \frac{d(dA_i(t)/dt)}{dt}).$$

When the right variables are assigned the right values (two bodies, one central gravitational force, etc.), the law of Areas is a special case of (3). Overall, the explanation shows how the explanandum law of Areas would change according to (3) and how it systematically depends on Newton's second law, the forces and the particular conditions cited in the explanans. More specifically, the explanation makes clear how the total swept area would vary were the mass, speed, position of the Earth different, were additional forces at play but also if additional bodies were included in the system. In short, (3) and the explanation including it indicate how to answer a range of what-if questions about counterfactual circumstances in which the explanandum have changed. Given the range of these questions and the invariance of the explanation, it is difficult to do better, because Newton's law and (3) cover all situations in classical physics and therefore all classical changes that can be brought about to the two-body system case.

Let us now turn to the second explanation (hereafter explanation 2). In order to clarify why it can be considered better, I shall give two versions of the explanation, one of which will be more pictorial. Let us start with the vectorial derivation. Because of relation (2), the law of Areas holds true if the intensity $\|\mathbf{L}_E\|$ of the angular momentum \mathbf{L}_E of the Earth about the Sun is constant. In virtue of relation (1), this happens when $(\mathbf{r}' - \mathbf{r}) \times d\mathbf{p}/dt = 0$, which is the case if $d\mathbf{p}/dt$ and $(\mathbf{r}' - \mathbf{r})$ are collinear. This is so because the only force at play is radial and the variation of momentum of a particle is along the direction of the force

exerted upon it, that is $d\mathbf{p}/dt=\alpha\mathbf{F}$, where α is real, not necessarily constant and not specified. Newton provides a more geometrical way to see the explanation:

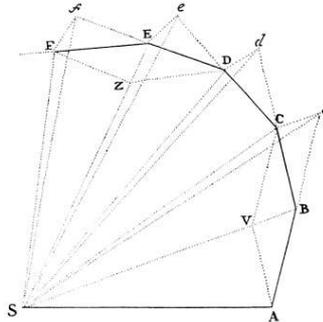


Figure 1: Geometrical demonstration of the law of Areas by Newton (1726)

The Earth's trajectory goes through A, B, C, etc. and the law of Areas holds true if the areas of SAB, SBC, etc. are equal. The explanation of each trajectory step is broken down into two parts. Firstly, if no force was at play, because of the inertia principle, the Earth would go from B to c in one time step and therefore the areas of SAB and SBc are equal. Secondly, if the Earth was motionless in B, because of the central gravitational force, it would go somewhere on (SB), say in V. By combining the two moves, the Earth finally goes to C, with $\mathbf{BV}=\mathbf{cC}$; therefore, the areas of SAC and SBc are also equal, which proves the result. The important point is that the numerical equality between the areas of SBc and SBC holds true whatever the position of V on (SB), and thus, provided that the change of momentum due to a force is along the force direction, that is, $d\mathbf{p}/dt=\alpha\mathbf{F}$.

How good is this second explanation? Firstly, it also exhibits the features emphasized by DN theorists in that it is a deductively valid argument in which a nomological component is essentially needed (as well as the description of some particular circumstances). It shows in addition that the whole content of Newton's second law is not required within the explanation. More precisely, the quantitative part of Newton's second law, which relates the values of forces and acceleration, and the quantitative description of the gravitational force, can be removed for the premises without altering the validity of the argument. Better, from a physical point of view, this removal brings some important piece of explanatory information because it indicates more specifically what in the physics is essential for the law of Areas to be the case. The quantitative aspect of the momentum variation is shown to be explanatorily irrelevant, which indicates that the law of Areas does hold true for all worlds with a dynamical law such that the variation of momentum is along the force direction. Explanation 1 does not provide this piece of explanatory information because it includes the above-mentioned irrelevant facts.

Accordingly, explanation 2 is also instrumental in answering what-if questions about what would happen if the intensity of the force were different, the time were discrete or the gravitational constant were to change with time. So, the corresponding explanatory generalization is also invariant under a large range of interventions.

4. Comparison between the two Explanations regarding Depth and Diagnosis about the Inadequacy of Woodward's Account

Let us now see how the two explanations fare comparatively according to Woodward's criterion of explanatory depth. As mentioned above, both explanations are invariant under a large range of interventions. As we saw, Woodward suggests assessing explanatory depth by comparing domains of invariance. In the present case, neither of the two explanations can then be said to be deeper than the other because neither of the two sets is a subset of the other. Indeed, explanation 1 directly yields answers to what-if questions about how the total swept area quantitatively changes when, say, non radial forces are at play or more bodies are involved, which explanation 2 does not (because it omits the quantitative part of Newton's second law). Conversely, explanation 2 explicitly indicates that the law of Areas would still hold true in circumstances in which Newton's second law would be violated – which explanation 1 does not – because it considers the whole law with its quantitative aspect to be explanatorily relevant. Overall, from Woodward's perspective, we have a situation with two good explanations whose explanatory depth cannot be compared because their domains of invariance only overlap partially. This is indeed a case which Woodward accommodates when he notes that the comparison of the domains of invariance of explanations “obviously yields only a partial ordering” because “for many pairs of generalizations, neither will have a range of invariance that is a proper subset of the other” (2003, 262-64).

My point is that this Woodwardian conclusion is not satisfactory. If one focuses upon the relevance of the explanatory material regarding the explanandum, explanation 2 is better than explanatory 1. It is indeed commonly agreed that an explanation of A should merely include explanatory information which is relevant to the occurrence of A (at least if one's epistemic goal is to provide an explanation of A that is as explanatory as possible, see section 6 for more comments about this restriction). As mentioned earlier, explanation 2 omits explanatory material which is irrelevant to the occurrence of the law of Areas, while explanation 1 does not. It is then no surprise that explanation 1 provides an answer to many what-if questions whose answers depend on this irrelevant material and cannot therefore be given by explanation 1. However, while these additional answerable questions contribute to extend the invariance of explanation 1, the ability to answer them should not be seen as a sign of the greater value of explanation 1 (quite the contrary!) because, as the Newtonian investigation described above shows, answering them requires certain causal information which is explanatorily irrelevant here. In brief, information present in fundamental causal laws, though true in all the corresponding models, should not necessarily be part of the explanations of all statement true in these models.

Let us now try to see more clearly why Woodward's account leads to irrelevant features being included in explanations to make them deeper. The reason seems to be that he requires an explanation to account for many counterfactual cases belonging to the same physical type, defined in terms of the explanandum variable appearing in the explanatory generalization, and of which the explanandum fact is an instantiation. But this compels him to include in the explanatory material not only facts which are explanatorily relevant to the target explanandum but also facts which are explanatorily relevant to the other values the

explanandum variable may take. But as the example shows, the explanatorily relevant facts for the latter and the former need not coincide. The conclusion then is that facts belonging to an identical type do not always have the same explanations or explanations of the same type.

Here, it is important to note that the explanandum type about which this conclusion is drawn (the variation of the swept area) is not the product of conceptual gerrymandering. So the conclusion needs to be rephrased more precisely and strongly like this – facts belonging to an identical *bona fide* physical type (corresponding to the explanandum variable of a genuine physical generality) do not always have the same set of explanatory relevant facts or explanations of the same type.

This conclusion has a counterpart in terms of whether domains of invariance are appropriate for the assessment of the depth of an explanation and which what-if erotetic requests are appropriate for this task. If we require an explanation of a target explanandum fact F to allow us to answer what-if questions about counterfactual circumstances corresponding to the invariance domain of some general and functionally described regularity of which the explanandum case is an instance, this may imply physical information that is explanatorily relevant for these circumstances (but not for F) being included in the explanatory material. Accordingly, even if these explanatory requests are by themselves scientifically legitimate, it may be illegitimate to judge the explanatory value or depth of an explanation of F by its capacity to answer these requests because the physical information necessary for this task may be explanatorily irrelevant regarding F – and this information should therefore not be included in a good explanation E of F . In short, what-if questions about some circumstances in the domain of invariance of the explanatory generalization used in the explanation E may still not be appropriate questions for testing the depth of E because the corresponding criterion is incompatible with a satisfactory treatment of the problem of relevance for explanations.

The conclusion regarding the evaluation of explanatory depth in terms of domain of invariance comes naturally. It is not legitimate to evaluate the depth of an explanation by assessing the domain of invariance of the generalization used therein. Performing well on the invariance criterion leads to the promotion of explanations of individual facts which are special cases of general explanatory patterns built on generalizations that are invariant on large domains... but it potentially also leads to the violation of the requirement of relevance for the explanations of these individual facts.

5. Another Criterion for Explanatory Depth

As can be inferred from the discussion of the example, a good explanation (which satisfies the criterion of explanatory relevance) still seems to provide answers to many *appropriate* what-if questions. Explanation 2 shows that the law of Areas would still hold true in many circumstances in which the quantitative part of Newton's second law or the intensity of the gravitational law would be different. It thereby enables an affirmative answer to the corresponding "would-the-explanandum-still-be-the-case" (in short "would-still" questions). For a derivative explanation, this set of circumstances in which the explanandum is shown by an explanatory argument to remain unchanged corresponds to

the set of situations in which the premises of the explanatory argument are true. Further, the more irrelevant information is removed from the premises, the weaker these explanatory premises and the wider the class of situations to which they apply. Let us call this class of situations the domain of strict invariance of the explanation (by contrast with Woodward's notion of domain of (large) invariance of the generalization employed in the explanation, hereafter "large invariance"). In this way, the above discussion leads to the following suggestion:

(S) The wider the domain of strict invariance of an explanation, the deeper the explanation.

Much more discussion would be required to develop this suggestion into a fully-fledged proposal about the nature of explanation. In particular, a critical comparison with notions discussed by Reichenbach or Salmon in different contexts such as the notions of broadest homogeneous reference class, maximal class of maximal specificity or exhaustiveness (Kitcher and Salmon 1989, 69, 104, 193) would be helpful. Nevertheless, the following remarks are relevant here. Firstly, (S) indicates how an explanation can be improved by expurgating its premises from irrelevant information but does not however indicate in general what type of information may be present in the premises for something to count as a potential explanation. Therefore, it should not be seen as an independent stand-alone criterion (otherwise, the best explanation would be the self-explanation of one fact by itself). Secondly, the domain that is described here should be distinguished from the scope of the laws or the domain of invariance of the generalization present in the premises, which characterize statements – strict invariance characterizes the explanation itself. Alternatively it can be seen as the domain of the explanatory generalization saying that when the premises hold true (in this or different worlds), so does the explanandum. Thirdly, just as for Woodward's account, this criterion is likely to describe only a partial order over explanations. Finally, it should be noted that the criteria of having a large domain of large invariance and of having a large domain of strict invariance go into two opposite directions. Indeed, explanations with large domains of general invariance require generalizations with a great deal of physical information packed in them; whereas explanations with large domains of strict invariance require premises with as little physical information as possible in their premises. So it does not seem possible to try to conciliate both criteria about the nature of explanatory depth.

6. Concluding Remarks: Generality and Manipulability *versus* Specificity and Relevance or the Contextual Choice of Epistemic Virtues in Scientific Practice

In this article, I have criticized the use of the size of the domains of invariance of the generalizations used in explanation to describe the depth of these explanations. I have argued that this characterization of the depth of explanations fares badly by the requirement of relevance, which explanatory explanations should primarily satisfy. To describe the depth of explanations I have proposed a different criterion based on the notion of strict invariance and the ability to answer "would-still" questions offered by explanations and I have emphasized that satisfying one criterion may run contrary to the satisfaction of the other.

One final word of caution is needed here. The above analysis covers the explanatory character of explanations of specific individual facts of which relevance is clearly a component. Like all other things, explanations may also have unspecific additional virtues, which may be philosophically unessential to them but often crucial to their use. In the present case, having wide large invariance is no doubt one such unessential virtue. Indeed, an explanation with wide large invariance, even if it is of average quality regarding explanatory relevance, does provide a functional pattern for a family of similar explanations: it offers the opportunity to explain many similar phenomena with the same pattern of reasoning, which yields a significant saving in scientific and cognitive means. Like any versatile tool, it is general and therefore this kind of explanation may prove useful (and philosophically attractive) without necessarily being optimal for specific explanatory tasks. Thus, finding such explanations is a scientifically legitimate (and difficult) task.

So should scientists favor in practice specific relevant explanations with wide domains of strict invariance over general explanations with wide domains of large invariance? I think there is no general answer to this question. Pace the philosophical interest for essential epistemic virtues, contextual interests are to prevail and be guided by scientific or practical needs. If for example you are interested in controlling optical rays within optical fibers or the trajectory of a car in various circumstances, there is little doubt that you will be interested in finding explanations with wide domains of large invariance so that you can determine how the rays or the cars will behave in a wide range of circumstances with one single functional relation and control them by adopting the external forcing. For some circumstances covered, it is likely that this single functional relation will contain unnecessary (irrelevant) information and for some specific cases you may even be using a sledgehammer to crack a nut but why should this be of any concern for you? For control purposes, it may be more convenient to use one single relation covering all cases rather than a larger and more cumbersome selection of relations which each specifically target a subset of circumstances.

Suppose now that you are interested in observing a green flash effect (optical phenomena occurring after sunset or before sunrise when a green spot is visible above the sun). Here, you may wish to learn about the circumstances in which you stand a good chance of observing a green flash effect and may therefore require as large a set of circumstances as possible. Therefore, knowing which circumstances will not alter the phenomenon (because they are irrelevant to the mechanism involved) is crucial. In this case, you will be interested in discarding any irrelevant information that restricts your knowledge of this set even if this also means leaving out of the explanans physical information that may be useful to answer questions about what would happen in close circumstances (in which no green flash effect is observed). So you may end up with an explanation that is not useful for manipulationist purposes because it is specifically targeted at the green flash effect. Perhaps this explanation will not even have a functional form (like above the explanation 2 of the law of Areas) but, because its explanans only describes the physical facts that are crucial for the green flash effect to happen and discards the others, it will be more explanatory and therefore more informative about the whole range of circumstances in which the observation can be made.

Please use the published version when quoting or citing

Philosophy of Science, Vol. 80, No. 5, December 2013, pp. 625-636.

In conclusion, Woodward's criterion for explanatory depth seems more appropriate to characterize explanations which are useful for control than those which are deeply explanatory.

Please use the published version when quoting or citing

Philosophy of Science, Vol. 80, No. 5, December 2013, pp. 625-636.

References

Hitchcock, Christopher, 1995. "Discussion: Salmon on Explanatory Relevance." *Philosophy of Science*, 62: 304-20.

Kitcher, Philip and Salmon, Wesley, 1989, *Scientific Explanation*. University of Minnesota Press, Minneapolis.

Newton, Isaac. 1726. *Philosophiae Naturalis Principia Mathematica*.

Woodward, James 1997. "Explanation, Invariance and Intervention." *PSA 1996 2*: S-26-41, S26-S41.

Woodward, James. 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford University Press.