The Structure of Causal Evidence Based on Eliminative Induction[1]

Wolfgang Pietsch[2], Philosophy of Science, Carl von Linde-Akademie, Technische Universität München, Germany

Abstract: It will be argued that in deterministic contexts evidence for causal relations states whether a boundary condition makes a difference or not to a phenomenon. In order to substantiate the analysis, I will show that this difference / indifference making is the basic type of evidence required for eliminative induction in the tradition of Francis Bacon and John Stuart Mill. To this purpose, an account of eliminative induction will be proposed with two distinguishing features: it includes a method to establish the causal irrelevance of boundary conditions by means of indifference making, which will be called strict method of agreement, and it introduces the notion of a background against which causal statements are evaluated. Causal statements thus become three-place-relations postulating the relevance or irrelevance of a circumstance C to the examined phenomenon P with respect to a background B of further conditions. To underline the importance of evidence in terms of difference / indifference making, I will sketch two areas, in which eliminative induction is extensively used in natural and engineering sciences. One concerns exploratory experiments, the other engineering design methods. Given that a method is discussed that has been used for centuries, I make no claims to novelty in this paper, but hope that the combined discussion of several topics that are still somewhat underrepresented in the philosophy of science literature is of some merit.

1. Introduction

Causation has been a thriving topic in recent decades, but the discussion has largely focused on methods of probabilistic causality and their application in complex contexts within the social sciences, economics, medicine, or artificial intelligence. Certainly, causality is at least as important in the natural and engineering sciences but there are a number of conceptual and methodological differences. Maybe most importantly, in these sciences the phenomena in question can usually be isolated under laboratory conditions, i.e. they can be produced in situations where one has good knowledge and control of all or at least most potentially relevant boundary conditions. Relatedly, experimental research in these sciences almost always happens under an assumption of determinism, i.e. that the phenomenon in question is fixed by a combination of the relevant boundary conditions.

The structure of causal evidence, i.e. evidence for causal relations, is determined by the empirical input that the corresponding scientific method requires. Therefore, one major concern in this paper will be to present a method, with which causality can be established in deterministic contexts. In this respect, Mill's method of difference is usually evoked which in turn forms part of a larger set of methods that are usually subsumed under the label *eliminative induction*. In the past, several accounts of eliminative induction have been

---

[1] forthcoming in a special issue of *Topoi* on causality and evidence, edited by Phyllis Illari and Federica Russo

[2] pietsch@cvl-a.tum.de

formulated, with precursors in the Middle Ages and culminating in the two most celebrated versions, namely Francis Bacon's method of exclusion (1620/1994, Bk. 2) and John Stuart Mill's methods of induction (1886, Bk. III, Ch. VIII). While both accounts have been heavily criticized, there is little doubt that they contain crucial elements of a scientific method that is widely used in experimental and engineering research. Somewhat curiously, as of today, no account of eliminative induction seems to exist that is generally agreed upon. For whatever reasons, the exercise appears to be more difficult than expected. We are thus faced with the slightly paradoxical situation that eliminative methods are employed successfully in both scientific and everyday contexts although we do not yet possess a fully worked-out systematic account.

Note that in the literature, eliminative inference refers to two distinct methodologies, which are only remotely connected. In this essay, the term 'eliminative induction' will be used in the more narrow sense of Bacon and Mill to describe an inductive method which aims to establish causal relations by examining a phenomenon under varying boundary conditions. In the process, conditions are either proven to be causally relevant to a phenomenon or are shown to be irrelevant and are thus eliminated. The terminology is generally attributed to Mill (1886, 256) and is for example taken up by Mackie (1980, 297). Call this eliminative induction$_1$, which will be presented in considerable detail in Section 3.

Today, the term 'eliminative inference' is often used in the much broader sense of eliminating any kind of rival hypotheses. Such eliminative inference$_2$ is in fact quite distinct from eliminative induction$_1$, both in form and spirit. First, the latter requires a specific framing of research questions, in which the relevance of boundary conditions can be systematically examined, while in eliminative inference$_2$ not much is presupposed about the form of hypotheses. Second, under certain premises that will be discussed later in the paper, some eliminative methods$_1$, notably the method of difference, allow for a *direct* identification of causes. By contrast, eliminative induction$_2$ always proceeds by a stepwise elimination of single or at most classes of hypotheses—a direct inference is usually impossible. Finally, eliminative inference$_2$ is closely linked to a Popperian hypothetico-deductive account of scientific method, which is generally hostile to induction. By contrast, eliminative induction$_1$ is meant to be a powerful inductive method, which under certain premises can reliably identify causal relations.

The structure of the paper is as follows. After a brief comparison between enumerative and eliminative induction as well as the corresponding types of evidence in Section 2, I will present an account of the core principles of eliminative induction in Section 3. Arguably, these are the *method of difference* and a complementary method, which will be called the *strict method of agreement*. The former allows establishing causal relevance, the latter causal irrelevance. On the basis of these methods, we show how the method of concomitant variations can determine a causal influence of varying degree and under which circumstances enumerative induction is appropriate. Also, I will shortly discuss the *problem of eliminative induction*, i.e. under which additional premises eliminative inferences are valid and thus can be used for predictions and interventions. Most importantly, the principle of causality is among these premises explaining why eliminative induction only works in deterministic contexts.

On the basis of this exposition, I will provide in Section 4 a detailed discussion of the structure of evidence for causal relevance and irrelevance. The form is difference / indifference making, i.e. a comparison of two situations which differ only in one potentially relevant boundary condition and in the possible impact on the phenomenon in question. Finally, Section 5 will assess to what extent eliminative induction and evidence in terms of difference / indifference making are used in scientific practice. First, I will argue that exploratory experimentation, which tries to map the causal structure of a yet unfamiliar phenomenon, closely mirrors the requirements for eliminative induction. Second, I will show that crucial design methods in the engineering sciences, geared at the optimization of artifacts, implicitly employ evidence of the mentioned type.

The whole paper will have the character of an overview trying to link up several topics that are still to varying degree underrepresented in the philosophy of science literature, in particular eliminative induction, scientific experimentation, and engineering design.

2. Causality and evidence

The nature of evidence with respect to a scientific method is usually discussed in the context of scientific confirmation. In the following, I will concentrate on two broad classes of inductive methods and their corresponding types of evidence, namely eliminative and enumerative induction. Eliminative induction has been defined in the introduction and will be discussed in much detail in Section 3. Enumerative induction constitutes the simplest kind of an inference from particulars to universals relying on the following rather primitive scheme: Starting from the observation that in a number of instances whenever C is given, P is found to be present as well, we infer the general law that C causes the phenomenon P. A probabilistic generalization of this approach is straight-forward. Enumerative induction has been known since the ancients, some version is discussed for example in Aristotle's *Posterior Analytics*[3].

In enumerative induction, the generic type of confirmation is instance confirmation. Observing a black raven is considered evidence for the corresponding universal hypothesis. A negative instance either falsifies the hypothesis or makes it less probable, depending on whether the original hypothesis was deterministic or statistical. Instance confirmation establishes a correlation between types of facts. Since simple correlation does not imply causation, correlational data must be supplemented with additional assumptions in order to get causal information. There exists an abundant literature on this problem. Arguably, the most sophisticated approaches concern Bayesian networks as developed in Pearl (2000) or Spirtes, Glymour and Scheines (2000). The mentioned additional premises range from graph-theoretical restrictions like minimality or faithfulness to requiring evidence based on interventions. Note that complex Bayesian networks can only be set up with probability distributions that correlate a large number of parameters. Such evidence in fact resembles the type required for eliminative induction, as we will see now.

---

[3] 'we may not proceed as by induction to establish a universal on the evidence of groups of particulars which offer no exception, because induction proves not what the essential nature of a thing is but that it has or has not some attribute' (Aristotle 1928, II,7, 92a36-b1).

In eliminative induction, the interest shifts from the mere repetition of instances to the examination of phenomena under as large a variety of circumstances as possible.[4] In other words, eliminative induction requires as evidence a much more detailed description of the observed events compared with enumerative induction. As a general rule, the larger the number of parameters considered, the better. A further difference is that in eliminative induction no fixed hypothesis is tested, but rather new evidence always extends the causal knowledge in stating that certain boundary conditions are or are not causally relevant to a phenomenon. Strictly speaking, positive instances in the sense of an exact repetition under exactly the same circumstances have no confirmatory value at all in eliminative induction (Keynes 1921, Ch. XX). Negative instances do not falsify, but just amend the causal knowledge by identifying boundary conditions under which the phenomenon fails to occur. In this way negative and positive instances are quite symmetric in their inferential power, much more so than in enumerative induction.

An interesting case, where the evidence required for an eliminative method is discussed in the recent literature, concerns debates regarding the evidential status of randomized controlled trials.[5] As is typical for eliminative induction, detailed attention is given to the boundary conditions, in particular regarding the careful choice of adequate study and control populations involving strategies like randomization, double blinding, and placebo controls. Debates about the required size of the populations, possible confounders, or external validity all underline the concern that neglected boundary conditions could affect the reliability and generalizability of the results.

While a thorough methodological analysis of various practices of scientific experimentation still largely remains a desideratum for philosophy of science, recent work has identified a type of exploratory experimentation that closely follows the rationale of eliminative induction (Steinle 1997, 2005; Anderson 1997; Waters 2007). In Section 5, the link between this type of experimentation and eliminative induction will be made explicit. Whether we are dealing with such exploratory experimentation or with randomized controlled trials, eliminative induction rather than enumerative induction forms the methodological basis of experimental practice, when it aims at causal knowledge.

Not surprisingly, many proponents of eliminative induction have been quite critical of enumerative induction, including the two most influential figures Francis Bacon and John Stuart Mill. The former writes that ‘the induction […], which proceeds by simple enumeration is a childish affair, unsafe in its conclusions, in danger from a contradictory instance, taking account only of what is familiar, and leading to no result.’ (Bacon 1620/1994, 21) Mill agrees: ‘[Enumerative induction] is the kind of induction which is natural to the mind when unaccustomed to scientific methods. […] It was, above all, by pointing out the insufficiency of this rude and loose conception of Induction that Bacon merited the title so generally awarded to him of Founder of the Inductive Philosophy.’ (Mill 1886, 204)

---

[4] In the 20th century, the classic accounts which most stress the relevance of boundary conditions are Keynes (1921) and Mackie (1980). More recently, the tradition was picked up in particular by Baumgartner and Graßhoff (2004). Russo (2007) also stresses the significance of variation in comparison with regularity.

[5] For a concise overview see Howick (2011) and further references therein.

Until about the end of the 19th century, Bacon was generally considered the founding figure of the modern experimental method. Only in the 20th century with its widespread distrust in universal accounts of scientific method, discussions of eliminative induction have almost entirely vanished with a few notable exceptions (e.g. von Wright 1951; Mackie 1980; Skyrms 2000; Baumgartner & Graßhoff 2004; Norton 2010[6]). In spite of this shift in perspective, eliminative induction remains as important as ever for the experimental practice of science and the production of artifacts in engineering. Let us now proceed to discuss the most important elements of this method.

3. Eliminative induction revisited

3a. The structure of causal statements

In eliminative induction a phenomenon P is examined under variable boundary conditions C1, …, CN, which are all assumed to be potentially relevant for P, i.e. for which it is yet unknown if they are causally relevant or not. In addition, there is a background B of further boundary conditions which are either held fixed, if they are relevant or potentially relevant for P, or they are allowed to vary, if they are known to be irrelevant for P.

Thus, causal statements are three-place-relations of the following form: 'Boundary condition CX is causally relevant / irrelevant to phenomenon P with respect to background B.' By definition: In the case of *causal relevance*, a change in CX *always* leads to a change in P. Similarly, in the case of *causal irrelevance*, a change in CX *never* has an impact on P. Cases, in which a change in CX only sometimes has an impact on P can be ignored if the background is always adequately formulated (assuming that every event is fully determined by its boundary conditions).

A number of difficulties arise. For example, it is not at all clear which boundary conditions to list among the C. In particular, one should be able to change each of the boundary conditions separately while holding all others fixed in order to apply the method of difference or the strict method of agreement. However, any CX will have a number of causal implications that automatically change with it. Also, in order to influence the state of a specific CX, a number of external circumstances have to be changed as well. Thus, it is by no means obvious, which boundary conditions to include among the C. Certainly, these *basic parameters* will generally involve a whole cluster of individual conditions.

The introduction of a background adequately mirrors the fact that causal statements are usually contextual, i.e. they are never or at least rarely strictly universal. Causal laws hold only *ceteris paribus*, i.e. if a usually infinite number of additional conditions are satisfied. A bullet, even if directed at the heart of someone, only kills in the absence of an infinite number of preemptive causes. When it is impossible to list all relevant boundary conditions, the introduction of a causal background, which isolates crucial factors with respect to a certain context, becomes mandatory.

---

[6] While Norton uses the term eliminative induction in the sense of eliminating hypotheses, he also discusses in quite some detail Mill's methods.

Note that according to this framework, the relevance or irrelevance of the conditions in B should in principle be relative to a background as well, just because all causal statements have the structure of three-place-relations. With respect to such further complications, let us be pragmatic and just assume that in standard applications the background is sufficiently robust. As in any empirical science, there always remains the possibility of unpleasant surprises.

3b. The method of difference and causal relevance

The method of difference allows determining the causal *relevance* of a boundary condition, for example a property, an object, or an event. In this method, two instances are compared which differ in only one of the circumstances CX and agree in all others. If in one instance both CX and P are present and in the other both CX and P are absent, then CX is causally relevant to P. More specifically, CX is causally relevant[7] to P with respect to a background B* consisting of the original background B plus all conditions C1, …, CN except of course CX. Notably, CX may fail to be causally relevant to P with respect to the original background B, since the definition of causal relevance in Section 3a requires that a change in CX *always* leads to a change in P, which may not be the case.

The method of difference shows that CX is a necessary[8] and sufficient condition for P with respect to the background B*. However, CX may fail to be a sufficient condition with respect to B, if one or more of the remaining conditions C1, …, CN turns out necessary for P as well. As pointed out already by Mill, CX may not even be necessary with respect to B, if P can be produced in several distinct ways.

It was one of John Mackie's (1965, 1980) great achievement to have pointed out the logical structure of causal statements in terms of necessary and sufficient conditions, though of course these ideas are already implicit in the writings of authors like von Wright or Mill. Mackie also stressed the context-relativity of causal statements by introducing the notion of a causal field (1980, 34-36). This concept, which Mackie borrowed from his teacher John Anderson (1938, xvi), essentially corresponds to the causal background with some minor differences in framing. Most importantly, Mackie uses the notion to distinguish between veritable causes and mere conditions in the field. In this way, he can account for the distinction between those causally relevant boundary conditions which are legitimate answers to why-questions and those which are not. In this paper we will not address these crucial pragmatic aspects due to lack of space.

3c. The strict method of agreement and causal irrelevance

Mill regards the method of difference as the most effective and most reliable for identifying causes: 'It thus appears to be by the Method of Difference alone that we can ever, in the way of direct experience, arrive with certainty at causes.' (1886, 258) In particular, the method of agreement for him has weaker inferential power and its results are less reliable. This method he construed in the following way: two or more instances are compared where P is always

---

[7] Note that the situation is symmetric if P and CX are exchanged. Therefore, additional constraints like time ordering have to be taken into account to establish a causal direction where necessary.

[8] Eliminative induction is not only able to identify the nomic necessity of causal laws, but also various other kinds of necessity, e.g. definitional necessity.

present but the circumstances C1, …, CN are either present or absent in various combinations. If the instances have only one circumstance CX in common, then this circumstance is a good candidate for the cause of P.

Besides identifying plausible candidates for causes, there is a further and arguably more important function of the method of agreement, namely that of identifying the *irrelevance* of boundary conditions. After all, Mill's version of the method can be read in two ways, as hinting at the causal relevance of CX or at the causal irrelevance of the other boundary conditions.

From this perspective, one can construct a version of the method of agreement that is analogical to the method of difference, is in its inferential power largely equivalent and just as widely used in experimental research. To distinguish it from Mill's general method of agreement, let me call it the *strict method of agreement*: Consider a situation where again *only one* circumstance CX is changed, but where the phenomenon P remains unchanged. We can then infer the causal irrelevance of CX for P with respect to the background B* consisting of B plus all conditions C1, …, CN except CX.

Let me add two remarks. First, as we will see in Section 3f, the strict method of agreement needs the same additional premises as the method of difference. In other words, the problem of induction for the method of difference is the same as for the strict method of agreement, justifying the claim of similar inferential power. Second, difficulties arise in connection with causal overdetermination. Consider a typical example, where someone is shot in the heart by two different persons A and B at the same time. Whose bullet then is the cause for death of the poor fellow? According to the strict method of agreement, neither A's nor B's bullet is causally relevant with respect to a background where the other person also shoots. According to the method of difference, they are both causally relevant, when the other person does not shoot. As an immediate consequence, causal relevance / irrelevance is undefined if the background leaves undetermined if the other person shoots. Finally, A's and B's bullets taken in conjunction are causally relevant with respect to plausible backgrounds. While this may sound odd, no formal contradictions result as long as one carefully keeps track of the background. The described situation in a way just constitutes the evidential signature for cases of overdetermination. Other counterfactual analyses, which fail to include background relativity and clear definitions of causal relevance and irrelevance, invariably lead to contradictions, when a bullet is identified as cause but then the counterfactual analysis fails in the case of overdetermination.

Counterintuitive features certainly remain, but then our intuitions are far from clear about cases of overdetermination—as for example Mackie has rightly stressed (1980, p. 47). For instance, it is strange that a composite event can be causally relevant to a phenomenon, while none of its component events is. But note that this occurs only on a specific level of coarse-graining. When we observe the trajectories of the bullets through the body according to a more fine-grained description, at least most of the overdetermination vanishes. If that were not the case one should eventually reconsider the description of the event. In particular, two events that have exactly the same effects on all levels of coarse-graining might just be the same event in the first place. With this we have barely scratched the surface of the highly complex possibilities for strategies of dealing with awkward causal signatures.

While the method of difference forms part of all standard accounts of eliminative induction, the strict method of agreement is not included. However, in scientific practice conclusions to causal irrelevance are ubiquitous. The reason we can include the strict method of agreement without further complications is intimately connected with the introduction of a background and the view of causal statements as three-place-relations as foreshadowed in Mackie's work.

3d. Derivation of the method of concomitant variations from the method of difference

Mill's method of concomitant variations stands out in spirit. While the other methods are concerned with the absence and presence of boundary conditions, the method of concomitant variations considers changes in degree and thus functional relationships. How a variation in degree might link up with the presence or absence of conditions or if there is any connection at all remains largely unclear in Mill's writings. To my knowledge, von Wright (1951, 83 & 161) for the first time pointed out the close relationship between the method of concomitant variations and the method of difference, but the remark is not fleshed out in much detail. So, let me try to suggest, what he may have had in mind.[9]

The method of concomitant variations can be derived from the method of difference with help of the following premises: i) that each circumstance CX, which is variable in degree, can be considered as the combined action of a large number of minute (in the limit infinitesimal) causes of the same type ΔCX; ii) that some *principle of composition of causes* can be assumed for the ΔCX, determining how they operate together (cf. also Mill 1886, Book III Ch. VI). An example is a simple independence assumption, i.e. independently of the present degree of CX, each ΔCX leads to the same increase ΔP of the phenomenon P, resulting in proportionality between CX and P. Of course, a functional relationship can never be fully proven experimentally and will always involve some interpolation, since the number of observations is always finite.

When all other circumstances are held constant and a small change ΔCX in the degree of CX results in a corresponding change of P, then the inference clearly relies on the method of difference. Therefore, the method of concomitant variations provides strong confirmation of causal dependence between CX and P, since every new observation of a change in degree counts as an independent application of the method of difference.

Obviously, there exists a corresponding version of the method of concomitant variations for the strict method of agreement: If CX can be changed in degree while no change in P is observed, this implies strong evidence for the causal irrelevance of CX for P.

Let me conclude with two remarks concerning the principle of composition of causes. First, if the influence of the ΔCX depends on the present degree of CX, more complex non-linear functional dependencies result. Second, in any causal analysis based on eliminative induction,

---

[9] Skyrms (2000, Sec. V.9) presents a similar story relying on a somewhat complicated reconstruction of continuous variables as families of physical qualities.

assumptions about the interaction of the various causes have to be made, which in first approximation are usually independence assumptions.

With the method of difference and the strict method of agreement plus the corresponding versions of concomitant variations, we now dispose of the tools to determine a set of conditions that are in their conjunction C sufficient for the production of P with respect to a certain B and where each condition forms a necessary part of C. In principle, we just have to examine for each circumstance C1, ..., CN, in which contexts it is relevant for P or not. The process leading to such a result could be called *parameter variation*, in that to gain full causal knowledge the whole parameter space has to be mapped, i.e. all possible combinations of parameter values have to be tried.

3e. The role of enumerative induction and the possibility of objective chance

In sketching an ideal process of eliminative induction, we have left numerous complications aside which for lack of space we cannot address here, for example the pragmatics of causation which causally relevant circumstances are singled out as causes in certain contexts. We should however briefly comment on how enumerative induction fits into the picture. As one might guess, this involves such intricate issues as the interpretation of probability or the foundations of statistics. Probably the most elaborate and convincing argument how enumerative induction is to be integrated into an eliminative perspective, can be found in Keynes (1921, Ch. XX).

Following Keynes, my basic take is that enumerative induction is less fundamental compared with eliminative induction, in that each application of enumerative induction must eventually be justified by means of eliminative methods. Arguably, enumerative induction works only in cases when we already have a clear conception about which circumstances are irrelevant, which are relevant and how potentially relevant circumstances change from one instance to the next. As an example, enumerative induction works especially well in the extreme case when P is in part due to objective chance. But such a situation can only be identified, if we have already excluded by eliminative induction all possible causes of P in terms of boundary conditions, and maybe more importantly, if we have identified all boundary conditions that must be instantiated such that P is exclusively due to chance.

Enumerative induction normally stands at the end of the process of eliminative induction when one already has a good understanding of which boundary conditions are relevant. It can be used to check for relevant boundary conditions in the background B whose influence is not yet accounted for, to estimate the influence of relevant boundary conditions which are beyond experimental control as in error theory, or to determine the presence of objective chance.

3f. The problem of eliminative induction

The problem of induction can be reformulated in terms of the Kantian question: Which additional premises would make an inductive inference reliable? Just as Kant’s epistemology is an enquiry about the a priori conditions for the possibility of experience (‘Bedingungen der Möglichkeit von Erfahrung’), we will now determine conditions for reliably identifying causal connections in the world. Thus, the problem of a specific inductive method consists in formulating and justifying these additional premises. Ideally, if a causal relation is identified by an inductive inference under these premises, this warrants predictions with respect to future instantiations of the same boundary conditions. Of course, an amount of uncertainty will always remain. However, the hope is that for some inductive methods, a measure for the quality of inductive evidence and thus of inductive inferences follows from the additional premises.The problem of enumerative induction was treated extensively by Hume[10] and is essentially unsolvable, since it requires a general *principle of the uniformity of nature* which obviously does not hold. General laws could be derived from observed regularities by assuming that these will continue to be valid. However, there obviously are many regularities that eventually cease to hold which raises the problem how to distinguish between reliable and unreliable generalizations (cp. Goodman 1983). If at all, this issue can only be resolved by keeping close track of the boundary conditions, i.e. by eliminative induction.

However, there is no reason to despair since the problem of eliminative induction is completely distinct from the problem of enumerative induction. For the method of difference, we now list three additional premises which will all be shown to be necessary conditions plus a further condition that is required for predictions: i) *There must be a complete cause for the phenomenon P, i.e. a condition that fully determines the occurrence of P (principle of causality).*[11] After all, if P were due to chance, then the method of difference would fail in the following way: Assume that two instances are observed which differ only in the presence of CX and in the presence of P. By the method of difference, we must then conclude that CX is causally relevant for P, while they are in fact totally unrelated and the occurrence of P actually due to pure chance.

ii) *The background B may contain only conditions that either are irrelevant and allowed to change or are potentially relevant and held constant (excluding of course those conditions that act directly on at least one of the C).* Again, otherwise the method of difference would lead to wrong conclusions. Assume a condition BX in the background which is a necessary condition for P and changes from one instance to the other along with P and the condition CX. The method of difference would again tell us that CX is causally relevant for P, while CX may in fact be totally unrelated with P and the absence of P really due to the absence of BX. A stronger version of this requirement which can frequently be found in the literature is that the circumstances C1, …, CN comprise all potentially relevant boundary conditions—which however is usually an unrealistic assumption. There will always be causally relevant

---

[10] On closer inspection, it is not difficult to show that in his treatment of induction, Hume almost exclusively had enumerative induction in mind. This shortcoming was already stressed by John Maynard Keynes in his *Treatise on Probability*: 'by emphasizing the number of instances Hume obscured the real object of the method. […] The *variety* of the circumstances […] rather than the number of them, is what seems to impress our reasonable faculties.' (1921, 233-234) Russo (2007) endorses a similar anti-Humean viewpoint emphasizing variation over regularity.

[11] A difficulty arises from the definitions of causal relevance and irrelevance as given in Sections 3b and 3c. In cases of overdetermination, a phenomenon may cease to have a cause on some level of description even though determinism holds in principle (cp. example in 3c). However, this does not constitute a problem, since in those cases the phenomenon is fixed, no chance can interfere, and thus cases of overdetermination are not subject to counterexamples of the method of difference as that given above. More generally, not all phenomena need to be caused, in particular *phenomena that cannot be otherwise, e.g. by definition, do not have causes* since the method of difference cannot be applied.

conditions in the background: e.g. stones only fall to earth, as long as there is no superplanet around the corner.

The requirements i) and ii) are of course familiar from the history of philosophy and scientific method. Due to lack of space we must abstain from a detailed examination of the many variants of these principles which have been proposed. The first i) has been discussed under a number of different names, for example as principle of causality, law of causality, or principle of sufficient reason. Intimately connected with the second requirement ii) are various locality constraints as used in physics and often evoked in discussions on causality. Such locality constraints restrict the potentially relevant circumstances to those that are in the spatio-temporal neighborhood of the phenomenon. Baumgartner and Graßhoff (2004, Ch. IX.2.4) present a version of the second requirement which is perhaps closest to the formulation given above. Keynes' principle of limited independent variety[12] (1921, Ch. 22) is intimately related to the second requirement, but in addition involves elements of premise iv) to be discussed below. In what is perhaps the clearest analysis of the problem of eliminative induction in the literature, von Wright terms the first requirement the deterministic postulate and some variant of the second, involving again elements of premise iv), the selection postulate (1951, Ch. V).

There are further premises that are less frequently discussed in the context of eliminative induction. iii) A requirement whose significance can hardly be overrated is that *the conditions C1, …, CN as well as the phenomenon P must be formulated in the correct language accurately capturing the causally relevant categories for the circumstances and phenomena.* This issue links up closely with Nelson Goodman's celebrated new riddle of induction or grue paradox. Suppose a phenomenon P is caused by the color green CX. Now from one instance to the other, the color is changed from green to blue and the phenomenon disappears. Using the correct language of green and blue, i.e. the correct reference classes for the colors, we can easily identify green as causally relevant for P by employing the method of difference. However, using grue we are misled and will conclude that either the presence of P is due to chance or identify a wrong cause if another unrelated circumstance changes simultaneously. Similar examples can of course be constructed if an inadequate language is employed for P or the other C1, …, CN. Note that in these examples the premises i) and ii) can both be satisfied, which shows that we are dealing with a separate issue.

When an eliminative method yields an unstable result, there are always various options. Instead of searching for boundary conditions that have not been taken into account, one can also reconsider the classification scheme, e.g. switch from grue to green in the example above. Eliminative induction is as much about determining a causally adequate ontology or classification system as it is about finding causal laws relating various facts. These are two sides of the same coin. Thus, eliminative induction is also a method to determine 'natural kinds' under the limitation that different classification schemes might be possible.

While iii) highlights a number of issues that make the process of eliminative induction extremely challenging, this must be considered a strength rather than a weakness of the

---

[12] 'the objects in the field, over which our generalisations extend, do not have an infinite number of independent qualities; that, in other words, their characteristics, however numerous, cohere together in groups of invariable connection, which are finite in number' (1921, 256)

proposed account. After all, the history of science is full of examples showing how questions of language are inextricably entwined with induction, not only in those sciences that deal explicitly with ‘species’ like chemistry or biology. Even if often looked down upon, classification belongs to the core of the scientific enterprise and generally ‘we cannot improve the language of any science, without at the same time improving the science itself; neither can we, on the other hand, improve a science, without improving the language or nomenclature which belongs to it’, as Lavoisier famously stated in the introduction to his antiphlogistic chemistry (1789/1965, xv).

iv) While the above premises are sufficient to identify causal relevance according to the method of difference, they do not guarantee that this knowledge can be used for inferences regarding the future. *If we are aiming at predictions, then the total cause should not be so complex as to exclude a reoccurrence of the relevant circumstances*. More specifically, the background B should again be available at future occasions (while of course irrelevant factors in B may have changed) and some set of sufficient conditions for P with respect to B should be reproducible and to some extent controllable. Then, the causal knowledge acquired in the process of eliminative induction can be used to reproduce, influence or merely predict the occurrence of P.

We have thus identified the crucial premises for the method of difference. Remarkably, the premises for the strict method of agreement are essentially the same. Maybe, this is not so surprising since both methods are complementary: after all, if just one circumstance is modified, either the phenomenon changes or not. The principle of causality must be assumed for the strict method of agreement because otherwise it might identify a circumstance CX as irrelevant that is in fact probabilistically relevant. Equally, counterexamples can be constructed if the background does not satisfy condition ii) and if the language is inadequate iii).

3g. Interventions and counterfactuals

The eliminative method presented here can provide some insights into concepts that play an important role in classic accounts of causality, in particular interventions and counterfactuals.

Regarding interventions, the method of difference and the strict method of agreement are silent on the origin of the evidence. Notably, they do not require that the situations of difference and indifference making are produced artificially. Thus, a separate ontology of interventions or manipulations as qualitatively distinct from mere observations is not necessary. Rather, pure observations can sometimes yield as conclusive and reliable causal knowledge as experiments. However, experimental knowledge is mostly superior since one can explicitly set up and control the boundary conditions in a way required for eliminative induction and does not have to depend on luck or chance.

Concerning counterfactual statements, eliminative induction provides the means to evaluate their truth by referring to actual situations in the real world. This, of course, is the basic idea behind the method of difference where the two situations instantiate as closely as possible a counterfactual with respect to each other. However, the counterfactual is never entirely realized since from one situation to the other always a number of boundary conditions change.

This objection can be somewhat circumvented with help of the strict method of agreement which allows to identify irrelevant boundary conditions. Thus, the counterfactual is realized in the world if it differs only in terms of irrelevant boundary conditions. Of course, strictly speaking, we will never get to the point, where we can be completely certain that all boundary conditions that changed were indeed irrelevant. Still, one can always further approximate such knowledge. No possible-world semantics is required to evaluate the truth of counterfactuals.

Thus, an account of causality based on eliminative induction differs from counterfactual accounts in that difference-making is evaluated not with respect to possible worlds but referring to actual situations in our world and it differs from interventionist and manipulationist accounts in that these situations do not have to result from interventions.

3h. Eliminative induction in probabilistic contexts

Thus far, we have presented eliminative induction as a method to determine causal relevance and irrelevance in deterministic contexts, which is justified mainly because the method presupposes the principle of causality (cp. Section 3f). But eliminative induction is in fact not restricted to deterministic settings. Rather, it is often possible to reformulate probabilistic contexts in a way that the problem becomes deterministic and then eliminative induction can be applied.

An excellent example concerns randomized controlled trials, which supposedly constitute the gold standard of the recent wave of evidence-based sciences. In such studies, it is usually not possible to determine causal connections on the level of individuals because the situations are too complex to directly apply eliminative induction. However, on the level of populations the problem can often be reformulated in a manner suitable for eliminative induction. As an example, consider the question if there is a causal relation between taking aspirin and curing headaches. On the level of individuals, the question cannot be answered by the method of difference, just because it is impossible to find two persons that differ only in terms of having taken an aspirin. However, given two large populations, to which subjects are allocated randomly, we can hope that the differences between individuals cancel out on the population level and that the two populations are on average similar with respect to all relevant boundary conditions. If then, one of these groups is told to take aspirin and the other must refrain from doing so, the method of difference and the strict method of agreement become applicable. The idea that randomization guards against ‘corruption by the causes of disturbances which have not been eliminated’, i.e. against unwanted variations in the boundary conditions, is already present in Ronald Fisher’s writings (2003, 19).

Another example, where a probabilistic context can be reframed in a deterministic manner, concerns quantum mechanics. Of course, this is the standard example of an allegedly indeterministic theory, in which eliminative induction might fail. However, the method turns out just as useful for experiments in microphysics as in deterministic macrophysics. After all, in quantum mechanics the deterministic part, which concerns the development of the wave function according to the Schrödinger Equation, separates nicely from the indeterministic part regarding the collapse of the wave function upon measurement. Therefore, all experiments that only concern the deterministic part can be treated by means of eliminative induction, i.e.

as long as one stays on the ensemble level of the probability distribution and does not consider individual quantum events. Consider for example double slit experiments involving a large number of particles. In those experiments, one does not observe single, indeterministic quantum events, but rather the full probability distribution resulting from the wave function which can be calculated deterministically from the boundary conditions by means of the Schrödinger Equation.

Much more needs to be said about the connection between eliminative induction and probability, but this must be left to another occasion.

4. Evidence as difference and indifference making

Let us briefly summarize what kind of evidence is required to establish causal relations by means of eliminative induction. For this, we just have to look at the fundamental methods, i.e. the strict method of agreement and the method of difference, and examine what kind of empirical input they require. [13] The answer is straight-forward: pairs of observations that differ in only one C of all circumstances that are potentially relevant to a phenomenon P. Regarding the phenomenon, two cases should be distinguished corresponding to the two methods: (i) First, the change in C has no effect on P, which constitutes evidence for the causal irrelevance of C to P. (ii) Second, the change in C prevents P from happening, which constitutes evidence for causal relevance of C to P. Under certain additional premises, which were pointed out in Section 3f, such evidence can be conclusive.

Thus, one is looking for situations of difference or indifference making. A generalization to several potentially relevant factors is straight-forward. Essentially, in order to establish complete causal knowledge, all possible combinations have to be tried. In case of parameters that can be changed in continuous degrees, such complete knowledge is of course not possible, but it can be reasonably approximated.

In principle, causal evidence of the type required for eliminative induction is also a three-place relation involving a background B, the potentially relevant parameters C and the phenomenon P. Notably, a systematic account of a specific evidence situation must mention (i) the state of causally relevant conditions in the background or at least a general description of the background or context, (ii) the parameters C which are changed or held constant and (iii) the respective impact on the phenomenon P. This rationale is followed quite closely in typical experimental reports[14], which start out with a sketch of the experimental set-up usually including a description of the scientific instruments that were used. Against this causal background, the phenomenon P is produced and as many potentially relevant boundary conditions C as possible are identified. Then the process of eliminative induction is carried

---

[13] Emphasizing the nature of evidence and scientific method when discussing causation is much in the spirit of the epistemic approach advocated by Williamson (2005) and Russo & Williamson (2011), who argue that an understanding of causality can only result from a thorough examination of causal epistemology.

[14] For examples I recommend a Google search on ‘how to write a lab report’ resulting in e.g. http://schools.cbe.ab.ca/b631/Science%20Web%20Page/Grade%209%20Science%20Web%20Page_files/9%20How%20To%20Write%20Up%20A%20Lab%20Report.pdf, accessed 29.8.2013.

out in that as many combinations of the parameters C as deemed reasonable are tried and the respective impact on P is recorded.

An important issue concerns the question whether evidence in terms of difference and indifference making is the only kind of evidence for causal relations or if other types, in particular mechanistic evidence, should be taken into account as for example Russo and Williamson have argued (e.g. 2007, 2011; see also Illari 2011). In this article, I do not want to take a stance on this issue. However, my intuition is that mechanistic evidence should in principle be accountable for in terms of difference making. After all, mechanisms often refer to physical processes and I cannot see how the relevant physical knowledge is established by other means than exploratory experimentation relying on difference making. In this view, the additional confirmatory strength of mechanistic evidence would result from linking the examined causal relation via the analysis in terms of more fine-grained laws with a larger body of evidence that was used to establish these fine-grained laws (cp. the discussion in Glennan 2011, Sec. 5).

In the following, we will establish that experimental science often proceeds on the basis of eliminative induction and consequently that in such instances evidence of the mentioned type is used.

5. Eliminative induction in experimental science and engineering

5a. Exploratory experimentation

Francis Bacon arguably made two major contributions to science: first, he played a decisive role in establishing a tradition of scientific experimentation which eventually led to the rise of modern science; second, Bacon exposed the weaknesses of the Aristotelian enumerative induction and replaced it by his more powerful inductive *method of exclusion* relying on a collection of instances in the three kinds of tables (Bacon 1620/1994). The two issues are closely related. I will now argue that eliminative induction provides the logic behind a large number of scientific experiments, in particular exploratory experiments interested in the causal structure of novel phenomena. Certainly, other types of experiments exist, which do not follow the logic of eliminative induction so closely, for example crucial experiments that aim to decide between competing theories like Arthur Eddington's solar eclipse experiment which confirmed a prediction of Einstein's theory of general relativity but was in disagreement with standard accounts of Newtonian mechanics or the Foucault-Fizeau experiment allegedly deciding in favor of the wave and against the particle nature of light. Notably, these experiments are not *exploratory*, they are *theory-driven*, testing the coherence of a specific theoretical view.[15]

---

[15] The distinction and terminology is Steinle's (1997; 2005, esp. Ch. 7), it is also used by Burian (1997). For a more recent discussion confer Waters (2007) and references therein. The basic idea of exploratory experimentation is also described in Vincenti (1993) from an engineering perspective. With increasing use of information technologies in science, novel research paradigms emerge, in particular data-driven approaches. For an attempt to relate these to an account of eliminative induction, see my draft 'Big Data – The new science of complexity' (http://philsci-archive.pitt.edu/9944/).

A number of features distinguish exploratory from theory-driven experimentation: (1) The former, but not the latter involves extensive parameter variation. (2) There is little theoretical knowledge about the examined phenomenon in exploratory experimentation, while in theory-driven experimentation a specific question is asked which is largely framed by the theoretical context, for example concerning the truth of a specific hypothesis, e.g. if light deviates in a gravitational field. Of course, exploratory experimentation is not theory-free. For example, the scientific instruments presuppose a lot of theory. More importantly, the question which parameters are considered potentially relevant is largely determined by theoretical intuitions. (3) A good way to understand the distinction is that it mirrors the different roles of experiments in a hypothetico-deductive versus an inductive approach.

A good example of exploratory experimentation regards Wilhelm Röntgen’s discovery of the X-rays (1895). At the beginning stands the detection of an unexpected effect that he then systematically examines by means of parameter variation. For example, he determines the extent to which the novel radiation passes through various bodies depending on width; the fluorescence caused by X-rays in various materials including the sensibility of photographic plates; the heat generated by X-rays; refraction, e.g. in prisms, and reflection on different types of surfaces; the decrease of intensity in vacuum depending on distance; the deviation in electric and magnetic fields; etc.

Obviously, we are witnessing simple eliminative induction. First, Röntgen works in a laboratory environment with complex experimental set-ups which together provide the stable background B. In particular, the background includes boundary conditions under which the novel phenomenon can be reliably produced. He then identifies potentially relevant parameters, guided by intuition and his knowledge about other types of radiation. The rest is just parameter variation, i.e. he tries to map the parameter space to an extent that he deems necessary varying one of the circumstances at a time. For example, by the strict method of agreement, he finds that electric and magnetic fields have no influence on the novel kind of radiation. By the methods of difference and of concomitant variations he finds the remarkable absorption properties of X-rays.

From a systematic perspective, exploratory experiments that try to map the causal structure of novel phenomena have a number of characteristics in common: (i) they happen under laboratory conditions, i.e. in an extremely well-known and well-controlled environment; (ii) experimenters always have some intuitions about which parameters are potentially relevant for the phenomenon in question; (iii) the use of scientific instruments ensures the controlled production of the phenomenon as well as the systematic variation of potentially relevant boundary conditions. Largely, there are two kinds of instruments, one concerned with the production of phenomena, the other with analysis / measurement, although of course these functions are often intertwined; (iv) exploratory experiments follow an analytic rationale, i.e. they try to isolate the phenomenon in question as much as possible from other interfering phenomena and to subsequently dissect it in as many independent sub-phenomena as possible; (v) ideally, only one boundary condition, one basic parameter, is changed at a time; (vi) during the experiment, the scientist tries to map the parameter space as far as it is thought to be relevant, i.e. to examine all possible combinations of parameters (which of course is often not strictly feasible, especially when dealing with continuous variables); (vii) an error analysis is carried out which generally includes the repeated realization of all measurements, usually for several times in a row; (viii) scientific experiments must be reproducible, i.e. yield the same results in other laboratories.

Let me briefly argue that all these characteristics of exploratory experiments ensure that eliminative induction is possible and yields good results. (i) and (ii) are supposed to guarantee the constancy of the background B as one of the crucial requirements for eliminative induction (cp. Section 3f). If eliminative induction is to yield reliable results then all conditions that are potentially relevant must be held constant except those in whose influence one is explicitly interested. That is why scientific experiments usually happen in closed and highly controlled environments that are shielded off as much as possible from external influences. (iii) In such an ideal environment, the phenomenon in question is produced and analyzed, usually by means of specialized scientific instruments. These usually rely on well-established technologies that in turn exploit reliable causal knowledge in connection with the phenomenon in question. At the intersection of the knowledge encoded in the scientific instruments, the new territory of the yet unexplored phenomenon arises.

Scientific instruments aim to realize items (iv) to (vi). They try to partition the phenomenon in suitable subphenomena which ideally are mutually independent. These suggest a choice of potentially relevant boundary conditions C1, …, CN. In the best case, the instruments allow to change the conditions independently to record the respective impact on the phenomenon, i.e. they allow for an application of the method of difference, the strict method of agreement or their continuous counterparts.

Steps (vii) and (viii) assess the quality of the inductive inference. In error analysis, one checks for both systematic and accidental errors, i.e. for relevant boundary conditions in the background that cannot be held constant but were thus far not accounted for. Regarding the former, the influence of parameters is estimated that are known to be relevant and that can be theoretically approximated but which are beyond experimental control. By contrast, the influence of accidental errors is usually estimated experimentally, by a repeated realization of the same measurement. While this may resemble enumerative induction, it can only be understood from the perspective of eliminative induction. These repeated measurements allow assessing the extent to which the background really remains constant or whether there are accidental fluctuations of relevant boundary conditions which are not yet accounted for. (viii) Finally, the requirement that scientific experiments should be reproducible also tests the quality of the causal knowledge gained by an experiment. In particular, it ensures that we have sufficient knowledge of the relevant boundary conditions and are capable to reproduce the background in order to produce the phenomenon in a different setting, e.g. in another laboratory.

In summary, conditions (i) to (iii) assure that the basic set-up required for eliminative induction is realized, in particular that the background remains constant and that one has good knowledge and control of potentially relevant boundary conditions with respect to this background. Conditions (iv) to (vi) denote the actual application of eliminative induction. And finally steps (vii) and (viii) help to assess the quality of the eliminative process.

5b. Design Methods

Francis Bacon famously made a link between scientific knowledge and man's power over nature: 'Human knowledge and human power meet in one; for where the cause is not known the effect cannot be produced.' (1620/1994, I.iii) As the quote indicates, the link is provided by causality, which Bacon establishes with his method of exclusion, i.e. essentially eliminative induction. Therefore, causal knowledge not only looms large in the experimental analysis of nature as described in the previous subsection but also with respect to the creation of objects and phenomena by design, i.e. causal knowledge is as important in experimental science as in engineering. In the following, a number of design methods from the engineering sciences will be analyzed to show that they heavily rely on eliminative induction and therefore on evidence of the type difference / indifference making.

In spite of the high hopes once connected with Ian Hacking's call for a 'new experimentalism' (1983), there is still comparatively little literature on the philosophical analysis of scientific experimentation (Radder 2003, 1). Even less exists on the methodology of the engineering sciences and on engineering design in particular. Still, a few historians of technology and practicing engineers have written on these subjects with the foundational and systematic interest required for good analytic philosophy (e.g. Vincenti 1993, Arthur 2009; compare also the enormous collected volume edited by Anthonie Meijers 2009). From this pioneering work I will largely draw in the following in order to connect several basic design methods with eliminative induction.

A number of authors have pointed out remarkable analogies between biological and technological evolution (e.g. Vincenti 1993, Arthur 2009; cp. also Dawkins 2006, Ch. 11). The economist and complexity scientist Brian Arthur distinguishes two modes of technological evolution (2009, Ch. 1). (i) The first is largely analogous to biological evolution: a functioning artifact is copied introducing slight variations and in case these result in improvement, the new design survives and older versions disappear. This mode leads to a gradual optimization of technologies. (ii) The second mode is prevalent in technological development but is rare in the biological evolution of large-scale organisms, while more relevant for simpler organisms like bacteria or viruses. This second mode consists in recombinations of large chunks of genes in biology or of lower-level technologies in engineering. This mode, which is arguably the crucial evolutionary mode when it comes to technological invention, makes use of two fundamental characteristics of technology: modularity and recursiveness.

Let me now discuss the two modes in turn and point out how they closely mirror the characteristics of eliminative induction.

(i) Like Darwinian evolution, the first mode of technological evolution relies on the introduction of small variations into a functioning artifact and a subsequent selection of those variations which lead to significant improvement within a specific environment. This mode results in the continuous optimization of technologies relative to their contexts of application. Several historians and philosophers of engineering and technology have extensively discussed this method (Vincenti 1993, Chs. 5 & 9; Arthur 2009, Chs. 1 & 9).

Walter Vincenti speaks of *parameter variation*, which supposedly is ‘so much taken for granted by engineers that they rarely call it by name’ (1993, 291). Parameter variation, in Vincenti’s account, is mostly identical with exploratory experimentation involving the same steps discussed in the previous section. It constitutes a method to gain causal knowledge of a phenomenon in absence of higher-level theory. The main difference lies in aim. While exploratory experiments are carried out to acquire further knowledge about the respective phenomenon, parameter variation in engineering design aims at the improvement of a technology through causal knowledge.

According to Vincenti, the method has a venerable tradition in the history of engineering: it ‘can be defined as the procedure of repeatedly determining the performance of some material, process, or device while systematically varying the parameters that define the object of interest or its conditions of operation. Apart from possible earlier use in science, the method goes back in engineering at least to the ancient Greek catapult designers, who established the best proportions for their full-sized devices “by systematically altering the sizes of the various parts of the catapult and testing the results”.’ (139)

Vincenti (1993, Ch. 5) describes in considerable detail an example from aeronautics, the well-documented experiments on aircraft propellers conducted by William Durand and Everett Lesley at Stanford University from 1916 to 1926. In these, propeller design was improved by trying ‘all possible combinations of values’ (148) of those parameters that were considered relevant to performance, i.e. essentially by applying eliminative induction. The variables were separated into two groups: the conditions of operation and the geometric properties of the propeller. The former included the forward speed of the propeller with respect to the air stream and the revolutions per unit time of the propeller; as relevant geometric properties Durand and Lesley identified the diameter of the propeller as well as five ratios in the geometry of the propeller including the mean pitch ratio, which measured the angular orientation of the blade section relative to the plane of propeller rotation, and the type of blade section, e.g. a straight or a curved blade outline. These parameters correspond to the conditions C1, …, CN of Section 3a.

Other parameters, that were known to be relevant, were held constant, including for example the number of blades of the propeller and the relative orientation of the blades. These constitute part of the background B of the experiment. Still further conditions are beyond experimental control even though they are known to be relevant including the viscosity and compressibility of the air or the elastic bending of the rotor blades. But these factors were estimated to have a negligible effect on performance and can thus also be considered as forming part of B.

In the next step, Durand and Lesley built a wind tunnel, in which different propellers could be tested under laboratory conditions that guaranteed the constancy of the background and an adequate control of the parameters C1, …, CN. They produced forty-eight propellers that covered the relevant range of the five geometric ratios and subsequently determined the efficiency of each propeller in the wind tunnel depending on forward speed and revolutions per unit time. In other words, they put the method of difference, the strict method of agreement, and the method of concomitant variations to work. Obviously, parameter variation is just exploratory experimentation with the goal of optimizing a specific artifact for a certain purpose.[16]

An extreme case of technological optimization by eliminative induction is the study of technological failure, when a technical artifact ceases to function as expected in a specific context of application. Failure analysis is a huge industry in the engineering literature and there exist a large number of analytical methods to find out the cause of a failure, e.g. change analysis, fault tree analysis, or MORT analysis. The details are irrelevant here. Suffice it to remark that all of these methods at least implicitly interpret the cause of failure in terms of difference making. The basic question in failure analysis is: which boundary condition(s) need to be changed in order to make the failure disappear resulting in an artifact, which functions properly in the context of the previous failure? Failure analysis thus is also an optimization method that aims to extend the domain of application of a technology.

That failure analysis relies on basic eliminative induction, can be seen in the following description by Louis Bucciarelli, another of the few engineers who have dared to venture onto the grounds of philosophy: ‘When things go wrong, when your product or system misbehaves and surprises you or, more ambiguously only suggests that something is out of order, the first task is to try to replicate the failure, to establish conditions such that, when we set the system in motion, the faulty behavior re-appears. In this way we can construct a fuller description of the problem, stating under what conditions and with what settings the fault occurs. A next step, if we are successful in replicating the malfunction—and this with some consistency—is to change conditions in some way and observe the result. We seek to make a relevant difference in conditions—relevant in that it alters the state of the product in some significant way—and then note if this alteration does, or does not, eliminate the failure. Our traditional strategy recommends that we change but one condition at a time, proceeding in this way until a cause of the failure is identified, i.e., the system stands corrected and now runs as it should.’ (Bucciarelli 2003, 25) Needless to say, that all kinds of practical complications arise in real-world settings. For example, often the failed artifact is unique and in order to reproduce the failure, one has to work with appropriate models. In summary, failure analysis is about using eliminative induction to map the boundary at which a phenomenon can still be produced.

A common characteristic of all methods of optimization is that they start with a largely functioning artifact as for instance in the propeller study. This is not a coincidence but rather results from the fact that most causal knowledge is contextual in the sense explained in Section 3a. Most importantly, we rarely have complete causal knowledge of all relevant boundary conditions. Thus, the causal knowledge of the technologies around us is almost always incomplete. Their functioning is rarely understood from first principles.

In the lack of complete understanding, we just copy what has been known to work and hope that it will continue to function in similar contexts. This copying without full understanding is an essential part of the human condition. From our earliest day, we copy from others, often unconsciously, from parents, friends, teachers etc. and gradually optimize when things go

---

[16] Additional difficulties arise, when the causal knowledge gained with the model is to be applied to the real object under operating conditions, i.e. to a propeller of a plane in flight. Engineers, much more than natural scientists, have to take care that the background is stable enough to allow for practical application.

wrong. When we copy in crafts or in technology, the causal background is usually taken for granted or considered the responsibility of others and not further analyzed. We can thus make use of causal knowledge that has been lost or was never known in the first place, but is carried on implicitly in functioning artifacts and procedures. In this way, artifacts and actions are important carriers of causal knowledge beyond the textual knowledge usually associated with science.

The idea that in the absence of higher-level knowledge, improvement is achieved by introducing small variations into functioning artifacts or organisms, is common both to technological and biological evolution. As pointed out, the reason lies in the structure of causal knowledge. Optimization by evolution essentially relies on eliminative induction and the corresponding type of evidence in terms of difference / indifference making.

(ii) The second mode of technological development is prevalent when it comes to genuine invention of artifacts that are to serve novel means or old means in novel ways. This mode of technological development is described in much detail in Brian Arthur's book 'The Nature of Technology' (2009, Ch. 2 & 9). The essential characteristics of technology which enable this second mode are modularity and recursiveness. These ensure that existing technologies can be regrouped and rearranged in an almost infinite number of ways. It is therefore apt to compare, as Arthur does, technology with a language of sentences, words, syllables and letters that can be recombined in endless ways.

When technologies are taken apart in suitable ways we end up with building blocks that can themselves be considered technologies. Modularity means that such subtechnologies generally function independently of the context into which they are embedded. For example, a laser works according to the same principles, whether it is built into a printer or into a CD player. [17] Recursiveness denotes that technologies can be considered at a hierarchy of several levels of resolution, i.e. technologies can be split into subtechnologies, subtechnologies into subsubtechnologies and so on. As an example, lasers are constructed from diverse technical elements, e.g. a light source, mirrors, an energy supply etc., that are technologies themselves and are often in turn composed of several subtechnologies.

The principles of modularity and recursiveness are well mirrored in the structure of eliminative induction and of causality in general. Boundary conditions can be combined into groups at various levels of resolution or coarse-graining, and on each level of description the method of difference and the strict method of agreement can be applied. Modularity is essential for eliminative induction, since it ensures that the dependence relations among the relevant parameters are not so complex as to render a causal analysis impossible.

Modularity should be understood as a research objective rather than a fact about technological systems. It ensures that a causal mechanism works with respect to a large variety of different backgrounds, in a large number of different contexts. Modularity has both an objective and a linguistic dimension, in that causal relations in nature often seem to be modular as they do not

---

[17] A certain notion of modularity has recently come under attack (e.g. in Cartwright 2009). However, this debate does not concern Arthur's basic observation that modularity, broadly understood, constitutes a pragmatically useful feature for technological invention.

depend on many background conditions, e.g. the rock drops to earth and does not much care about what happens around it. But this modular structure of the world can only become apparent, if the phenomena are described in an adequate language.

The mentioned analogy between language and technology thus has deeper structural reasons, and the link is provided by causality. Modularity depends crucially on the classification of phenomena, on the development of a suitable language that adequately mirrors the causal structure of the world—as also discussed in Section 3f in connection with Goodman's problem of grue. It is therefore not surprising that the recursive and modular structure of causal relations is mirrored in a recursive and modular structure of our language.

In summary, both modes of engineering design, optimization and recombination, rely heavily on eliminative induction. Thus, a considerable portion of man's power over nature, as manifest in experimental science and in engineering, is due to this type of induction and the corresponding evidence of difference and indifference making.

6. Conclusion

An account of a specific type of evidence was presented that is prevalent to establish causal relations in deterministic contexts, namely difference and indifference making with respect to a background of boundary conditions. This type of evidence was then linked to a particular scientific method, the eliminative induction in the tradition of Bacon and Mill. I tried to sketch the most important ingredients of an account that is both defensible and relevant to scientific practice. The central element remains the method of difference establishing causal relevance. As further characteristics, all causal statements are made context-relative and a method to establish causal irrelevance is introduced, referred to as strict method of agreement. Such eliminative induction and the corresponding type of evidence were shown to be widely used in science, in particular in exploratory experimentation and methods of artifact design.

Acknowledgments

I am much grateful for helpful comments from two anonymous referees as well as the editors Phyllis Illari and Federica Russo.

References

Anderson, John. 1938. "The Problem of Causality." *Australasian Journal of Psychology and Philosophy* 16(2): 127-142.

Aristotle. 1928. *Posterior Analytics.* Transl. Geoffrey R.J. Mure. In William D. Ross (ed.). *The Works of Aristotle. Vol. 1*. Oxford: Oxford University Press.

Arthur, Brian W. 2009. *The Nature of Technology*. New York, NY: Free Press.

Bacon, Francis. 1620/1994. *Novum Organum*. Chicago, Il: Open Court.

Baumgartner, Michael & Gerd Graßhoff. 2004. *Kausalität und kausales Schließen*. Norderstedt: Books on Demand.

Bucciarelli, Louis L. 2003. *Engineering Philosophy*. Delft: Delft University Press.

Burian, Richard. 1997. "Exploratory Experimentation and the Role of Histochemical Techniques in the Work of Jean Brachet, 1938–1952." *History and Philosophy of the Life Sciences* 19:27–45.

Cartwright, Nancy. 2007. *Hunting Causes and Using Them.* Cambridge: Cambridge University Press.

Dawkins, Richard. 2006. *The Selfish Gene.* Oxford: Oxford University Press.

Fisher, Ronald A. 2003. *The Design of Experiments.* In *Statistical Methods, Experimental Design, and Scientific Inference.* Oxford: Oxford University Press.

Franklin, Allan. 2012. "Experiment in Physics." *The Stanford Encyclopedia of Philosophy (Spring 2012 Edition)*, http://plato.stanford.edu/archives/spr2012/entries/physics-experiment/

Glennan, Stuart. 2011. "Singular and General Causal Relations: A Mechanist Perspective." In Phyllis M. Illari, Federica Russo & Jon Williamson (eds.). *Causality in the Sciences* (789-817). Oxford: Oxford University Press.

Goodman, Nelson. 1983. *Fact, Fiction, and Forecast.* Cambridge, Ma: Harvard University Press.

Hacking, Ian. 1983. *Representing and Intervening.* Cambridge. Cambridge University Press.

Herschel, John F. W. 1851. *Preliminary Discourse on the Study of Natural Philosophy*. London: Longman, Brown, Green, and Longmans.

Howick, Jeremy. 2011. *The Philosophy of Evidence-Based Medicine.* Chichester: Wiley-Blackwell.

Hume, David. 1739/2000. *A Treatise of Human Nature*. Oxford: Oxford University Press.

Hume, David. 1748/2007. *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

Illari, Phyllis M. 2011. "Mechanistic Evidence: Disambiguating the Russo-Williamson Thesis." *International Studies in the Philosophy of Science* 25(2):139-157.

Keynes, John M. 1921. *A Treatise on Probability*. London: Macmillan.

Lavoisier, Antoine. 1789/1965. *Elements of Chemistry*. New York, NY: Dover.

Mackie, John L. 1965. "Causes and Conditions." *American Philosophical Quarterly* 12:245-265.

Mackie, John L. 1980. *The Cement of the Universe.* Oxford: Oxford University Press.

Meijers, Anthonie (ed.). 2009. *Philosophy of Technology and Engineering Sciences. Handbook of the Philosophy of Science, Vol. 9.* Amsterdam: Elsevier.

Mill, John S. 1886. *System of Logic*. London: Longmans, Green & Co.

Norton, John. 2010. "A Survey of Inductive Generalization". http://www.pitt.edu/~jdnorton/papers/Survey_ind_gen.pdf (Accessed 02/28/2012).

Pearl, Judea. 2000. *Causality. Models, Reasoning, and Inference.* Cambridge: Cambridge University Press.

Pruss, Alexander R. 2006. *The Principle of Sufficient Reason*. Cambridge: Cambridge University Press.

Radder, Hans (ed.). 2003. The Philosophy of Scientific Experimentation. Pittsburgh, PA: Pittsburgh University Press.

Röntgen, Wilhelm K. 1895. Eine neue Art von Strahlen. Würzburg: Stahel'sche K. Hof- u. Universitäts-Buchhandlung.

Russo, Federica. 2007. “The Rationale of Variation in Methodological and Evidential Pluralism”. *Philosophica* 77:97-124.

Russo, Federica & Jon Williamson. 2007. “Interpreting Causality in the Health Sciences.” *International Studies in the Philosophy of Science* 21:157-170.

Russo, Federica & Jon Williamson. 2011. “Epistemic Causality and Evidence-Based Medicine.” *History and Philosophy of the Life Sciences* 33(4):563-581.

Spirtes, Peter, Clark Glymour & Richard Scheines. (2000). *Causation, Prediction and Search*. Cambridge, MA: M.I.T. Press.

Sprenger, Jan. 2011. “Hypothetico-deductive confirmation.” *Philosophy Compass* 6:497-508.

Skyrms, Brian. 2000. *Choice and Chance*. Belmont, CA: Wadsworth.

Steinle, Friedrich. 1997. “Entering New Fields: Exploratory Uses of Experimentation.” *Philosophy of Science* 64:S65–S74.

Steinle, Friedrich. 2005. *Explorative Experimente.* Stuttgart: Franz Steiner Verlag.

Vincenti, Walter. 1993. *What Engineers Know and How They Know It.* Baltimore: Johns Hopkins University Press.

Waters, C. Kenneth. 2007. “The Nature and Context of Exploratory Experimentation.” *History and Philosophy of the Life Sciences* 29(3): 275-284.

Weber, Marcel. 2012. “Experiment in Biology.” *The Stanford Encyclopedia of Philosophy (Spring 2012 Edition)*, http://plato.stanford.edu/archives/spr2012/entries/biology-experiment/

Williamson, John. 2005. *Bayesian Nets and Causality: Philosophical and Computational Foundations.* Oxford: Oxford University Press.

von Wright, Georg H. 1951. *A Treatise on Induction and Probability*. New York, NY: Routledge.

von Wright, Georg H. 1957. *The Logical Problem of Induction.* New York, NY: Barnes & Noble.