

Causal explanation beyond the gene: manipulation and causality in epigenetics *

Jan BAEDKE

Received: 15.11.2011

Final version: 29.3.2012

BIBLID [0495-4548 (2012) 27: 74; pp. 153-174]

ABSTRACT: This paper deals with the interrelationship between causal explanation and methodology in a relatively young discipline in biology: epigenetics. Based on cases from molecular and ecological epigenetics, I show that James Woodward's interventionist account of causation captures essential features about how epigeneticists using highly diverse methods, i.e. laboratory experiments and purely observational studies, think about causal explanation. I argue that interventionism thus qualifies as a useful unifying explanatory approach when it comes to cross-methodological research efforts. It can act as a guiding rationale (i) to link causal models in molecular biology with statistical models derived from observational data analysis and (ii) to identify test-criteria for reciprocal transparent studies in different fields of research, which is a shared issue across the sciences.

Keywords: causation; explanation; intervention; epigenetics; methodology; observational studies.

RESUMEN: Este artículo trata de la relación entre explicaciones causales y metodología en una disciplina biológica relativamente joven, la epigenética. Basándome en casos de la epigenética molecular y ecológica, muestro que la concepción intervencionista de la causalidad desarrollada por James Woodward capta algunos rasgos esenciales del modo en que los epigenetistas conciben la explicación causal usando métodos sumamente diversos: e.g., experimentos de laboratorio o estudios observacionales. Defiendo que el intervencionismo es útil como aproximación unificadora a la explicación cuando se trata de empresas investigadoras transdisciplinarias. Puede servir como guía para (i) conectar los modelos causales en biología molecular con los modelos estadísticos derivados del análisis de datos observacionales y (ii) para identificar criterios de prueba para estudios recíprocos en diferentes ámbitos de investigación, un problema de interés común en diferentes ciencias.

Palabras clave: causalidad; explicación; intervención; epigenética; metodología; estudios observacionales.

The right order for experience is first to kindle a light, then with that light to show the way, beginning with experience ordered and arranged, not irregular or erratic, and from that deriving axioms, and from the axioms thus established deriving again new experiments.

—Francis Bacon, 1620

1. *Unifying epigenetics and evolutionary biology*

The past few decades have seen an important expansion of our understanding of inheritance and its underlying causal processes as established by genetics and the Mod-

* Thanks to Michael Anacker, Dan Brooks, Fridolin Gross, Maria Kronfeldner, Frank Paris, Jessica Pahl, Helmut Pulte, Kirsten Schmidt, the session audience at the ISHPSSB meeting (2011), the Bielefeld Biolosophy meeting group, and three anonymous reviewers of this journal for constructive comments on earlier versions of this paper. Financial support by the Ruhr University Research School (RURS) and the ISHPSSB is gratefully acknowledged.



ern Synthesis. To cope with this development, the time is right to broaden the spectrum of philosophical analysis on causation and causal explanation in biology, and to focus on disciplines and phenomena responsible for this expansion. Thus epigenetics, a relatively young field of research that challenges gene centrism, takes center stage.

The notion of epigenetics has radically changed throughout the decades. Today epigenetics basically represents a heterogeneous field that focuses on non-genetic inheritance phenomena: “The term ‘epigenetics’ defines all meiotically and mitotically heritable changes in gene expression that are not coded in the DNA sequence itself” (Egger et al. 2004, 457). Epigenetic inheritance thus includes developmentally induced highly stable non-genetic cell heredity (Jablonka and Lamb 1995, 2005) and less stable transmission between generations of organisms (Jablonka and Raz 2009). Heritable epigenetic variation can appear with a degree of autonomy from the DNA-level. Thus epigenetics challenges gene centrism and asks for a broader notion of heredity that should be taken into consideration for inheritance and evolution.

Epigenetic inheritance systems, like chromatin-marking systems (Suzuki and Bird 2008) or RNA-mediated gene silencing (Siomi and Siomi 2009), are highly sensitive to environmental influence. They constitute a second layer of biological information, mainly responsible for editing and realising genetic information, by switching genes on or off. These mechanisms, important for *development*, are involved in *heredity* (information transfer) as well¹. Because of this dual nature, inducible epigenetic inheritance has often been associated with Lamarckian ‘soft inheritance’ or inheritance of acquired characteristics (Gissis and Jablonka 2011; Jablonka and Lamb 1995, 2005, 2008; Richards 2006; Sano 2010).

This claim recently has been accompanied by a common call for an ‘Extended Synthesis’ or even ‘post-Darwinian Synthesis’ which includes epigenetics². Such a theory is based on a ‘development first’-view on evolution, which emphasizes the evolutionary significance of developmental responsiveness. It has to be supported by a new methodological and explanatory framework that revises traditional disciplinary boundaries, especially those between molecular and developmental biology on the one hand and evolutionary biology on the other. This framework has yet to be developed.

Despite the vast amount of new data and epigenetic heredity models coming from manipulative experiments in molecular biology, some critics still claim that the transgenerational transfer of epigenetic information does not have an impact on natural population dynamics and is therefore irrelevant for evolutionary biology. Their argument relies on the fact that, contrary to genetic variation, body-to-body transfer of epigenetic variation shows relatively limited heredity because of unstable states and high mutation rates (Hall 1998; Pál and Hurst 2004; Walsh 1996). To face these critics, additional research efforts stretching out to ecologically relevant circumstances have

¹ In contrast to those molecular biologists investigating the developmental dimension of epigenetic processes (e.g., their role in regulation of DNA-expression), (molecular) epigeneticists often focus on explaining how these processes contribute to (transgenerational) information transfer, i.e. heredity.

² See, e.g., Cabej 2008; Gissis and Jablonka 2011; Jablonka and Lamb 2005; Pigliucci 2009; Pigliucci and Müller 2010.

to be made. Such observational studies are crucial to estimate how stable or invariant transgenerational epigenetic effects are in natural populations.

It is at this point where epigenetics and evolutionary biology face a major explanatory and methodological problem. If epigenetics want to become a discipline involved into evolutionary biology, it has to be able to interrelate molecular biologists' causal explanations with statistical explanations of ecologists (Griesemer 2011; Richards et al. 2010). This problem has to be solved in order to link observational studies in the recently emerging field of ecological and evolutionary epigenetics to causal reasoning in experimental (molecular) epigenetics. In the philosophical literature on causation the molecular biologists' view on causal explanation is commonly treated as a mechanistic account of explanation (Bechtel and Abrahamsen 2005; Darden 2005; Machamer et al. 2000; Tabery 2004). According to such an account, explanation is achieved when one shows how a phenomenon is produced by a mechanism (statistically relevant correlations alone are considered to be insufficient for a causal explanation). In contrast, ecologists (and epidemiologists) develop statistical models to make inferences from observational data (Royle and Dorazio 2008; Shipley 2000); until now many ecological mechanisms are not well known (Raerinne 2011).

I will argue that an essential stepping stone to unite developmental and evolutionary biology in general and epigenetic subdisciplines in particular is to establish a unifying way of causal explanation guiding research on joint phenomena. James Woodward's interventionist account of causation (Woodward 2003) will be presented as a good candidate to solve this translation or unification problem in epigenetics. I will show (i) that it provides an explanatory framework essential to link explanations in ecological observational data analysis to manipulation-based causal explanations in molecular biology and (ii) that it has the capacity to identify test-criteria and research strategies for reciprocal transparent studies. Thus interventionist explanation can act as a rationale guiding causal reasoning in different fields of research. In molecular and ecological epigenetics it provides a crucial heuristic tool to coordinate transdisciplinary, cross-methodological investigations on causal relationships.

2. *The interventionist account of causal explanation*

The idea that causal relationships are somehow exploitable for purposes of control and manipulation has intuitive appeal. Francis Bacon came up with this idea in his *Novum Organon* (Bacon 1994). According to Bacon, we can only uncover and explain the hidden causal structures that nature hides from us by interacting with them (i.e. manipulating putative causes). This account of causal explanation has been further developed in the 20th century by philosophers (Menzies and Price 1993, von Wright 1971) as well as scientists (Freedman 1997, Holland 1986, Pearl 2000). It is commonly called the *manipulationist account of causal explanation*. The discussion of causal explanation in epigenetics presented below will mainly focus on James Woodward's interventionist account (Woodward 2003, 2004, 2007; Woodward and Hitchcock 2003)³. The

³ In contrast to the notion of manipulation, Woodward refers to a special kind of manipulation relevant to causal explanation. These manipulations are called interventions: Interventions are not based on

general applicability of interventionism to causal reasoning in biology has been demonstrated by Waters (2007) and Woodward (2001, 2010).

According to Woodward, the ‘stability’ of a generalization under a certain kind of manipulations, called interventions, on the value of its variables is what matters in scientific explanation. This basic idea of his interventionist account can be described as follows: *A variable X causes variable Y iff {(i) there is a possible manipulation (intervention) of some value of X which will change the value of Y (or the probability distribution of Y) and (ii) this relationship would remain stable}*.

This formula can be summarized under the slogan “invariance under interventions” (Woodward 2003, 250). Type-causal claims, the type of causal level Woodward’s account suits best, relate variables that can take more than one value. In this sense, *X* causes *Y*, if *X* can take the values {ingestion of aspirin, no ingestion of aspirin} and *Y* takes the values {relief from headache does occur, relief from headache does not occur} and if interventions on *X*, i.e. changes in the value of *X*, are accompanied by systematic changes in the value of *Y* (or the probability distribution of *Y*). Every time we wiggle the *X* cause, the effect *Y* will accordingly wiggle.

In addition, the following criteria have to be met in order to derive a generalization correctly⁴: Interventions should not change *Y* via a route that does not go through *X*. Nor should it affect other causes of *Y*, e.g., an off-path variable, say *Z*, which is not on the pathway *X* – *Y*, but connects these variables through another causal pathway. Only those variables that are intermediate causes of *Y* (if there are any) should be correlated with the intervention.

For instance, let us consider an experiment designed to determine whether treatment (variable *T* that takes one of the two values 1 and 0) with a drug causes recovery (*R*) from a disease by killing (*K*) a sort of bacterium. To determine the validity of the experiment we need to rule out the possibility that there are causes of *R* that are correlated with an intervention (*I*) on *T*, and that affect *R* independently of the causal pathway *I* → *T* → *K* → *R*. As will be argued below, such an experimental design (i.e. fixing off-path variables by additional manipulations) provides the methodological basis for causal reasoning in molecular studies on epigenetic inheritance systems.

As condition (ii) above indicates, the test for systematically changing invariant relationships between variables relies on a counterfactual notion of invariance: “[I]t has to do with whether a relationship *would* remain stable *if*, perhaps *contrary to actual fact*, certain changes or interventions occur” (Woodward 2003, 279; emphasis added). Creating these hypothetical, counterfactual situations is guided by asking *what-if-things-had-been-different-questions* or *w-questions*, as Woodward calls them. He argues that any successful causal explanation ought to be able to answer *w-questions* by presenting a “hypothetical or counterfactual experiment that shows us that and how manipulation of the factors mentioned in the explanation [...] would be a way of manipulating or alter-

any kind of human agency, which is often thought to be necessary for these interventions to be carried out. Thus Woodward’s interventionist account offers a non-reductive account of causal explanation, which has to be distinguished from other reductive manipulationist accounts.

⁴ For a more detailed description of these criteria, see Woodward 2003, 94-99.

ing the phenomenon explained” (Woodward 2003, 11). Following this explanatory approach, scientists derive generalizations from counterfactual situations which would be “invariant [i.e. they would hold] under some appropriate set of interventions” (Woodward 2003, 15).

Giving the experiment described above an interventionist/Woodwardian reading, the causal generalization “treatment with the drug causes recovery from a disease by killing a sort of bacterium” is invariant under a set or range of interventions that could be performed on the cause variable (treatment) to change the value of the effect variable (recovery). Concerning this set of manipulations, his generalization has explanatory power.

Recently some philosophers of science have argued that the description of mechanisms (rather than invariant dependencies) adequately accounts for the explanatory practice of scientists (Bechtel and Abrahamsen 2005; Darden 2005; Machamer et al. 2000; Tabery 2004). Usually identifying (components of) mechanisms shows *why* or *how* something works – such mechanistic explanations are bottom-up explanations describing the internal causal structure of a phenomenon. Especially in molecular biology, this idea to explain by identifying and elucidating mechanisms has been quite popular⁵. Does this mechanistic strategy contradict Woodward’s claim that causal explanations uncover patterns of systematic counterfactual dependence? As some have argued (Craver 2007; Glennan 2005; Woodward 2002), describing a mechanism of a phenomenon is consistent with the causal-interventionist way of explaining. Interventionist explanations might be considered as rather weak or superficial compared to mechanistic ones, because they often lack to convey how (or why) an observed dependency relation produces a phenomenon. Thus, mechanistic explanations can be understood as a complement to the interventionist strategy in terms of clarifying the internal causal structure underlying a certain systematic dependency relation.

Does the interventionist account of causal explanation presented in this section suitably describe causal reasoning in experimental and observational epigenetic research? This question will be addressed in the following section.

3. Epigenetic niggling in molecular biology: Multifactorial experimentation and artificiality

If we want to assess the theoretical impact epigenetics could have on the notion of heredity as used by neo-Darwinians, the assumed Lamarckian processes underlying transgenerational epigenetic effects have to be rendered more precisely. A first step to do so is to estimate the degree of sensitivity to environmental influences (i.e. developmental responsiveness) these systems show, before turning to inheritance phenomena. Manipulating environmental variables and observing corresponding effects in (transgenerational) phenotypes is perhaps the easiest way to achieve this goal. Many molecular studies on epigenetic mechanisms intervene on environmental variables by using pharmacological or toxicological agents (Anway et al. 2005), by giving a particu-

⁵ Paradigmatic examples of mechanisms discussed in philosophy of biology include protein synthesis, cellular metabolism, photosynthesis, as well as synaptic plasticity. For a more comprehensive list, see Raerinne 2011.

lar nutrition (Vastenhouw et al. 2006) or by exposing organisms to an enriched environment, including novel objects and elevated social interactions (Arai 2009) to induce certain (heritable) phenotypic effects. These methods take into account the high degree of developmental responsiveness and variability in epigenetic systems. In addition, *genetic* and *environmental* off-path variables have to be controlled for while wiggling epigenetic processes to preclude that any genes or environmental factors have causal influence on the observed (transgenerational) phenotypic effects. Thus multifactorial experimentation, a tool well known from genetics, takes the next level in epigenetics: Epigeneticists have to manipulate three different types of variables – epigenetic, environmental and genetic ones.

But are these manipulations in molecular epigenetics really interventions in the sense described by Woodward's account? And does epigeneticists' causal reasoning rely on counterfactuals? These two crucial questions will be addressed now by pointing to a lab study done by Vastenhouw et al. (2006), which exhibits a commonly used design in experimental epigenetics. Nadine Vastenhouw and her colleagues investigated long-term (i.e. transgenerational) gene silencing in the nematode *Caenorhabditis elegans* mediated by a process known as RNA interference (RNAi)⁶. They used a mutant strain with a gene (*gfp*) expressing green fluorescent protein (GFP) under the control of a germline-specific promoter; worms which expressed GFP in the germline fluoresced green when exposed to ultraviolet illumination. They then fed the animals bacteria that express double-stranded RNA (dsRNA) homologous to the *gfp*-DNA-sequence. This dsRNA triggered silencing of the gene *gfp* in the germline and lead to a *heritable* 'weaker-green phenotype' with reduced or no GFP expression. By selecting for this new phenotype gene silencing could be inherited for at least 80 generations. Figure 1 shows two path diagrams of the assumed causal system before and after the change in nutrition.

This experiment by Vastenhouw and her colleagues can easily be given an interventionist reading: The causal generalization G "environmentally induced gene silencing (i.e. epigenetic variation) produces a new phenotype that is heritable over many generations of sexual reproduction" is invariant under intervention on the value of the environmental (cause) variable to change the value of the transgenerational phenotype.

⁶ In RNA interference small RNA molecules bind to other specific messenger RNA (mRNA) molecules and either increase or decrease their activity (e.g., by mRNA degradation, which prevents mRNA from producing a certain protein). In addition, these small molecules can regulate the activity of genes directly on the DNA-level (e.g., by inducing DNA methylation). RNAi is involved in directing development and defending cells against parasitic genes. For a detailed description of RNAi-mediated gene silencing pathways, see Siomi and Siomi 2009.

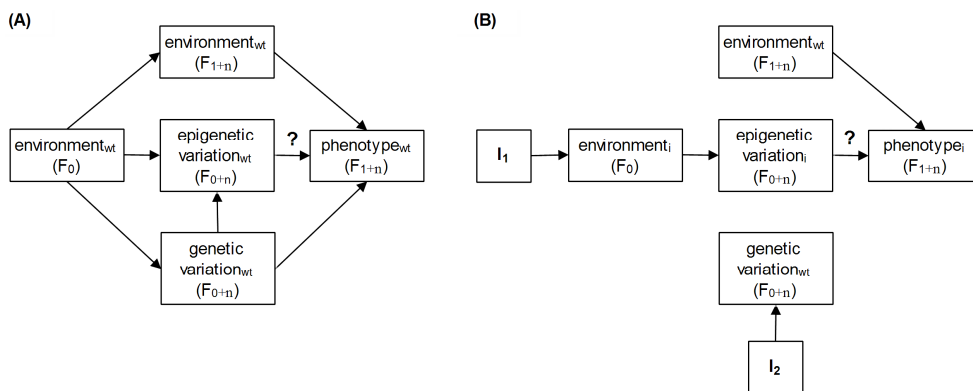


Figure 1: Two path diagrams representing the controlled experiment by Vastenhouw et al. (2006) on transgenerational gene silencing in the nematode *C. elegans* through manipulation of environmental, epigenetic and genetic variables. (A) The hypothetical causal system before experimental manipulation; (B) the causal system during (and after) manipulation. The question mark indicates the relationship under investigation (i.e. inheritance of phenotypically relevant epigenetic variation). Environment_{wt} represents the environment before manipulation, and variable environment_i represents changed nutrition *only* in the parental generation F₀. In path diagram (A) epigenetic as well as genetic variation is assumed to show wild-type values (wt); the transgenerational phenotype_{wt} can be induced by environment_{wt} in every generation anew without inheritance. In (B) intervention 1 (I₁; changing nutrition in F₀) changes the phenotype of the progeny (to phenotype_i); causal off-path variables (environment and genetic variation) are controlled for by I₁ and I₂ (i.e. using genetically identical organisms).

According to Woodward, this generalization rightfully claims explanatory power, because the manipulations performed in this experiment show the specific kind of causal structure described by the notion of intervention:

- (i) intervention 1 (I₁) on the environment variable avoids affecting other (environmental and genetic) causes of the phenotype – by feeding *C. elegans* bacteria expressing double-stranded-RNA *only* in the F₀ generation the new phenotype cannot be induced by the environment in every generation anew (without inheritance of genetic or epigenetic variation taking place) and by using non-mutagenic substances a change in the genetic material can be prevented (see two arrow-breaking events on the pathways ‘environment (F₀) → environment (F_{1+n})’ and ‘environment → genetic variation’ in figure 1B);
- (ii) intervention 2 (I₂) on the genetic variation variable represents using a strain of genetically identical organisms and selecting for the induced transgenerational phenotype (phenotype_i) in every generation (starting from F₁), thus fixing the genetic variation variable to value {wt} (see two

arrow-breaking events on the pathways ‘*genetic variation* → *epigenetic variation*’ and ‘*genetic variation* → *phenotype*’ in figure 1B).

By means of these interventions Vastenhouw et al. (2006) were able to demonstrate invariance of causal claim G. G, i.e. the dependency relation between RNAi induced epigenetic variation and the new heritable phenotype, remains stable during changes upon the inducing event. This is a very interesting dependency relation they were able to demonstrate, since it seems to exhibit invariance across the Weismann barrier⁷. In addition, G emphasizes the causal relevance of transgenerational epigenetic inheritance for heredity phenomena independently of other causal factors, like genes.

However, it remains to be shown that this interventionist view on causation and causal explanation underlying this typical methodology in experimental epigenetics is accompanied by counterfactual reasoning, i.e. epigeneticists’ seek to answer Woodwardian *w-questions* by proving systematic *counterfactual* dependencies. In the contemporary literature on epigenetic inheritance *w-questions* are often not as explicitly formulated as done by Conrad Hal Waddington, who coined the term ‘epigenetics’ in the late 1930s:

Manipulation of the variables in a model, or simulation, is intended to provide answers to the critical question, ‘*What would happen if...*’. (Waddington 1977, 207; emphasis added)

However, we find several papers in experimental epigenetics containing (hypothesized) counterfactual explanations in the form of “if the value of an environmental/epigenetic/genetic variable *X* would be changed, the value of the (transgenerational) phenotype *Y* would change accordingly”⁸. These findings indicate that interventions performed in experimental epigenetics are backed up by a counterfactual explanatory approach that guides development of experimental design applicable to answer Woodwardian *w-questions*.

Thus the interventionist approach to causation suitably describes the way molecular epigeneticists think about causation and causal explanation. However, these scientists seem not to regard invariant generalizations as explanatorily sufficient, since they seek to articulate the *mechanisms* explaining how an observed dependency relation arises, too. For example, Vastenhouw and her colleagues ask: “Is RNAi the mechanism behind the initial silencing?” (Vastenhouw 2006, 882)⁹. Others hope that including characterization of mechanisms underlying dependency relationships will “lead to a

⁷ According to this general principle, the manipulation on the environmental variable performed by Vastenhouw and her colleagues should not be able to trigger epigenetic variation in the germline that leads to a new, highly stable transgenerational phenotype. But yet, as the germline-specific promoter of the silenced gene *gfp* and interventions on values of off-path variables ensure, this seems to be possible.

⁸ For examples of counterfactuals in experimental epigenetics, see Anway *et al.* 2005, esp. 1467; Holmquist and Ashley 2006, various pages; Kelly and Aramayo 2007, various pages; Molnar *et al.* 2010, esp. 874; Rassoulzadegan *et al.* 2006, 472-473.

⁹ The invariant relationship between induction and transgenerational effect observed in this study does not elucidate, if RNAi is the mechanism producing the initiation of gene silencing (see causal arrow on the pathway ‘environment → epigenetic variation’ in figure 1) *as well as* the inheritance or maintenance of silencing (see question mark in figure 1).

more complete and better defined [causal] picture” (Rassoulzadegan et al. 2006, 473). It is not surprising that in molecular studies on epigenetic inheritance phenomena we find such ‘mechanism talk’. But explaining a phenomenon by describing a mechanism for that phenomenon is just a second explanatory step for epigeneticists that has to be distinguished from interventionist explanation¹⁰. As will be demonstrated below, the interventionist strategy of causal explanation alone is sufficient to establish (non-mechanistic) causal explanations precisely enough to develop new methodologies and to guide further research – even across disciplinary boundaries – on (the mechanisms of) the phenomenon under study. However, this special heuristic role of Woodwardian explanations does not render mechanistic explanation redundant in epigenetics.

The experimental study discussed above shows why multifactorial manipulation on different levels (environmental, epigenetic and genetic) is a necessary methodological tool in order to correctly prove invariance under intervention in molecular epigenetics. These lab-documented invariance states provide valuable insight into causal relationships within a certain set of variables. However, the relevant question addressed here is whether these dependency relations in artificial epigenetic systems occur under natural conditions as well¹¹.

The loss of similarity between natural and artificial epigenetic systems in epigenetics is addressed by critics who claim that epigenetic variation likely has no impact on natural populations (e.g., Pál and Hurst 2004). This critique heavily relies on the argument that molecular epigeneticists use invalid counterfactuals, or rather that they use generalizations holding *solely* in counterfactual situations. They do not assume that such counterfactuals are too vague to lead to an appropriate experimental design (since they are in fact able to do so), nor do they claim that generalizations on causal relationships in epigenetics are not invariant in the sense described by Woodward¹². In fact, it is argued that the artificial situations going along with epigenetic causal reasoning are not counterfactual in nature, because there is no real-world counterpart for them.

¹⁰ For reasons of space it cannot be discussed here, what it means in detail for epigeneticists to explain by articulating mechanisms, e.g., if epigeneticists’ mechanistic explanations are consistent with interventionist definitions of (representations of) mechanisms (see Craver 2007; Glennan 2005; Woodward 2002).

¹¹ It should be mentioned that this distinction between artificial and natural epigenetic systems is not a clear-cut one. In biology artificiality comes in degrees – it appears in field studies as well (e.g., by choosing a suitable population or ecosystem). Here ‘artificial conditions’ refers to research practices based on physical manipulation.

¹² As the case of Vastenhouw et al. 2006 demonstrates in an exemplary manner, causal explanation in experimental epigenetic studies focuses on Woodward’s notion of invariance rather than on the notion of stability, advocated by Sandra Mitchell (e.g., Mitchell 2000). According to Mitchell, a generalization is called *stable* if it holds in many possible background conditions. Testing *stability* is rarely used in epigenetics (an exception is Jablonka and Raz 2009 – this survey should give proof that a certain epigenetic dependency relation can be found in a variety of species). For a discussion on the relevance of Woodward’s notion of invariance and Mitchell’s notion of stability in biological causal reasoning, see Woodward 2010.

To illustrate this crucial problem, let's have a look at another experimental study on RNA interference: Rassoulzadegan et al. (2006) showed that the injection of small RNAs into fertilized eggs causes a heritable epigenetic modification (paramutation) in male mice, which results in white tail-tips and paws. However, they had to admit: "The *initial event* inducing paramutation *is not known*" (Rassoulzadegan et al. 2006, 473; emphasis added). There is no 'real environment'-value known that the inducing cause variable could take.

However, some phenotypes caused by natural (heritable) epigenetic variation are already known (Bossdorf et al. 2008), but identifying relevant causal relationships between epigenetic variation and environmental or genetic variables remains tricky in the absence of experimental analysis¹³. Without manipulative tools available, distinguishing a stable epigenetically caused phenotype from a genetically caused one is nearly impossible. Thus still most of the (heritable) variation in epigenetic systems known is brought about in the lab under artificial situations.

To gain knowledge of real-world causal processes underlying natural epigenetic variation (i.e. the naturally operating counterparts of artificial conditions in the lab) and to estimate the evolutionary relevance of epigenetic inheritance, it is important that the population or ecological level is made assessable to causal reasoning, too. This goes along with the problem mentioned earlier of linking causal (molecular) models based on counterfactual reasoning and multifactorial manipulation with quantitative (i.e. statistical) models derived from observational data analysis. I will outline a solution to this 'translation problem' below.

4. *Getting out of the lab: Observational studies and the solution to the 'translation problem'*

How frequently are epigenetic variants environmentally induced and how often are they transmitted between generations? What impact does this transmission have on natural population dynamics? To answer these essential evolutionary questions, the subfield of population or ecological epigenetics has recently emerged. It includes different lab and field study techniques, like studies conducted under controlled conditions (e.g., greenhouse studies; see Bossdorf et al. 2008; Johannes et al. 2008) and purely observational studies on natural populations (Herrera and Bazaga 2011). Only the last-mentioned account, also known as 'natural experiment' (Diamond 1986), addresses this question fully, since it is the only one avoiding manipulation of the population under study. In contrast to epigeneticists working in the lab, ecologists performing natural experiments use *non*-experimental methods, like regression equations, to infer causation from statistical relevance.

This methodological distinction leads us to the essential questions: Do ecology epigeneticists infer causation from correlation patterns in observational data on natural populations by means of the molecular epigeneticists' view of causal explanation, i.e.

¹³ As Vastenhouw and her colleagues showed, such an experimental analysis does not necessitate articulating molecular mechanisms in order to conclude whether an induced phenotypic effect is caused by a genetic or an epigenetic modification – demonstrating invariance under intervention is sufficient to do so.

the interventionist account? Does interventionism provide epigenetics with reliable criteria to develop methodological strategies that enable these scientists to translate causal molecular models into statistical models and vice versa? I will argue that the answer to both questions is “yes”.

At first sight this might look strange, simply because we are not able to manipulate purely observational data. This is well known at least since Robert Lucas (1976) argued that performing interventions (e.g., policy making) on a system under study always disturbs the stability of this systems. If we nevertheless claim that in observational studies it is useful to think about causation as being potentially exploitable for purposes of manipulation, it is therefore necessary to present a notion of invariance under intervention that:

- (i) accounts for statistical methods used in the relevant (sub)discipline,
- (ii) is a counterfactually defined notion in the sense described above and
- (iii) is backed up by ‘context information’ that tells us in which observational studies
- (iv) the interventionist rationale should be chosen over non-manipulative, non-counterfactual ways of causal explanation.

Recently Federica Russo (Russo 2011, 2012; Russo et al. 2011), focusing on structural equation modeling techniques in observational studies in the social sciences, argued against condition (i), that Woodwardian invariant generalizations are not central to causal inference in observational contexts, since testing the condition of invariance *does not* take manipulation as methodologically fundamental:

Rather than manipulation, the basic idea or rationale underpinning causal analysis [in observational and in experimental contexts] is that some form of joint *variation* between variables of interest has to be evaluated. (Russo et al. 2011, 52; emphasis in original)

According to Russo, the notion of (co-)variation of variables acts as a precondition to the notion of manipulation and not vice versa. Therefore in non-experimental contexts the first question to be asked is whether a data set reveals meaningful co-variation between putative cause and effect variables. Subsequently, scientists can use further tests, like the one for invariance, to determine whether the observed co-variation is in fact causal¹⁴.

Against condition (ii), Russo argues that even if statistical techniques which manipulate values of variables in a data set are used in observational studies to determine whether associations are causal, such invariance-tests are not counterfactual ones: We do not test whether generalizations *would* remain stable if we were to intervene, but rather whether they *are* in fact stable across subpopulations or different partitions of the

¹⁴ For a thorough discussion of Russo’s rationale of variation, see Russo 2009. Russo claims that this rationale is central to experimental contexts as well: “[V]ariation not only guides causal reasoning in observational settings, but does so also in *experimental* ones” (Russo 2012, 136; emphasis added). This claim is questionable at least for disciplines like epigenetics, where most of the (co-)variation known is brought about in the lab by manipulation in the first place.

data set (Russo 2011, 2012). Thus, according to Russo, it is rather a *non*-counterfactual concept of invariance that establishes causal reasoning in observational contexts¹⁵.

To argue in favour of condition (i) to (iii), I will point to what may be the first truly ecological epigenetics study – a natural experiment done by community ecologists Herrera and Bazaga (2011) – and especially to the way interventionist reasoning is linked to statisticians' modeling techniques in this case. Two modeling tools essential for observational studies will be addressed here: statistical conditioning and specification of structural models¹⁶.

Statistical conditioning implies, simply put, observing a population over a long period of time and then statistically controlling for a certain variable on paper without any physical changes in the field population. Thus we separate the data into groups that are homogenous with respect to the controlled variable, e.g., by controlling for potentially confounding variables.

Structural equation modeling (SEM) is a multivariate statistical analysis method often used in quantitative population research in ecology or evolutionary biology (Pugesek et al. 2003; Shipley 2000). SEM models allow translating a set of hypothesized causal relationships between variables into a model concerning patterns of statistical dependencies. It offers methods for modeling complex systems with many assumed causal relationships between latent (not directly observed) variables. As in regression dependent variables in structural equations regress on independent variables, meaning that the dependent variables are being predicted by the independent ones.

Testing a proposed theoretical model against empirical data is known as the *confirmatory* aspect of SEM. If a model is difficult to fit on the data (e.g., because theory is vague) and repeated model modification is necessary, SEM enters into the *explanatory* mode, generally referred to as 'specification search' (Leamer 1978). It is this explanatory mode that matters most for SEM in ecological epigenetics, since the epigenetic theory of natural transgenerational inheritance is young and likely necessitates modification¹⁷.

Carlos Herrera and Pilar Bazaga (2011) observed that epigenetic variation in a wild population of the southern Spanish plant *Viola cazorlensis*, a long-lived non-model organism, is significantly correlated with long-term (i.e. two decades) differences in herbivory (i.e. magnitude of browsing damage), but only weakly with herbivory-related DNA sequence variation (for correlation pattern, see figure 2A)¹⁸. This study com-

¹⁵ For a detailed discussion of Russo's concept of causal explanation in the social sciences, see Mateiescu 2012.

¹⁶ I will focus on manipulative methods of building and modifying structural equation models rather than on 'interventionist' structural model testing, since there is already a vast literature on this subject (see, e.g., Freedman 1997; Pearl 2000; Woodward 2003; Tanaka et al. 2011).

¹⁷ This is the case, because "[u]ntil a few years ago, epigenetics was a field of research that had nothing to do with ecology and that virtually no ecologist had ever heard of" (Bossdorf and Zhang 2011, 1572).

¹⁸ The epigenetic variation in this species refers to multilocus differences in DNA-methylation patterns among individuals measured by using methylation-sensitive molecular markers; DNA sequence variation refers to multilocus genetic differences among individuals exclusively in those loci known to be significantly related to herbivory. For a review on this study, see Bossdorf and Zhang 2011.

bines two types of natural experiments in community ecology (see Diamond 1986): (i) a natural trajectory experiment – collecting herbivory data of a 20-year long series of annual herbivory for each of the plants in the same population; (ii) a natural snapshot experiment – observing a final steady state of the community under study (i.e. comparing information on genetic and epigenetic characteristics of the plants *after* 20 years and differences in substrate type and long-term herbivore).

From the correlation pattern (depicted in figure 2A) they derived the final causal thesis that natural epigenetic variation is likely to be, in addition to genetic and environmental variables, an (at least partly) independent cause, which influences interactions between plants and herbivores (i.e. herbivory damage) in the field. But what kind of causal reasoning supports this derivation?

The amount of empirical data and reliable theoretical models usable as a starting point for this ecological study was very low, since “[n]ext to nothing is known [...] on the magnitude and ecological correlates of epigenetic variation of non-model organisms in natural environments“ (Herrera and Bazaga 2011, 1676). Therefore they used a combination of different model modification techniques to build a structural equation model that increasingly fits the observed data:

First they applied a method of ‘statistical manipulation’, referred to as statistical conditioning above, to build a causal model. Therefore they intervened on a part of the assumed causal structure by changing the environmental substrate type variable to a certain *non-actual* value $\{i\}$. The substrate type variable correlates with the measured association between the variable of special interest (i.e. epigenetic variation and herbivory damage). It takes the values $\{\text{rock, cliffs, sandy soil}\}$ among individuals in the population under study. Statistically holding this variable fixed at value $\{i\}$ helps to stabilize the total effect on the potential output variable (herbivory damage)¹⁹. Figure 2B shows a path diagram of the causal system after statistical manipulation.

In a second methodological step, directly leading from statistical conditioning, Herrera and Bazaga (2011) designed a competing model situation that includes four alternative, most likely causal models varying in the fit to the data. These models correspond to different theoretical positions on the potential causal role of epigenetic variation. This model specification procedure makes concessions to the fact that the theory of natural epigenetic inheritance is relatively vague, i.e. backed up by very little empirical data. Therefore they changed the causal relationships between variables in the model in various ways (linking the variables through different causal routes). Each time the model was re-evaluated using standard goodness of fit statistics. By this modification procedure they were able to rule out a couple of non-fitting, hypothesised causal models.

¹⁹ For a mathematical optimization procedure to intervention effects that reduce the variances of the output variable in SEM, see Tanaka et al. 2011.

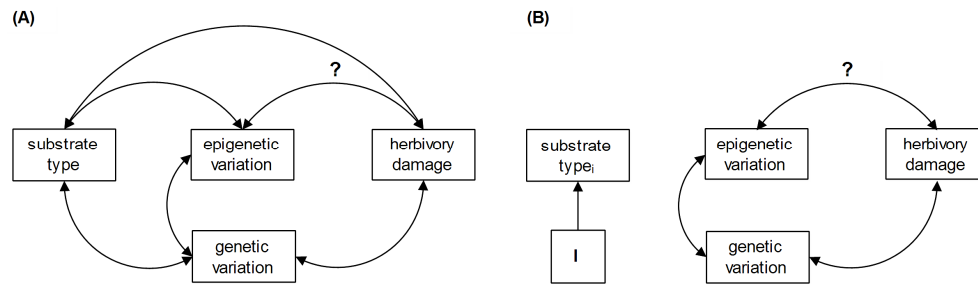


Figure 2: A simplified path diagram representing the correlation pattern observed by Herrera and Bazaga (2011) in a wild population of *V. cazorlensis* before and after statistical manipulation on an environmental variable. (A) The causal system before manipulation; (B) the causal system after manipulation (i.e. statistical population_i). The question mark indicates the assumed causal relationship epigeneticists are primarily interested in. In diagram (B) an intervention (I) changes the value of the substrate type variable, holding it fixed at value {i} (depicted as three arrow-breaking events). Substrate type variable does not change in value in population_i.

But unfortunately, in the end, there were two different models left over with equally good fit to the data. These two models are represented together in one path diagram in figure 3.

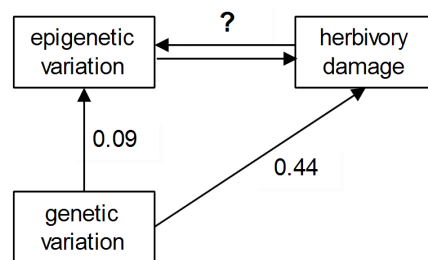


Figure 3: A simplified path diagram representing the two best fitting causal models of Herrera and Bazaga (2011). The question mark indicates a nonrecursive (bidirectional) relation that clarifies the difference between the two causal models (each contains one of the recursive (unidirectional) relationships 'epigenetic variation → herbivory damage' and 'herbivory damage → epigenetic variation'). Both causal models (i.e. the bidirectional relation) could in fact be correct. Path coefficient 0.09 indicates a relatively weak causal influence (only 9% of relevant epigenetic variation explained by herbivory-related genetic variation); path coefficient 0.44 (44% of herbivory damage explained by herbivory-related genetic variation).

Are these model building and model specification procedures used by Herrera and Bazaga (2011) guided by an interventionist view on causal explanation similar to the

one we find in experimental epigenetics? The method of statistical manipulation, i.e. changing the substrate type variable to a certain *non-actual* value, initially used in this ecological epigenetics study, can easily be understood as being governed by Woodwardian *w-questions* like “What would happen if the plants would not live on their natural substrate type (e.g., rocks)?”, which help to create a counterfactual, artificial population (population_i, depicted in figure 2B) where the value of this environmental variable does not change naturally²⁰:

Because we were interested in the herbivory-genetic-epigenetic causal structure *after* controlling for ecological, substrate-related effects, the *residuals* after controlling for the average effect of substrate on herbivory were used [...] *rather than the raw herbivory data*. (Herrera and Bazaga 2011, 1679; emphasis added)

This counterfactual population_i (plants living in an environment which does not interact with them through substrate types) is compared with the natural population to see whether the observed relationship between epigenetic (and genetic) and herbivory variation would remain stable or invariant. This procedure is closely linked to Woodward’s concept of ‘weak invariance’ (Woodward 2003, chapter 6.15 and 7.8). Referring to observational data, i.e. an epidemiological study on the causal relationship between smoking and lung cancer done by Cornfield et al. (1959), the notion of weak invariance means, simply put, that the generalization “smoking causes lung cancer” would remain stable across different partitions of the data set (e.g., across groups of people with different genetic backgrounds or socioeconomic conditions). These subpopulations of the data set are brought about by controlling for potentially confounding variables. Recently Russo (2011, 2012) argued that this notion of weak invariance, essential for SEM in observational studies, is not the standard-Woodwardian notion of *counterfactual* invariance, but rather a kind of *factual* invariance (dealing just with observed, ‘real-world’ data). But, in fact, it is a counterfactual one: As shown above, testing (weak) invariance means testing whether the generalization “epigenetic variation causes difference in herbivory damage” would remain stable, if we were to bring about a certain *counterfactual situation* by controlling for the substrate type variable. It is this notion of counterfactual invariance that is crucial both for Woodward’s account on causal inference in SEM (Woodward 2003, chapter 7.2) and for SEM model building techniques in ecological epigenetics, where testing for weak invariance provides information about which variables should be included or excluded from a causal model. However, like in molecular epigenetics ecological epigeneticists seem not to regard invariant generalizations as explanatory sufficient, since they seek to articulate the mechanisms underlying a dependency relation as well (see Herrera and Bazaga 2011, 1685)²¹.

As the bidirectional relationship in figure 3 correctly indicates, the manipulative model building and modification procedures used in this study carry with them a cer-

²⁰ The notion of population in this case in fact refers to a *statistical* population, but I understand it as being closely linked to the general notion of (natural) population by biologists. For counterfactuals in this paper, see, e.g., Herrera and Bazaga 2011, 1863, 1865.

²¹ On the connection between mechanistic and causal-interventionist explanations in ecology, see Păslaru 2009; Raerinne 2011.

tain coarse-grainedness or fuzziness, since the correct causal relationship between epigenetic variation and ecologically relevant traits must remain to some degree unclear. This is likely to be a general problem ecological epigenetics has to deal with in future natural snapshot experiments. In ‘snapshot’ data sets phenotypes measured in the field at a particular time reflect both heritable variation as well as environmentally induced phenotypic plasticity (i.e. less stable and reversible variation)²². For this reason it is difficult to decide between alternative models trying to explain the process leading to adaptive heritable epigenetic variation.

Does this mean that proving weak invariance of the association ‘epigenetic variation – herbivory damage’ does not causally explain at all, since it offers no more information than the measured correlation of these variables? From the perspective of Herrera and Bazaga (2011), already before the model modification procedure was started all four models introduced offer causal explanations: They “*parsimoniously explain* the observed associations between herbivory, genotype and epigenotype” (Herrera and Bazaga 2011, 1682; emphasis added). Do these ecologists simply forget that correlations by themselves are not explanatory? I claim that the dependency relation under study (see question mark in figure 3) in fact causally explains, although in a weak or superficial sense, because there is a statistical intervention during which this special association remains invariant. Such a manipulation specifies the causal structure underlying the observed association and supplies *auxiliary causal evidence*²³: It offers knowledge of whether certain causal factors (e.g., habitat parameters) are causally relevant for the association under study – as long as the dependency relation persists during intervention on the values of these variables (i.e. changing them to non-actual values) they can be excluded from a causal explanation. The dependency relation between two variables holds in some specific causal situations or settings brought about by an intervention, while in others it does not. In this sense, proving Woodwardian counterfactual invariance offers crucial information for model building in observational data analysis²⁴.

Of course, proving weak invariance among variables in observational snapshot data sets necessitates additional experimental approaches to tease apart different causal models and to disentangle plastic from heritable epigenetic variation:

²² In plants epigenetic variation can be induced by environmental stresses, including herbivory damage. This variation is not necessarily heritable.

²³ The need to seek for auxiliary causal evidence supplied by ‘invariance knowledge’ has been discussed in other observational research fields as well, e.g., in epidemiology (Lilienfeld and Lilienfeld 1980, chapter 12).

²⁴ In addition, this procedure helps to minimize the total effect of the negligible causal background on the relevant variables – a crucial method to estimate the influence of causally relevant variables (included in the model) on the causal association under study. In the case of Herrera and Bazaga 2011, after performing this interventionist strategy, only 9% of multilocus epigenetic variation could be explained by herbivory-related genetic variation. From this finding, they derived the final superficial causal explanation, that some values of the epigenetic variation variable (likely) are related to herbivory *independent* of genetic variation. Thus genetic variation would not be a common cause of the association while entailing this special sort of epigenetic variation.

A comprehensive research programme in ecological epigenetics must include molecular studies and controlled experiments, but also field studies that test whether epigenetic patterns in natural populations are consistent with theoretical predictions and the results of more controlled, but less realistic, experiments. (Bossdorf and Zhang 2011, 1573)

Such a program can be developed on the basis of the interventionist account, acting here as a “rationale underpinning causal reasoning” (Russo 2012, 131): Understanding causation and causal explanation as being closely linked to the notions of invariance under intervention and to counterfactual reasoning, like Vastenshouw et al. (2006) and Herrera and Bazaga (2011) do, helps to unify different methods in experimental and observational studies and to develop reciprocal transparent models in ecological epigenetics and molecular biology.

A consistent interventionism-based research strategy especially makes sense, referring to ‘context information’-condition (iii) above, if we have little information about the observed system (and its underlying mechanisms) before building causal models and if a unification of methods makes sense (i.e. if we are able to link observational studies to a thriving field of research that heavily relies on manipulation). Both are the case in epigenetics: Here causal modeling in the young subdiscipline of ecological epigenetics has major advantages in being connectable to reliable causal explanations coming from experimental studies to secure understanding of complex correlational patterns in the field²⁵.

Even when the derived causal model shows a certain fuzziness, as in the case of structural equation modeling above, following manipulative experiments can buck up smoothly and compensate this handicap by continuing to uncover the correct causal relationship. Thinking about interventions and counterfactual situations in order to test for invariance in observational studies directly leads to a particular experimental design in additional controlled experiments. It is this property that lies at the center of the idea to specify test-criteria or strategies for reciprocal transparent studies. Here are some of these heuristic research strategies relevant for epigenetics²⁶:

- (a) *Synchronizing lab experiments, field studies and natural experiments: W-questions*, like “What would happen if the plants did not live on rocks?“, asked in observational studies to create counterfactual statistical populations, can act as starting points to design connectable hybrid lab-field studies or common environment studies (e.g., greenhouse studies with fixed environmental variables). These investigations can also perform the test for invariance under intervention, although by using different manipulation

²⁵ Condition (iii), justifying the reasonableness of interventionist reasoning in observational contexts, is also compiled by the observational study on the association between smoking and lung cancer (Cornfield et al. 1959), discussed by Woodward. As Woodward correctly describes, despite epidemiological evidence on the stability of this association, “[a] similar stable association is found among laboratory animals exposed to tobacco smoke” (Woodward 2003, 312). Initially limited knowledge of underlying biochemical mechanisms could thus be compensated by a unifying explanatory approach (underlying statistical and physical manipulation techniques) to secure cumulative evidence.

²⁶ This list does not present an exhaustive catalog of methodological guidelines. It should rather provide the reader insight into the heuristic spectrum of interventionist reasoning (in epigenetics).

tools. In addition, methodological constraints in manipulative common environment approaches can have reflexive influence on natural environment studies, especially on considerations which natural population or which ecosystem to choose. Is it, for example, possible to preserve relevant real-world values of inducing environmental variables of a given ecosystem in a certain common environment approach?

- (b) *Generating transferable information*: Natural snapshot experiments ‘counterfactual population modeling’ likely leads to fuzzy causal models and new unanswered *w-questions*. But at this point following experiments or hybrid lab-field studies can take over by proceeding observational information about which parameter combinations have to be held fixed and which have to be wiggled. Thus the interventionist rationale leads to heuristics for further model specification and testing²⁷.
- (c) *Choosing (non-)model organisms*: The interventionist rationale helps to choose suitable model or non-model organisms, e.g., it denies choosing very long-lived organisms (like Herrera and Bazaga 2011 did) where testing invariance under physical manipulation may not be feasible.

Although there are other accounts on the market trying to address the topic of connecting observational and experimental epigenetic research²⁸, I claim that picking an interventionist explanatory strategy (based on the notion of invariance under intervention and on counterfactual reasoning) and manipulative methodologies is a pretty good choice, because it enables scientists to establish transdisciplinary, reciprocal transparent research with greater simplicity – an issue which is crucial for this quite young field to benefit from causal knowledge coming from lab experiments.

5. Conclusion

Biology, if done in the lab, is manipulationists’ business. As recent experimental work on epigenetic inheritance phenomena clearly indicates, the interventionist account elucidates the way how epigeneticists in molecular biology think about causation and causal explanation. Based on this explanatory framework, experimental epigenetics has developed a comprehensive methodological framework that rests upon manipulation

²⁷ In the case discussed above, the final model yields the new counterfactual “[i]f the methylation states of at least some of the MSAP markers [indicating multilocus epigenetic variation] associated with the resistance of *V. cazorlensis* to mammal herbivory were heritable, then the association between epigenotype and herbivory would translate into herbivore-driven selection on epigenotypes” (Herrera and Bazaga 2011, 1685; emphasis in original). In addition, new *w-questions* like “What would happen if all environmentally labile epigenetic variation would be controlled?” instantly arise. This special question could be answered by a Vastenhouw et al. (2006)-like removal experiment that fixes the herbivory damage variable inducing epigenetic change at value {no herbivory damage} after generation F_0 (e.g., by using fences in a field experiment).

²⁸ For example, see Tal et al. 2010. This rather non-interventionist account on epigenetic contribution to covariance between relatives considers trajectory information about the number of opportunities for epigenetic reset between generations and assumptions about environmental induction to determine heritable epigenetic variance and epigenetic transmissibility.

of inducing environmental variables and multifactorial experimentation (especially to control off-path variables on different levels). However, epigeneticists' interventionist causal claims holding solely under conditions that could be brought about in a living system by means of physical manipulation conflict with the interest of biologists to explain causal relationships in naturally operating systems. This contradiction is enhanced by the fact that epigenetics often deals with phenomena which are difficult to study without manipulation, because we cannot distinguish between natural (transgenerational) phenotypic effects brought about by genetic, epigenetic or environmental causes.

To escape this conflict between emphasizing manipulation and explaining natural associations, expanding research efforts into ecologically relevant circumstances have to be made. In addition, this step is essential to estimate the impact of the epigeneticists' Lamarckian framework on neo-Darwinism. I have argued that these efforts in ecological epigenetics go along with the crucial challenge to articulate reliable criteria and research strategies to interrelate molecular biologists' causal models with statistical models on observational data.

As the discussed epigenetic natural experiment shows, proving counterfactual dependencies among variables offers crucial causal knowledge for model building and modification in observational data analysis. These interventionist explanations can act as a rationale guiding further causal reasoning to unify methods from different fields of research. It is important to mention that arguing for this heuristic value of invariant generalizations does not imply claiming that the test for invariance under intervention fully captures what epigeneticists regard as explanatory, neither in experimental nor in observational contexts. But these explanations are sufficient to establish transdisciplinary communication on causal relationships to infer causal claims. In this sense, demonstrating Woodwardian systematic dependence qualifies as a minimal (explanatory) consensus. Even with no mechanistic explanation yet available, preceding interventionist explanation provides a crucial heuristic tool to coordinate further research efforts across disciplinary boundaries on (the mechanisms of) the phenomena under study.

This tool can take full effect if we have little knowledge about the underlying physico-chemical structure and repeated model specification is necessary. Such an interventionism-based heuristic research strategy enables ecological epigeneticists to point to an appropriate experimental design of additional manipulative experiments even if causal explanation in observational studies is rather fuzzy or superficial. Thus the 'interventionist rationale' helps to bridge the explanatory and methodological gap between molecular and evolutionary epigenetics.

Finally it should be stressed that this special 'translating capacity' of interventionism is not necessarily limited to biological research: Although physical manipulations are not day-to-day business in the social sciences, observational studies on social phenomena can intensify using this heuristic tool to bring causal modeling into agreement with experimental social sciences' methodologies.

REFERENCES

- Anway, Matthew D., Andrea S. Cupp, Mehmet Uzumcu, and Michael K. Skinner. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308: 1466–1469.
- Arai, Junko, Shaomin Li, Dean M Hartley, and Larry A. Feig. 2009. Transgenerational rescue of a genetic defect in long-term potentiation and memory formation by juvenile enrichment. *The Journal of Neuroscience* 29: 1496–1502.
- Bacon, Francis. 1994. *The Novum Organon*. Ed. and trans. Peter Uhrbach and John Gibson. Chicago, La Salle: Open Court. Originally published 1620.
- Bechtel, William, and Adele Abrahamsen. 2005. Explanation: A mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 421–441.
- Bosssdorf, Oliver, Christina L. Richards, and Massimo Pigliucci. 2008. Epigenetics for ecologists. *Ecology Letters* 11: 106–115.
- Bosssdorf, Oliver, and Yuanye Zhang. 2011. A truly ecological epigenetics study. *Molecular Ecology* 20: 1572–1574.
- Cabej, Nelson R. 2008. *Epigenetic principles of evolution*. Dumont: Albanet.
- Cornfield, Jerome, William Haenszel, E. Cuyler Hammond, Abraham M. Lilienfeld, Michael B. Shimkin, and Ernst L. Wynder. 1959. Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute* 22: 173–203.
- Craver, Carl F. 2007. *Explaining the brain*. Oxford: Clarendon Press.
- Darden, Lindley. 2005. Relations among fields: Mendelian, cytological and molecular mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 349–371.
- Diamond, Jared M. 1986. Overview: Laboratory experiments, field experiments, and natural experiments. In *Community ecology*, ed. Jared M. Diamond and Ted J. Case, 3–22. New York: Harper & Row.
- Egger, Gerda, Gangning Liang, Ana Aparicio, and Peter A. Jones. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 429: 457–463.
- Freedman, David. 1997. From association to causation via regression. *Advances in applied mathematics* 18: 59–110.
- Gissis, Snait, and Eva Jablonka. 2011. *Transformations of Lamarckism*. Cambridge: MIT Press.
- Glennan, Stuart. 2005. Modeling mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: 443–464.
- Griesemer, James. 2011. The relative significance of epigenetic inheritance in evolution: Some philosophical Considerations. In *Transformations of Lamarckism*, ed. Snait Gissis and Eva Jablonka, 331–344. Cambridge: MIT Press.
- Hall, Brian K. 1998. Epigenetics: Regulation not replication. *Journal of Evolutionary Biology* 11: 201–205.
- Herrera, Carlos M., and Pilar Bazaga. 2011. Untangling individual variation in natural populations: Ecological, genetic and epigenetic correlates of long-term inequality in herbivory. *Molecular Ecology* 20: 1675–1688.
- Holland, Paul W. 1986. Statistics and causal inference. *Journal of the American Statistical Association* 81: 945–960.
- Holmquist, Gerald P., and Terry Ashley. 2006. Chromosome organization and chromatin modification: Influence on genome function and evolution. *Cytogenetic and Genome Research* 114: 96–125.
- Jablonka, Eva, and Marion J. Lamb. 1995. *Epigenetic inheritance and evolution: The Lamarckian dimension*. Oxford: Oxford University Press.
- . 2005. *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge: MIT Press.
- . 2008. Soft inheritance: Challenging the modern synthesis. *Genetics and Molecular Biology* 31: 389–395.
- Jablonka, Eva, and Gal Raz. 2009. Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology* 84: 131–176.
- Johannes, Frank, Vincent Colot, and Ritsert C. Jansen. 2008. Epigenome dynamics: A quantitative genetics perspective. *Nature Reviews Genetics* 9: 883–890.
- Kelly, William G., and Rodolfo Aramayo. 2007. Meiotic silencing and the epigenetics of sex. *Chromosome Research* 15: 633–651.
- Leamer, Edward E. 1978. *Specification searches: Ad hoc inference with nonexperimental data*. New York: Wiley.

- Lilienfeld, Abraham M., and David E. Lilienfeld. 1980. *Foundations of epidemiology*. 2nd edition. Oxford: Oxford University Press.
- Lucas, Robert. 1976. Econometric policy evaluation: A critique. In *The Phillips curve and labor markets*, ed. Karl Brunner and Allan H. Meltzer, 19–46. Amsterdam: North-Holland.
- Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. Thinking about mechanisms. *Philosophy of Science* 67: 1–25.
- Mateiescu, Sebastian. 2012. The limits of interventionism – causality in the social sciences. In *Probabilities, laws, and structures*, ed. Dennis Dieks, Wenceslao J. Gonzalez, Stephen Hartmann, Michel Stöltzner, and Marcel Weber, 141–156. Vol. 3 of *The philosophy of science in a European perspective*. Dordrecht: Springer.
- Menzies, Peter, and Huw Price. 1993. Causation as a secondary quality. *The British Journal for the Philosophy of Science* 44: 187–203.
- Mitchell, Sandra D. 2000. Dimensions of scientific law. *Philosophy of Science* 67: 242–265.
- Molnar, Attila, Charles W. Melnyk, Andrew Bassett, Thomas J. Hardcastle, Ruth Dunn, and David C. Baulcombe. 2010. Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science* 328: 872–875.
- Pál, Csaba, and Laurence D. Hurst. 2004. Epigenetic inheritance and evolutionary adaptation. In *Organelles, genomes and eukaryote phylogeny*, ed. Robert P. Hirt and David S. Horner, 353–370. London: Taylor and Francis.
- Páslaru, Viorel. 2009. Ecological explanation between manipulation and mechanism description. *Philosophy of Science* 76: 821–837.
- Pearl, Judea. 2000. *Causality: Models, reasoning, and inference*. New York: Cambridge University Press.
- Pigliucci, Massimo. 2009. An extended synthesis for evolutionary biology. *Annals of the New York Academy of Sciences* 1168: 218–228.
- Pigliucci, Massimo, and Gerd B. Müller. 2010. *Evolution: The extended synthesis*. Cambridge: MIT Press.
- Pugsek, Bruce H., and Adrian Tomer. 2003. *Structural equation modeling: Applications in ecological and evolutionary biology*. Cambridge: Cambridge University Press.
- Raerinne, Jani. 2011. Causal and mechanistic explanations in ecology. *Acta Biotheoretica* 59: 251–271.
- Rassoulzadegan, Mino, Valérie Grandjean, Pierre Gounon, Stéphane Vincent, Isabelle Gillot, and François Cuzin. 2006. RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. *Nature* 441: 469–474.
- Richards, Christina L., Oliver Bossdorf, and Massimo Pigliucci. 2010. What role does heritable epigenetic variation play in phenotypic evolution? *BioScience* 60: 232–237.
- Richards, Eric J. 2006. Inherited epigenetic variation – revisiting soft inheritance. *Nature Reviews Genetics* 7: 395–401.
- Royle, J. Andrew, and Robert M. Dorazio. 2008. *Hierarchical modeling and inference in ecology*. San Diego: Academic Press.
- Russo, Federica. 2009. *Causality and causal modelling in the social sciences*. New York: Springer.
- Russo, Federica. 2011. Correlational data, causal hypotheses, and validity. *Journal for General Philosophy of Science* 42: 85–107.
- . 2012. On empirical generalisations. In *Probabilities, laws, and structures*, ed. Dennis Dieks, Wenceslao J. Gonzalez, Stephen Hartmann, Michel Stöltzner, and Marcel Weber, 123–139. Vol. 3 of *The philosophy of science in a European perspective*. Dordrecht: Springer.
- Russo, Federica, Guillaume Wunsch, and Michel Mouchart. 2011. Inferring causality through counterfactuals in observational studies – some epistemological issues. *Bulletin de Méthodologie Sociologique* 111: 43–64.
- Sano, Hiroshi. 2010. Inheritance of acquired traits in plants. *Plant Signaling & Behavior* 5: 346–348.
- Shipley, Bill. 2000. *Cause and correlation in biology*. Cambridge: Cambridge University Press.
- Siomi, Haruhiko, and Mikiko C. Siomi. 2009. On the road to reading the RNA-interference code. *Nature* 457: 396–404.
- Suzuki, Miho M., and Adrian Bird. 2008. DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics* 9: 465–476.
- Tabery, James G. 2004. Synthesizing activities and interactions in the concept of a mechanism. *Philosophy of Science* 71: 1–15.

- Tal, Omri, Eva Kisdi, and Eva Jablonka. 2010. Epigenetic contribution to covariance between relatives. *Genetics* 184: 1037–1050.
- Tanaka, Kentaro, Atsushi Yagishita, and Masami Miyakawa. 2011. Application of mathematical optimization procedures to intervention effects in structural equation models. *Technical Report, Department of Industrial Engineering and Management – Tokyo Institute of Technology*, no. 2011-2 (July), www.me.titech.ac.jp/technicalreport/h23/2011-2.pdf [accessed February 23, 2012].
- Vastenhouw, Nadine L., Karin Brunschwig, Kristy L. Okihara, Fritz Muller, Marcel Tijsterman, and Ronald H. A. Plasterk. 2006. Gene expression: Long-term gene silencing by RNAi. *Nature* 442: 882.
- von Wright, Georg Henrik. 1971. *Explanation and understanding*. New York: Cornell University Press.
- Waddington, Conrad H. 1977. *Tools for thought*. London: Jonathan Cape.
- Walsh, J. Bruce. 1996. The emperor's new genes. *Evolution* 50: 2115–2118.
- Waters, C. Kenneth. 2007. Causes that make a difference. *Journal of Philosophy* 104: 551–579.
- Woodward, James. 2001. Law and explanation in biology: Invariance is the kind of stability that matters. *Philosophy of Science* 68: 1–20.
- . 2002. What is a mechanism? A counterfactual account. *Proceedings of the Philosophy of Science Association* 2002: S366–S377.
- . 2003. *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- . 2004. Counterfactuals and causal explanation. *International Studies in the Philosophy of Science* 18: 41–72.
- . 2007. Causation with a human face. In *Causation, physics, and the constitution of reality*, ed. Huw Price and Richard Corry, 66–105. Oxford: Oxford University Press.
- . 2010. Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology and Philosophy* 25: 287–318.
- Woodward, James, and Christopher Hitchcock. 2003. Explanatory generalizations, part I: A counterfactual account. *Noûs* 37: 1–24.

JAN BAEDKE studied philosophy and biology at Ruhr University Bochum. He works as a PhD fellow of the Ruhr University Research School (supported by the German Research Foundation, DFG) at Bochum University. His research interests include the philosophy and history of biology, causal reasoning and methodology. He currently publishes some papers on pictures in the sciences as well.

ADDRESS: Department of Philosophy, Ruhr University Bochum, Universitätsstr. 150, 4801 Bochum, Germany. E-mail: jan.baedke@rub.de