⋆

EnviroGenomarkers: the interplay between mechanisms and difference making in establishing causal claims

⋆

Federica Russo[a] & Jon Williamson[b]
[a] Center Leo Apostel, Vrije Universiteit Brussel &
Centre for Reasoning, University of Kent
[b] Philosophy, University of Kent

October 23, 2012

Abstract

According to Russo and Williamson (2007, 2011a,b), in order to establish a causal claim of the form 'C is a cause of E', one typically needs evidence that there is an underlying mechanism between C and E as well as evidence that C makes a difference to E. This thesis has been used to argue that hierarchies of evidence, as championed by evidence-based movements, tend to give primacy to evidence of difference making over evidence of mechanisms, and are flawed because the two sorts of evidence are required and they should be treated on a par.

An alternative approach gives primacy to evidence of mechanism over evidence of difference-making. In this paper we argue that this alternative approach is equally flawed, again because both sorts of evidence need to be treated on a par. As an illustration of this parity we explain how scientists working in the 'EnviroGenomarkers' project constantly make use of the two evidential components in a dynamic and intertwined way. We argue that such an interplay is needed not only for causal assessment but also for policy purposes.

§1 Introduction

Russo and Williamson (2007, 2011a,b) argued that, in the health sciences, in order to establish a causal claim one normally needs to establish both that the putative cause makes a difference to the putative effect (e.g., it raises the probability of the effect, conditional on states of the effect's other direct causes) and that there exists an underlying mechanism linking the putative cause to the putative effect that can explain this difference making. Evidence of difference making is required because causal claims are used for prediction and control, and one can only predict an effect on the basis of the cause, or control the effect by manipulating the cause, if the cause makes a difference to the effect. Evidence that there is an underlying mechanism is required because causal claims are used to explain, but in order to explain some phenomenon one needs to point to the (functioning of the) mechanism responsible

for it; so invoking a cause as an explanation for an effect is only successful to the extent that the cause is a part of the mechanism responsible for the effect. Evidence of mechanisms is also useful to show that the difference-making relationship is not spurious, to rule out potential confounders, and to extrapolate causal claims to new populations or individuals.

It is worth making two points clear from the very start. First, we need evidence that *there is* a plausible mechanism, but we do not necessarily need to know the mechanism in any detail. In practice, there are cases in which very strong statistical evidence—which is ostensibly evidence of difference making rather than of mechanisms—can licence some action (e.g., a public health action) even though the detailed mechanism has not been worked out yet. An example of this is the relation between asbestos exposure and cancer, where there is good statistical evidence but by no means a full description of the mechanism of disease causation. However, in this case we do have evidence *that there is* a mechanism. We know that the statistical correlation is unlikely to be spurious or confounded, we know a lot about the toxic properties of asbestos, and we know about analogous causal mechanisms such as that between smoking and lung cancer; all this evidence makes the existence of an underlying mechanism between asbestos and cancer sufficiently plausible to warrant the causal claim and hence also the corresponding public policy interventions.

Second, difference making and mechanisms are *normally* required, but they are not necessary and sufficient conditions for causality. This is in line with Bradford Hill's guidelines for causal assessment, where his nine indicators of causality were meant to be neither necessary nor sufficient for causality (for a discussion, see, e.g., Russo and Williamson (2011a)). Difference making is not required in cases where no difference can be made—such as when the effect is bound to occur anyway, so the cause cannot raise its probability any further. Mechanisms are not required in cases where there is nothing that can be linked by a mechanism—such as when the cause and/or the effect is an absence.

This epistemological and methodological thesis—i.e., the thesis that one normally needs evidence of both difference making and the existence of an underlying mechanism in order to establish a causal claim—has become known as the Russo-Williamson Thesis, or RWT, and has generated some controversy (Weber, 2009; Broadbent, 2011; Campaner, 2011; Clarke, 2011; Howick, 2011; Illari, 2011a; Darby and Williamson, 2011; Gillies, 2011; Russo and Williamson, 2011a,b). The main consequence of RWT is that neither sort of evidence—difference-making nor mechanistic—has primacy over the other. This goes against standard accounts in philosophy of causality, which tend to give primacy to one or other evidential component. This also goes against the diktats of the evidence-based movements in medicine and public policy, which tend to place difference-making evidence above mechanistic evidence in their evidence hierarchies (Russo and Williamson, 2011a). But, as we shall see in this paper, one cannot simply respond by turning the evidence hierarchies on their heads, treating mechanistic evidence as superior to difference-making evidence.

In this paper, we shall investigate a case study that both supports RWT and illustrates the parity of the two kinds of evidence. The project 'Genomic Biomarkers of Environmental Health'—or, for short, 'EnviroGenomarkers'—investigates the effects of environmental agents on a number of diseases, looking at biomarkers of exposure and of disease. The EnviroGenomarkers methodology establishes relationships on the basis of a subtle interplay between difference-making and mechanistic evidence, neither of which trumps the other. In §2 we introduce the EnviroGeno-

markers project and its methodology. In §3 we review two key mechanistic views of causality, the process-tracing approach and the complex-systems approach, showing how elements of both are involved in the mechanisms studied by EnviroGenomarkers. In §4 we argue that difference making is required in addition to mechanisms in order to understand EnviroGenomarkers.

§2 EnviroGenomarkers

§2.1. The project

EnviroGenomarkers is a project investigating the effects of environmental exposure on various diseases by using -omic technologies and biomarkers. We will argue in §4 that the methodology used in EnviroGenomarkers to establish causal claims conforms with our suggested interplay between difference making and mechanisms.

EnviroGenomarkers[1] is a European FP7 network with eleven partners from six European countries: National Hellenic Research Foundation, Greece; University of Maastricht, Netherlands; Imperial College London, United Kingdom; Umeå University, Sweden; Istituto per lo Studio e la Prevenzione Oncologica, Italy; University of Crete, Greece; University of Utrecht, Netherlands; Istituto Superiore di Sanità, Italy; National Public Health Institute (KTL), Finland H.; University of Leeds, United Kingdom; Lund University, Sweden. Scientists working in the project are studying the role of environmental agents in breast cancer and Non-Hodgkin's lymphoma, and in childhood diseases including allergy, neurological and immune diseases, and thyroid disruption. The underlying idea is to measure the effects of environmental agents through the evolution of biomarkers that predict the increased risk of the aforementioned diseases.

There are a number of reasons why EnviroGenomarkers is of potential interest to researchers interested in causality (scientists and philosophers alike). To begin with, the science is happening right now; consequently, the project gives an insight into the epistemology of causality that can otherwise be masked by a retrospective, historical approach to scientific discovery. In particular, EnviroGenomarkers is an excellent test case for RWT, which is an attempt to make sense of how contemporary science establishes causal relationships.

Another reason why EnviroGenomarkers is of interest is that it bridges 'levels', as it investigates both the macro (environmental) and the micro (biomarker) level. EnviroGenomarkers tries to solve the problem of measuring the influence of the environment onto the molecular level. The importance of environmental factors has long been established in epidemiology. Yet the question is still open as to how exactly quantify and explain the effect of environmental agents on disease. Rappaport and Smith (2010) note that, typically, scientists concentrate *separately* on the various categories of environmental exposure (e.g., air and water pollution, dietary habits and obesity, stress and behaviour, types of infection, ...). Rappaport and Smith argue that this is the wrong approach: different exposure categories should be studied together rather than separately, and to understand the action of environmental exposure we have to change our concept of environment. They suggest considering also *the body* as the environment, where various active chemicals act as exposures. They thus coin the term 'exposome' to refer to the totality of environmental expo-

[1] See the official website of the project: http://www.envirogenomarkers.net/.

sures, the ones coming from 'outside' and from 'inside', so to speak. The concept of exposome is also core to EnviroGenomarkers, as we shall see below.

Finally, EnviroGenomarkers uses evidence coming from innovative '-omic' technologies, which are thought to promise the missing links between environmental exposure and disease. -Omic technologies study complete sets of biological molecules, instead of a single biological structure (such as a protein or gene) which is the approach of traditional molecular biology. Such technologies allow researchers to detect changes in metabolism or gene expression of cells or tissues in response to exposure to some agent or class of agents by studying, for instance, gene expression profiling (transcriptomics), epigenetic changes in DNA (epigenomics), or the metabolites in a specified biological sample (metabolomics). Thus, an interesting question is: what evidence are -omic technologies in fact able to generate and, how is this evidence to be used to establish causal claims? We shall attempt to answer this question in §5, in the light of the preceding arguments of this paper.

The methodology of the project can be condensed in the phrase 'meeting in the middle', which was first put forward by Vineis and Perera (2007) and subsequently developed by Chadeau-Hyam et al. (2011). Simply put, 'meeting in the middle' means finding biomarkers of exposure and biomarkers of disease outcome, and *then* finding the 'intermediate' biomarkers that link exposure and disease, which are located in the middle of the causal network from exposure to disease.

'Meeting-in-the-middle' involves combining the results of prospective and retrospective studies. From prospective studies, scientists extract information about preclinical biomarkers related to particular exposures. From retrospective studies, they extract information that backtracks from clinical disease to preclinical response to exposure. Then they try to find the overlap, that is, those biomarkers that are good predictors of disease and that are associated with exposure. For instance, the aforementioned work by Chadeau-Hyam et al. (2011) describes a pilot study using data from the European Prospective Investigation into Cancer and Nutrition (EPIC). Researchers compared spectra of plasma samples from 24 cases of colon cancer cases, and 19 cases of breast cancer against 43 controls. Those plasma samples were collected on average 7 years before appearance of cancer. The comparison between cases and controls allowed researchers to identify a putative list of intermediate biomarkers linking exposure and disease.

The project is original in that it offers a new perspective on biomarkers. Biomarkers are not analysed synchronically but diachronically. This involves drawing the epidemic curve of disease and tracing the evolution of early biomarkers of exposure until disease develops. This allows models of infectious diseases to be extended to chronic diseases. For a discussion, see Galea et al. (2010); Vineis and Chadeau-Hyam (2011).

It is worth noting that EnviroGenomarkers is both exploratory and confirmatory. On the one hand, the project aims to discover *new* biomarkers for disease. On the other hand, the project also aims to validate the results of existing studies. For instance, Chadeau-Hyam et al. (2011) mention that metabolic profiling has been carried out using urine samples, whilst a number of epidemiological cohort studies collected blood samples. It is an important question to establish the extent to which studies using different biological specimens lead to coherent results.

§2.2. The interpretation of the project

Clearly, EnviroGenomarkers attempts to establish causal claims linking environmental exposure and certain diseases. It is less clear, however, what such causal claims would amount to. There is in fact a vigorous debate among philosophers of science about the meaning of causal claims and about the evidence required to establish them. We lack space to present the debate in detail and we refer the reader to Russo and Williamson (2007, 2011a,b) for a thorough discussion.

We shall just note that the literature is polarised around two main ideas. One is that causation is a matter of difference making; the other is that causation is matter of mechanisms. Simply put, according to the first view, to establish whether a chemical, say benzene, is carcinogenic we need to know whether exposure makes a difference to cancer rates; for instance, statistical analyses of data may reveal that individuals exposed to benzene have a higher risk of cancer. According to the second view, instead, to establish the same causal relationship we need to know the mechanism linking benzene exposure and cancer that can explain occurrences of cancer in terms of exposure to benzene.

This question of how to interpret causal claims equally applies to EnviroGenomarkers. There are two prima facie options. The first one emphasises the idea of finding biomarkers that are good predictors of disease. The second one emphasises instead the idea of tracing the evolution of early, pre-clinical biomarkers of exposure until the development of disease. We shall now lay out these two options in turn, focusing on the latter option in the remainder of this paper.

¶ *Biomarkers as good predictors of disease.* According to this interpretation, in EnviroGenomarkers scientists are after *chains of difference making*: they hunt for those biomarkers that are good predictors of disease. This would mean that difference making is *sufficient* for their purposes: all that is needed is the identification of the chain of difference-making relations from exposure to disease via intermediate biomarkers. We would argue that this is a *mis*interpretation of EnviroGenomarkers.

In fact, Vineis and Perera (2007) see studies in molecular epidemiology as providing evidence of *mechanisms*—evidence which is required by, e.g., the International Agency for Research on Cancer (IARC), but that has not been provided by traditional epidemiological studies. They say:

> When combined with the best of the earlier validated biomarkers of dose, effect, and susceptibility, such new markers have the potential to add considerably to knowledge about the *mechanistic pathways* that relate pathogenic exposures to disease onset and also to serve as informative early markers of disease risk. [Vineis and Perera (2007, p. 1955), emphasis ours.]

That EnviroGenomarkers cannot be interpreted solely in terms of difference making serves to illustrate RWT: the EnviroGenomarkers project is concerned with determining causality, and, as RWT makes clear, evidence of difference making *alone* cannot establish causality.[2] Vineis and Perera (2007, p. 1961) are explicit that biomarkers have to be in some causal pathway from exposure to disease:

[2]Although, as we mentioned earlier, when commenting on the available evidence linking asbestos exposure and cancer, very strong statistical evidence can be enough, in conjunction with background knowledge of analogous mechanisms, to establish the existence of a mechanism as well as of difference making, and hence may licence the causal claim and its corresponding public health interventions.

> One of the main challenges with intermediate biomarkers is to understand whether they belong to the causal pathway between exposure and disease, whether they are simply a side effect of exposure or disease, or whether their measurement is confounded by some other exposure. For example, it is likely that certain mutations are genuine intermediate markers in the causal pathway, whereas others are a consequence of the disease, such as genomic instability that arises in cancer cells.

In response, one might accept that mechanistic evidence is required in conjunction with difference-making evidence to establish causal claims, but one might argue that evidence of difference making is sufficient to determine a mechanism. This move would be available to those such as Glymour and Cheng (1998), who conceive of mechanisms simply as chains of difference making.

We would not want to deny that evidence of difference making—produced, for example, by a series of well conducted randomised controlled trials—is sometimes sufficient to *make plausible* the existence of a corresponding mechanism. However, it is rarely enough on its own to *establish* the existence of an underlying mechanism: for that, theoretical knowledge of parts of a putative mechanism, or of analogous mechanisms, is normally also required. Moreover, not all mechanisms are accompanied by difference making—for examples see Williamson (2005, §7.3), Williamson (2009, §10) and the discussion of gene knock-out experiments in §4 below. Hence evidence of difference making alone is in general insufficient to establish a causal claim. It is thus for good reason that EnviroGenomarkers seeks more than chains of difference making.

It is worth reiterating that we do *not* deny that difference making plays a role in the discovery and identification of mechanisms. Rather, we deny that difference making is all there is to causal assessment. Likewise, we do not deny that difference making plays a role in the discovery and validation of biomarkers, but that difference making alone is sufficient for these tasks (see also below §4).

¶ *Tracing the evolution of biomarkers.* According to this interpretation, scientists working in EnviroGenomarkers are concerned with tracing the process leading from exposure to disease through signals of the biomarkers. Thus Vineis et al. (2009) draw an analogy between low-dose environmental exposures and clinical vulnerability on the one hand, and the evolution of biomarkers of disease on the other hand. Let us explain further.

One problem with assessing environmental exposures is that doses are almost always low. However, this does not mean that they have no effect. On the contrary, environmental exposures have serious effects, but in the *long run*, namely when coping mechanisms of the body are unable to counteract the effects of environmental exposures, thereby leading to a change in clinical state. Thus, to estimate the long-term effects of low-dose environmental exposures one needs to trace the evolution of vulnerability and exposure events until clinical manifestations appear. The same idea can be applied to biomarkers of exposure and of disease:

> The concept of acquired "clinical vulnerability" is related to previous insults/pathophysiological changes that predispose to disease. Intermediate markers and specifically '-omics' could be particularly useful in tracing the "history" of such insults and in reflecting the cumulative effect of different exposures. (Vineis et al., 2009)

This strongly resembles, at least prima facie, the Salmon-Dowe process-tracing approach to causality outlined below (§3.1). Simply put, the process-tracing approach identifies (physical) processes or bio-chemical chains—i.e., certain kinds of physical mechanism—as constitutive of causal relations.

The question therefore arises: does evidence of such mechanisms suffice to establish causal relations? We will argue not. In the remainder of the paper, we will argue that we need an interplay between the two evidential components—difference-making and mechanisms—both for causal assessment and for policy purposes.

After providing an introduction to the mechanistic approaches to causality in §3, we shall argue in §4 that we need difference making as well as mechanisms to interpret the causal claims of EnviroGenomarkers.

§3
Mechanistic approaches

Mechanistic approaches to causality hold that, loosely speaking, causality is to be analysed in terms of some physical connection between the cause and the effect. These approaches come in two main variants. *Process-tracing* approaches understand the connection as a process described in terms of the low-level physical quantities involved. *Complex-systems* approaches understand the connection in terms of the complex organisation and activities of different entities. These two kinds of approach appeal to different understandings of mechanisms, and in turn these different notions of mechanisms reflect distinct ways of theorising about causal relations.

§3.1. Process-tracing approaches

Process-tracing is the view according to which causation is cashed out in terms of physical processes possessing certain characteristics—to be specified—that make them causal. That is to say, in this approach A causes B just in case there is an appropriate kind of physical process linking A to B. The crucial problem is that some physical processes may not be causal. For instance, our intuition is that billiard balls moving and colliding constitute causal processes, whilst aeroplanes' shadows crossing on the ground do not.

Process-tracing has a long tradition in the philosophy of causality. In this section, we sketch its developments from the first discussions of Russell (1913, 1948) up to the most recent formulation of Boniolo et al. (2011). This view was very popular in the Eighties and Nineties, especially thanks to the influential works of Salmon (1984, 1997) and of Dowe (1992, 2000), whence the label 'Salmon-Dowe process view', customarily used in the literature. In the rest of the paper, we shall also refer to 'Salmon-Dowe' for convenience.

Before starting the round up of process-tracing approaches, it is worth noticing that process-tracing theorisers had physics in mind. This is important because some scholars argue that physical processes do not exhaust the meaning of causation in, e.g., biology (see for instance the position of Machamer et al. (2000) also discussed below) or social science (see for instance Russo (2009)) and some others advocate pluralism on the grounds that different concepts of causality suit different scientific contexts (see for instance Weber (2007)). Consequently, although process-tracing is a viable interpretation of EnviroGenomarkers *prima facie*, this hypothesis deserves closer investigation, which we undertake in the remainder of the paper.

We here list the main developments of the process-tracing view, in chronological order. In so doing, we avoid any technical details, which can be found in Williamson (2011, §2), and just concentrate on the core ideas. The difference between the various process-tracing accounts goes beyond terminological variations. The terms chosen by advocates of process tracing are meant to grasp different aspects of physical reality that are key to understand what a causal process is. These differences turn out not to be directly relevant to our argument; notwithstanding these differences, the persisting idea is tracing the evolution of a physical process.

¶ *Causal lines.* Despite the attack on causality, in his famous paper 'On the notion of cause', Russell argued that the metaphysically loaded notion of causation could be explicated using the notion of 'causal lines', that is space-time trajectories that persist in isolation from other things (Russell, 1913, 1948).

¶ *Mark method.* Reichenbach was interested in explaining the asymmetry of time by appealing to the asymmetry of causality. His core idea was that if a causal process is marked at the beginning, the mark would be found at the end of the process, but not vice versa. This meant, in his view, that causal processes are those processes in which the mark propagates from the beginning to the end (Reichenbach, 1956).

¶ *Processes and conserved quantities.* The 'combined' Salmon-Dowe view states that processes are world lines of objects, and *causal* processes are those that transmit conserved quantities (e.g., mass-energy, linear momentum, or charge) after an interaction between two (causal) processes (Salmon, 1984, 1997; Dowe, 1992, 2000).

¶ *Processes and extensive quantities.* This recent account is a follow-up of the Salmon-Dowe account. It is different in that it holds that to discriminate between causal and non-causal processes, one need to appeal to the transmission of extensive quantities, not conserved quantities. This approach is able to account for causation in stationary cases, which the Salmon-Dowe approach could not do (Boniolo et al., 2011).

Despite the visible differences between the aforementioned versions of process tracing, there is a constant thread: causal processes are *physical* processes that can be traced by employing the Reichenbachian mark method or by identifying the transmission of conserved or extensive quantities in the later development of Salmon-Dowe and Boniolo et al.

The question then arises as to how this mechanistic approach to causality is related to difference-making accounts. The answer is that it depends quite a lot on the particular version of process tracing. In Reichenbach's and in the early Salmon's approach, causal processes involved mark transmission, which was given a difference-making, counterfactual account: simply put, a physical process is deemed causal if, *were* it to be marked, that mark *would* be propagated along the process. However, the late Salmon's and Dowe's approaches wanted to eradicate the counterfactual aspect; thus they abandoned the mark transmission criterion and instead appealed to the possession of conserved quantities. Causal processes did not have a counterfactual characterisation any longer. However, causal interactions were still supposed to make a difference to the conserved quantities possessed by the interacting causal processes.

§3.2. Complex-systems mechanisms

The process-tracing approach was specifically developed to capture causation in physical contexts. The proponents of *complex-system mechanisms*, often called 'mechanistas', wanted instead to develop an account of causation more suitable to other sciences such as biology. A notable example is Machamer et al. (2000, p. 7):

> Although we acknowledge the possibility that Salmon's analysis may be all there is to certain fundamental types of interactions in physics, his analysis is silent as to the character of the productivity in the activities investigated by many other sciences. Mere talk of transmission of a mark or exchange of a conserved quantity does not exhaust what these scientists know about productive activities and about how activities effect regular changes in mechanisms.

Whence the need, according to them, for a *complex* mechanism. The three main contending definitions of mechanisms are the following:

¶ *Machamer, Darden and Craver* "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions." (Machamer et al., 2000, p. 3)

¶ *Glennan* "A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations." (Glennan, 2002, p. S344)

¶ *Bechtel & Abrahamsen* "A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena." (Bechtel and Abrahamsen, 2005, p.423)

Illari and Williamson (2012) propose the following as a potential consensus definition that ought to be acceptable to all proponents:

> A mechanism for a phenomenon consists of entities and activities organised in such a way that they are responsible for the phenomenon.

Here again, the differences between these positions are not essential to our argument about the constant interplay between difference making and mechanisms in EnviroGenomarkers, but let us explain these positions a bit further.

Some scholars insist more on the 'organisational' aspects of mechanisms to emphasise that mechanisms have a structure and that it is the structure that has explanatory power. Others focus on their 'elements', on the grounds that mechanisms with no entities would eventually collapse into Salmon-Dowe processes or world lines.

Illari and Williamson relax some of the requirements of the above-mentioned definitions (notably, regularity, complex systems, and start-finishing conditions) and offer a definition that grasps the essential aspects of mechanisms. Talk of the mechanism being 'responsible' for its associated phenomenon is not normally intended to invoke a causal notion of responsibility. Rather, the key point is that a mechanism

explains the phenomenon, and it explains the phenomenon in virtue of the nature of the parts and activities and their organisation. Indeed, the mechanism and the phenomenon that it explains may not be the sort of thing that can stand in a causal relation, since causal relations are typically taken to relate spatio-temporally disjoint events; neither is a mechanism an event nor need its associated phenomenon be disjoint from it.

Stock examples used in the literature to illustrate these definitions are the mechanism of protein synthesis and the mechanism of circadian rhythms. What is clear is that in such cases a 'simple' physical process explicated in terms of world lines à la Salmon-Dowe will not do justice to the complex architecture of entities and activities that is doing the work. We agree with this.

According to the complex-systems approach, it is a mechanism in the above sense—as opposed to a Salmon-Dowe process—that provides the crucial connection between cause and effect. On this view, one event causes another if, and only if, they are linked by a complex-system mechanism that accounts for the putative effect by invoking the putative cause. If this view were correct, in order to establish a causal claim it would suffice to establish that the putative cause and effect are linked by an appropriate complex-systems mechanism. In contrast, in our view causal assessment needs evidence of difference making alongside evidence of mechanisms. So complex-systems approaches are prone to neglect, by and large, the crucial role of difference-making evidence.

¶ One can argue that there isn't such a big divide between complex-system mechanisms and process tracing. The reason is that complex-system mechanisms and Salmon-Dowe processes share at least the goal of cashing out causation in terms of some kind of physical link between the cause and the effect (Williamson, 2011). Illari (2011b) makes the point that they are all, albeit in slightly different ways, accounts of *production*. An account of causal *production* says how the cause, in given circumstances, produces the effect. Thus, process-tracing approaches explain causal production by pointing to physical processes and their interactions. When two billiard balls collide, the collision (the cause) produces a change in the direction of the trajectories of the balls (effect). Complex-systems approaches explain causal production by pointing to the organisation of the entities and activities involved in the mechanism. In this sense they don't need to be in sharp contraposition.

According to Illari (2011b), an *information-transmission* account of production is able to reconcile Salmon-Dowe processes with complex-system mechanisms. Briefly put, Illari's view is that information transmission provides a general account of production: causal production is transfer of information from the cause to the effect. This can be linked to complex-systems approaches. Mechanisms, in Illari's view, ought to be seen as information channels in which the transfer of information does the 'production' job. Difference making does other jobs, such as guiding the choice of the information channels that interest us.

While there isn't such a big divide from the perspective of being *accounts of production*, in the disciplines where they have been originally developed (physics and biology), the two accounts do capture different notions of mechanism.

§3.3. Mechanisms in EnviroGenomarkers

Interestingly, in order to capture the mechanisms of EnviroGenomarkers one needs to appeal to both Salmon-Dowe processes and complex-system mechanisms, as we

shall now explain.

The process from exposure to disease is conceptualised, in EnviroGenomarkers, in terms reminiscent of a Salmon-Dowe process:

> The ultimate goal of using "-omics" technologies to identify environmental causes of disease is to derive an integrated view of the biological processes involved in the continuum from exposure to disease. (Vineis et al., 2009).

Those familiar with the causality literature would recognise the idea of one-off processes of the kind of Billy and Suzy throwing stones.[3] Instances of exposure, for example instances of ionising radiation reaching the human body, are in a sense analogous to instances of stone-throwing, and better understood in terms of one-off Salmon-Dowe processes rather than complex-systems mechanisms. This is because instances of exposure can be seen as world lines that carry and exchange conserved quantities, but there seems to be no obvious systematic *mechanism* of exposure, involving a stable arrangement of parts organised in such a way that, e.g., radiation reaches the body.

Yet, while the process leading from exposure to the body may resemble a Salmon-Dowe process, that is not the end of the story. Disease causation is much more complex than the one-off process leading from Billy throwing the stone at the bottle to its shattering. It usually takes many instances of exposure to cause disease—up to the moment in which a threshold of 'clinical vulnerability' is reached—and what goes on within the human body—involving the complex-systems mechanisms for cell metabolism, cell repair, cell death and so on—very much determines whether and when disease will occur.

On the other hand, neither are these complex-systems mechanisms the end of the story. When the complex-systems mechanisms for maintaining the integrity of the body fail, various processes are set in action that can lead ultimately to disease. These processes may be better understood as Salmon-Dowe processes than as mechanisms for disease, due to their unstable and irregular nature. Thus while one might perhaps say that a particular kind of cancer has a mechanism for tumour growth, it may be more natural to conceptualise a haemorrhage as a Salmon-Dowe process.

The general picture is then that repeated exposures to environmental agents (Salmon-Dowe processes) interact with regulatory (complex-systems) mechanisms for maintaining the integrity of the body; if eventually these regulatory mechanisms fail, further (complex-systems) mechanisms or (Salmon-Dowe) processes can be instigated which lead to disease.

¶ With this background, we can now ask the question of whether physical mechanisms, or for that matter bio-chemical mechanisms, suffice to establish causal relations in EnviroGenomarkers.

[3] This is a stock example in the philosophical literature, especially in discussions of causal overdetermination. For instance:

> Suzy and Billy, expert rock throwers, are engaged in a competition to see who can shatter a target bottle first. They both pick up rocks and throw them at the bottle, but Suzy throws hers a split second before Billy. Consequently Suzy's rock gets there first, shattering the bottle. ... Suzy's throw is a cause of the shattering, but Billy's is not. (Hall, 2004)

§4
Do mechanisms suffice?

We now tackle the question of whether one can interpret the claims of EnviroGenomarkers purely in terms of mechanisms in either of the above senses. Our answer is no: in Envirogenomarkers, the causal claims are made on the basis of evidence of difference making as well as evidence of mechanisms.

¶ *The interplay between mechanisms and difference making.* Scientists in EnviroGenomarkers are certainly interested in tracing mechanisms from exposure to disease via various intermediary biomarkers. Yet, a conclusion of type 'The evolution of biomarker X of exposure to environmental agent Y is evidence that Y is a cause of disease Z' is not *just* based on evidence involving a traced process, but rather on highly intertwined considerations about difference making and of theoretical plausibility. For instance, Chadeau-Hyam et al. (2011, p. 86) write:

> Taken together, our results suggest that meaningful relationships can be found using our data analysis strategy on metabolic profiling and are consistent with the epidemiological literature relating to colon cancer.

Here, while 'consistency with the epidemiological literature' provides theoretical plausibility which refers, inter alia, to plausible mechanisms explaining the correlations, the 'data analysis strategy' provides statistical evidence of difference making.

It is worth noting that within epidemiology two fields can be distinguished: descriptive and analytic epidemiology. Whilst the former is primarily concerned with finding the risks of disease and exposure for a given population, the latter is mainly concerned with testing hypotheses that *explain* risks of exposure and disease, that is with *how* and *why* disease spreads. In this sense, analytic epidemiology is concerned with investigating mechanisms of disease development. For a discussion, see Russo (2012).

The quest for causality brings these two fields together. On the one hand we invoke causes to explain disease. Hence there normally needs to be some underlying mechanism linking the cause and effect that can explain the effect in terms of the cause. On the other hand, we invoke causes to predict and control disease. Prediction and control is of course not possible unless the cause makes a difference to the effect. Hence the use of causal claims for explanation, prediction and control requires both mechanisms and difference making together. EnviroGenomarkers is engaged in determining causal relationships—i.e., in prediction and control as well as explanation—hence it brings together both descriptive and analytic epidemiology. Witness Vineis and Perera (2007, p.1955):

> When combined with the best of the earlier validated biomarkers of dose, effect, and susceptibility, such new markers have the potential to add considerably to knowledge about the mechanistic pathways that relate pathogenic exposures to disease onset and also to serve as informative early markers of disease risk.

The point is that there is a need to gain knowledge of both mechanistic pathways and predictors of disease at the same time, by finding biomarkers that are causal intermediaries between exposure and disease.

The need for 'theoretical plausibility' to back up difference-making considerations is very much in line with Bradford Hill's guidelines for causal assessment in medicine (Hill, 1965). Bradford Hill listed nine issues that ought to be considered. The following items concern evidence of mechanisms: temporality, theoretical plausibility, coherence, experimental evidence, analogy; whilst the following items concern evidence of difference making: strength of association, consistency, dose-response relationship, experimental evidence. Bradford Hill did not intend to formulate a check list, but rather an inventory of guidelines, none of which has the status of *sine qua non* condition. In other words, causal assessment needs a wise interplay of difference making and of mechanistic considerations (on this point, see Russo and Williamson (2011a)).

The idea of a 'wise interplay' between difference making and mechanisms can be supported by appealing to paradigmatic cases in the history of medicine (e.g., the discovery of *Helicobacter Pylori* causing gastric ulcer), to current medical practice (e.g., the procedures of the International Agency for Research on Cancer or various types of postmortem examination), and to theoretical considerations concerning the need of mechanisms for explanation and of difference making for prediction and control (Russo and Williamson, 2007, 2011a,b).

¶ *Public health policy.* One might ask whether mechanisms could be sufficient to set up effective public health policies. The answer is no. What is needed, again, is a wise interplay between difference-making and mechanistic considerations, in line with RWT.

Difference making and mechanisms are both needed to inform policy because they have roughly different roles: difference making provides information about what works for whom in what circumstances, while mechanisms tell us what paths to intervene upon. In practice, however, this distinction between the respective roles of difference making and mechanisms is rather blurred (for a discussion, see Russo (2011)). Russo (2012) tackles the role of difference making and of mechanisms in public health. In that paper, Russo points out that policy science is concerned with establishing the very basis of policy actions and that evidence-based policy has been developed in order to provide an answer to this need (see for instance Brownson et al. (2003, 1999) and Killoran and Kelly (2010)). Yet, evidence-based policy has left by and large unanswered the question of *what* evidence is in fact needed. Much discussion is devoted to the methods to *assess* evidence, but not *what* evidence has to be assessed. RWT says that causal assessment needs two evidential components: evidence of difference making and evidence of mechanisms. Russo (2012) argues that these two evidential components are needed for policy making too.

A good example is the MEND programme.[4] MEND is a public health programme, established in 2004, that aims to teach children and their families how to live healthier lives. MEND targets children in the age ranges of 2–4, 5–7, and 7–13, but also the parents of overweight or obese children, thus aiming to positively change their own and their children's habits concerning nutrition and lifestyle. In MEND one can report an obese child to the project officers and thus try to get the whole family involved in the programme. MEND proved to be quite successful in fighting obesity. Targeting the right groups and individuals, which requires having and using the right difference-making evidence, is part of the success. In other words, part of the success lies in a correct identification of the *relevant* causal variables, which is

[4] MEND stands for Mind, Exercise, Nutrition . . . Do it!. See http://www.mendprogramme.org/.

exactly the job of difference-making evidence. But the success of the MEND programme also depends on evidence of mechanisms: in the recognition that, to reduce child obesity, we may need to intervene on the child's eating behaviour and physical activity (biological factors) *and* on the parents of the child (socio-psychological factors).

In EnviroGenomarkers, difference-making information can tell one how to partition the population for different possible public health interventions. However, difference-making evidence produced in EnviroGenomarkers needs to be combined with further evidence for policy purposes. For instance, in the study by Chadeau-Hyam et al. (2011), metabolic profiling was used to analyse plasma samples in cases and controls in order to assess the possible effects of some dietary compounds. Researchers found out that metabolomic signatures were associated with colon cancer and dietary fibre intake (which is, according to other epidemiological studies, a protective factor). But how are such results to be used for public health purposes?

We need to reinterpret the information gathered from the studies on biomarkers in order to decide how to act to reduce the burden of disease. It would be ideal to figure out what portion of the population is most affected by colon cancer and what their dietary habits are in order to set up a targeted public health action. At the same time we can also act in a less targeted way and try to induce changes in dietary habits such that fibre intake is increased in the whole population. An example in this direction is the '5 a day' campaign sponsored by the UK National Health Service.[5] Actions like this aim to make people aware of the little changes that make a difference. Having five portions of fruit and vegetables a day makes diet much healthier and this contributes towards preventing a number of diseases.

¶ *Mechanisms without difference making.* Let us consider one further reason why evidence of difference making is important in addition to evidence of mechanisms. There are some causal relationships in which the cause makes no difference to the effect, although there is an underlying mechanism linking cause and effect. In the literature on causality, these cases are known as violations of the *faithfulness condition*, an assumption often made to facilitate causal modelling, which implies that, if one variable is a direct cause of another in a causal model then the two are probabilistically dependent conditional on the effect's other direct causes. One source of such cases is gene knock-out experiments (Steel, 2007, §4.4.2). In certain cases of genetic redundancy, the action of two genes can be mutually exclusive, but lead to the same effect—see, e.g., Scarff et al. (2004). Thus when one of the pair of genes fails, the other is expressed and the effect is caused via this back-up mechanism. The problem is that the effect depends on neither of these genes in isolation, since when one fails the other kicks in. This problem occurs generally when two back-up causes are mutually exclusive, and so no one of these causes can be held fixed to obtain a dependence between the other cause and the effect (Williamson, 2005, §7.3).

Now in cases such as these, there is no point in intervening on the cause (e.g., knocking out a gene) to prevent the effect, because the effect will happen anyway. Similarly, one cannot use the presence of the cause as a predictor of the presence of the effect, since the effect would be present in the absence of the cause. One cannot use causal claims for prediction and control if the cause makes no difference to the effect. In order to avoid these pathological cases, one needs to insist on evidence

[5] See the campaign website, where lots of information and tips are given (http://www.nhs.uk/Livewell/5ADAY/Pages/5ADAYhome.aspx).

of difference making before attempting to predict or control. Hence this provides another reason why one might demand evidence of difference making even when the underlying mechanisms are known.

Note that EnviroGenomarkers avoids this sort of pathological case because putative biomarkers that mediate between exposure and disease are chosen on the basis of their being difference makers. The question for EnviroGenomarkers is rather whether this difference making is substantiated by underlying mechanisms, that is whether a putative biomarker is a causal or spurious difference maker.

§5
Discussion and conclusion

In this paper we have used the EnviroGenomarkers project to illustrate the thesis that one needs evidence of difference making as well as evidence of mechanisms to achieve sound causal assessment and policy. These two evidential components are on a par—neither trumps the other in establishing causal claims or in setting up policy actions. We have also argued that different ways of theorising about mechanisms—process tracing and complex systems—are needed in order to interpret the mechanisms investigated by the EnviroGenomarkers project.

The advent of -omic technologies in molecular research has opened up new horizons for understanding the relations between environmental exposure and (some) diseases. The promise of -omic technologies has been stated as follows:

> Omic technologies offer great potential to identify biomarkers. (Vineis and Chadeau-Hyam, 2011)

> '-Omics' tools can be directly applied to samples from an epidemiologic case-control or cohort study to better characterise intermediate pathways, potentially providing the 'missing links' among exposures, genes, and diseases. (Vineis et al., 2009)

> Using these '-omics' technologies to inform hypothesis-directed pathway-based approaches to molecular epidemiology and to help direct genome-wide exploratory analyses into more promising directions. [...] one might think of the "-omics" data as providing the missing link among exposure, genes, and disease. (Thomas, 2006, p. 490)

The analysis of the methodology employed in EnviroGenomarkers suggests that -omic technologies can indeed generate evidence that there is a process tracing the evolution of a biomarker from exposure through the development of disease. Yet, -omic technologies are unable to generate evidence that, *alone*, is sufficient for causal assessment and for policy, for reasons discussed earlier in the paper. We pointed out that both causal assessment and policy need evidence of difference making *and* evidence of mechanisms. Evidence coming from -omic technologies does indeed help in both respects (difference making and mechanisms), but it does not exhaust the evidence needed for either task.

Consider causal assessment. As EnviroGenomarkers scientists have acknowledged, results coming from -omics analyses need to be substantiated by theoretical plausibility (i.e., existing knowledge of mechanisms at the molecular level). As for public health policy, we noted earlier that the identification of biomarkers may help

with identifying the correct groups to target in a policy intervention. However, results of studies on biomarkers *alone* are insufficient to warrant setting up policy actions.

The question remains open whether other more mechanistically oriented modelling can be used to model the relations between environmental exposure in EnviroGenomarkers. A possible candidate for such 'mechanistically oriented' modelling is the Recursive Bayesian Net (RBN) formalism (Casini et al., 2011). While Bayesian nets are often used to capture difference-making relationships, RBNs can also model the various levels of hierarchical organisation present in complex-systems mechanisms, to yield models that can be applied to explanation as well as prediction and control.

In sum, -omic technologies are particularly good at helping to generate hypotheses. The relations between environmental exposure and disease are, at the molecular level, yet to be fully understood. The promise of these technologies is high. At the same time, the gains, if projects such as EnviroGenomarkers prove successful, are high too. In any case, the EnviroGenomarkers project can shed some light on contemporary methods for causal inference, and serves as an interesting exemplar of a controversial philosophical thesis, RWT.

Acknowledgements. We thank the scientists that have been involved in the EnviroGenomarkers project, and especially Paolo Vineis and Marc Chadeau-Hyam, for spending time with us discussing the philosophical underpinnings of their project. We also thank Phyllis Illari and Lorenzo Casini for their comments, suggestions, and insights. Financial support from the British Academy and the UK Arts and Humanities Research Council is gratefully acknowledged. F. Russo is currently 'Pegasus Marie Curie Fellow' of the Research Foundation – Flanders.

References

Bechtel, W. and Abrahamsen, A. (2005). Explanation: a mechanist alternative. *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 36:421–441.

Boniolo, G., Faraldo, R., and Saggion, A. (2011). Explicating the notion of 'causation': the role of extensive quantities. In Illari, P., Russo, F., and Williamson, J., editors, *Causality in the sciences*, pages 502–525. Oxford University Press.

Broadbent, A. (2011). Inferring causation in epidemiology: mechanisms, black boxes, and contrasts. In Illari, P., Russo, F., and Williamson, J., editors, *Causality in the sciences*, pages 45–69. Oxford University Press, Oxford.

Brownson, R. C., Baker, E. A., Leet, T. L., and Gillespie, K. N. (2003). *Evidence-Based Public Health*. Oxford University Press, New York.

Brownson, R. C., Gurney, J. G., and Land, G. H. (1999). Evidence-based decision making in public health. *Journal of Public Health Management Practice*, 5(5):86–97.

Campaner, R. (2011). Understanding mechanisms in the health sciences. *Theoretical Medicine and Bioethics*, 32:5–17.

Casini, L., Illari, P. M., Russo, F., and Williamson, J. (2011). Models for prediction, explanation and control: recursive Bayesian networks. *Theoria*, 26(1):5–33.

Chadeau-Hyam, M., Athersuch, T. J., Keun, H. C., Iorio, M. D., Ebbels, T. M., Jenab, M., Sacerdote, C., Bruce, S. J., Holmes, E., and Vineis, P. (2011). Meeting-in-the-middle using metabolic profiling—a strategy for the identification of intermediate biomarkers in cohort studies. *Biomarkers*, 16(1):83–88.

Clarke, B. (2011). *Causality in medicine with particular reference to the viral causation of cancers*. PhD thesis, Department of Science and Technology Studies, University College London, London.

Darby, G. and Williamson, J. (2011). Imaging technology and the philosophy of causality. *Philosophy & Technology*, 24(2):115–136.

Dowe, P. (1992). Wesley Salmon's process theory of causality and the conserved quantity theory. *Philosophy of Science*, 59(2):195–216.

Dowe, P. (2000). Causality and explanation: review of Salmon. *British Journal for the Philosophy of Science*, 51:165–174.

Galea, S., Riddle, M., and Kaplan, G. A. (2010). Causal thinking and complex system approaches in epidemiology. *International Journal of Epidemiology*, 39:97–106.

Gillies, D. A. (2011). The Russo-Williamson thesis and the question of whether smoking causes heart disease. In Illari, P., Russo, F., and Williamson, J., editors, *Causality in the Sciences*, pages 110–125. Oxford University Press, Oxford.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69 (Supplement):S342–S353.

Glymour, C. and Cheng, P. W. (1998). Causal mechanism and probability: A normative approach. In Oaksford, M. and Chater, N., editors, *Rational Models of Cognition*. Oxford University Press, Oxford.

Hall, N. (2004). Two concepts of causation. In Collins, J., Hall, N., and Paul, L., editors, *Causation and counterfactuals*, pages 225–276. MIT Press, Cambridge MA and London.

Hill, B. (1965). The environment of disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58:295–300.

Howick, J. (2011). Exposing the vanities—and a qualified defence—of mechanistic evidence in clinical decision-making. *Philosophy of Science*, 78(5):926–940. Proceedings of the Biennial PSA 2010.

Illari, P. M. (2011a). Disambiguating the Russo-Williamson thesis. *International Studies in the Philosophy of Science*, 25(2):139–157.

Illari, P. M. (2011b). Why theories of causality need *Production*: an information transmission account. *Philosophy & Technology*, 24(2):95–114.

Illari, P. M. and Williamson, J. (2012). What is a mechanism? thinking about mechanisms *accross* the sciences. *European Journal for Philosophy of Science*, 2(1):119–135.

Killoran, A. and Kelly, M. P., editors (2010). *Evidence-based public health. Effectiveness and efficiency*. Oxford University Press, New York.

Machamer, P., Darden, L., and Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, 67:1–25.

Rappaport, S. M. and Smith, M. T. (2010). Environment and disease risks. *Science*, 330:460–461.

Reichenbach, H. (1956). *The direction of time*. University of California Press, Berkeley and Los Angeles, 1971 edition.

Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelian Society*, 13:1–26.

Russell, B. (1948). *Human knowledge*. George Allen & Unwin, London.

Russo, F. (2009). *Causality and causal modelling in the social sciences. Measuring variations*. Methodos Series. Springer, New York.

Russo, F. (2011). Causal webs in epidemiology. *Paradigmi*, XXXIX (1):67–98.

Russo, F. (2012). Public health policy, evidence, and causation. Lessons from the studies on obesity. *Medicine, Health Care, and Philosophy*, 15(2):141–151.

Russo, F. and Williamson, J. (2007). Interpreting causality in the health sciences.

International Studies in the Philosophy of Science, 21(2):157–170.

Russo, F. and Williamson, J. (2011a). Epistemic causality in medicine. *History and Philosophy of the Life Sciences*, 33:389–396.

Russo, F. and Williamson, J. (2011b). Generic versus single-case causality: the case of autopsy. *European Journal for Philosophy of Science*, 1(1):47–69.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press, Princeton NJ.

Salmon, W. C. (1997). Causality and explanation: a reply to two critiques. *Philosophy of Science*, 64(3):461–477.

Scarff, K. L., Ung, K. S., Nandurkar, H., Crack, P. J., Bird, C. H., and Bird, P. I. (2004). Targeted disruption of spi3/serpinb6 does not result in developmental or growth defects, leukocyte dysfunction, or susceptibility to stroke. *Molecular and Cellular Biology*, 24(9):4075–4082.

Steel, D. (2007). *Across the Boundaries, Extrapolation in Biology and Social Science*. Oxford University Press, Oxford.

Thomas, D. C. (2006). High-volume "-omics" technologies and the future of molecular epidemiology. *Epidemiology*, 17(5):490–491.

Vineis, P. and Chadeau-Hyam, M. (2011). Integrating biomarkers into molecular epidemiological studies. *Current Opinion in Oncology*, 23(1):100–105.

Vineis, P., Khan, A. E., Vlaanderen, J., and Vermeulen, R. (2009). The impact of new research technologies on our understanding of environmental causes of disease: the concept of clinical vulnerability. *Environmental Health*, 8(54).

Vineis, P. and Perera, F. (2007). Molecular epidemiology and biomarkers in etiologic cancer research: The new in light of the old. *Cancer Epidemiology, Biomarkers & Prevention*, 16(10):1954–1965.

Weber, E. (2007). Conceptual tools for causal analysis in the social sciences. In Russo, F. and Williamson, J., editors, *Causality and probability in the sciences*, pages 197–213. College Publications, London.

Weber, E. (2009). How probabilistic causation can account for the use of mechanistic evidence. *International Studies in the Philosophy of Science*, 23(3):277–295.

Williamson, J. (2005). *Bayesian nets and causality: philosophical and computational foundations*. Oxford University Press, Oxford.

Williamson, J. (2009). Probabilistic theories. In Beebee, H., Hitchcock, C., and Menzies, P., editors, *The Oxford Handbook of Causation*, pages 185–212. Oxford University Press, Oxford.

Williamson, J. (2011). Mechanistic theories of causality. *Philosophy Compass*, 6(6):421–447.