

Laws and Mechanisms in The Human Sciences

Rui Sampaio
Faculty of Human and Social Sciences
of the University of the Azores
rjsamsil@gmail.com

Abstract According to an influential epistemological tradition, science explains phenomena on the basis of laws, but the last two decades have witnessed a neo-mechanistic movement that emphasizes the fundamental role of mechanism-based explanations in science, which have the virtue of opening the “black box” of correlations and of providing a genuine understanding of the phenomena. Mechanisms enrich the empirical content of a theory by introducing a new set of variables, helping us to make causal inferences that are not possible on the basis of macro-level correlations (due to well-known problems regarding the underdetermination of causation by correlation). However, the appeal to mechanisms has also a methodological price. They are vulnerable to interference effects; they also face underdetermination problems, because the available evidence often allows different interpretations of the underlying structure of a correlation; they are strongly context-dependent and their individuation as causal patterns can be controversial; they present specific testability problems; finally, mechanism-based extrapolations can be misleading due to the local character of mechanisms. At any rate, the study of mechanisms is an indispensable part of the human sciences, and the problems that they raise can be controlled by quantitative and qualitative methods, and an epistemologically informed exercise of critical thinking.

Keywords Laws, Mechanisms, Causality, Reductionism, Human Sciences.

DOI 10.2478/kjps-2018-0004

According to an influential epistemological tradition, science explains phenomena on the basis of laws, but the last two decades have witnessed the emergence of a neo-mechanistic movement that emphasizes the fundamental role of mechanism-based explanations in science. This

epistemological turn results, on the one hand, from the limitations of a strictly nomological approach to science and, on the other hand, from a growing awareness of the methodological virtues of mechanistic explanations. In fact, a genuine scientific understanding of the phenomena cannot be reduced to lawlike regularities or “black boxes” of unexplained correlations; it requires something more. In Elster’s words: “To explain is to provide a mechanism, to open up the black box and show the nuts and bolts, the cogs and wheels of the internal machinery. (...) A mechanism provides a continuous and contiguous chain of causal or intentional links.”¹ The mechanical image used in this passage should not mislead the reader; the key point of a mechanistic explanation is the decomposition of a phenomenon into its constitutive components and respective activities, and many different types of entities can be considered components of a mechanism. For instance, imitation is a fundamental mechanism in the process of socialization, whose explanation involves reference to social and psychological entities.

The consideration of mechanisms in science is not, in itself, a novelty. The new mechanist turn in philosophy of science has some historical antecedents both in the ancient philosophy (as illustrated by atomism) and in early modern philosophy, where the likes of Galileo, Descartes, Gassendi and Boyle proposed, in opposition to the Aristotelian ontology and science, a new scientific program motivated by the intention to explain the natural phenomena (including biological phenomena) on the basis of matter and motion alone, thereby avoiding an appeal to obscure forms or essences. There are, however, differences between the new and the early modern mechanistic philosophy. First of all, the new mechanists are motivated by a dissatisfaction with the epistemological inheritance of logical empiricism, in particular with its account of law-based account of scientific explanation, which contrasts with their preference for local and particular mechanisms, whereas some important variants of the early modern mechanistic philosophy were committed to nomological explanations.² Glennan, one of the most prominent advocates of

1 Elster, 1983b, 24.

2 Roux (2018, 29) distinguishes in this context Descartes’s choice of nomological explanations from Gassendi’s “corpuscular” explanations, where “the burden of explanation

the mechanistic turn in philosophy of science, points out two other differences between the early modern and contemporary mechanists: the “New Mechanists are not committed to atomism either metaphysically or methodologically” and recognize the existence of important differences between mechanisms and machines.³

At any rate, according to the dominant view in the modern philosophy of science, scientific explanations rely in an essential way on laws of nature, but in domains like biology or psychology, where the study of mechanisms is traditionally very important, this nomological account of science is unsatisfactory. The situation in biology was clearly described by Francis Crick:

“Biology has its “laws”, such as those of Mendelian genetics, but they are often only rather broad generalizations, with significant exceptions of them. The laws of physics, it is believed, are the same everywhere in the universe. This is unlikely to be true in biology. (...) What is found in biology is mechanisms”⁴.

In a similar vein, Bechtel and Wright claimed that “when psychologists (...) offer explanations that go beyond the empirical laws or effects they identify, they frequently suggest that such explanations model a *mechanism*.”⁵ The reference to ‘effects’ in this context is explicitly inspired by Cummins, who had emphasized that psychological generalizations are usually called effects, not laws, a terminological tendency that reflects their lack of explanatory power.⁶

In what follows, I will start precisely by stressing some limitations of a nomological account of scientific explanation (1) and show the nature and methodological relevance of mechanism-based explanations (2), as

lies on the properties of insensible magnetic corpuscles.”

3 Glennan, 2017, 6-7.

4 Crick, 1988, 138.

5 Bechtel and Wright, 2009, 119. The authors also stress the distinctive character of many psychological mechanisms as information-processing or behaviour-regulating mechanisms.

6 Cf. Cummins, 2000.

well their own limitations or critical problems (3). Finally, I will argue that, first appearances notwithstanding, mechanistic explanations are not committed to epistemological reductionism or, in the case of the social sciences, to methodological individualism (4).

1. Limitations of Nomological Approaches to Science

As mentioned above, the idea that scientific explanations appeal to laws played a fundamental role in the history and philosophy of science, but it has not been easy to formulate a model of scientific explanation based on this idea. The most famous attempt is Hempel's "covering law" model of scientific explanations, according to which the *explanandum*, the phenomenon to be explained, should be inferred from two sets of premises (the *explanans*); premises expressing general laws and premises mentioning particular facts or antecedent conditions. If the laws contained in the premises were deterministic, the explanation would be called deductive-nomological; if the laws were statistical or probabilistic, the explanation would be either inductive-statistical or deductive-statistical.⁷ The model is elegant, but faces two serious difficulties. On the one hand, it does not preserve the asymmetric character of causal relations; according to a standard objection, Hempel's model allows us to explain the length of a flagpole's shadow using as premises the height of the flagpole and the angle of elevation of the sun or, alternatively, to explain the height of the flagpole using as premises the size of the shadow and the angle of elevation of the sun... Both explanations fit Hempel's model, but only the first one is a genuine explanation. A second source of difficulties lies in the possibility of situations where one can infer true conclusions from true, but explanatorily irrelevant premises, although they contain (true) general laws and (true) antecedent conditions.

Apart from technical problems of particular models, there are general limitations of a nomological account of science. To begin with, it is possible to give examples of good scientific explanations that do not mention laws; Darwin's theory of evolution is a case in point. The *mechanism* of

7 Hempel, 1965, 380 and ff.

genetic variation and differential reproduction is the key explanatory factor in Darwin's theory.

Probabilistic laws raise further explanatory problems, because they do not say that a phenomenon had to happen. Von Wright clearly expressed this point: "It seems to me better not to say (...) that the inductive-probabilistic model explains what happens, but to say only that it justifies certain expectations and predictions."⁸

If we consider another type of laws, the much discussed *ceteris paribus* laws, we face again serious difficulties. In complex domains, like those that study the human action and behaviour, our lawlike generalizations usually have exceptions, because of the interference of causal factors that cannot be specified in advance. Should we accept this particular type of laws; laws that are true in normal conditions ("everything being equal"), but admit of exceptions? Critics reject them, arguing that they are vacuous, vague and untestable.⁹ Vacuous, because to say that "*ceteris paribus*, all Fs are Gs" seems to be similar to saying that "all Fs are Gs, except when they are not"... Vague, because cp (*ceteris paribus*) clauses lack determinate truth conditions. Untestable, because failed predictions can be interpreted as exceptions covered by the cp clause. On the other hand, a supporter of cp laws may argue that they describe real tendencies (in a necessarily idealized way), that they can be tested in idealized settings, and that exceptions to nomological generalizations are acceptable, provided that they can be object of independent explanations.¹⁰ The latter point is particularly important, because it suggests a criterion to distinguish between testable and untestable cp laws. If unexplained exceptions to a law are simply dismissed with the help of cp clauses, one can legitimately conclude that the law in question is, as Popper would say, unfalsifiable; but if we can show how a familiar *mechanism*, for instance, generated a particular exception to a cp law, such an exception would not endanger the testability and refutability of these laws. As we see, the consideration of mechanisms may have the merit of solving disputes regarding cp laws.

8 Von Wright, 1971, 14.

9 For a critical view on cp laws, cf. Earman, Roberts and Smith, 2002.

10 For a defence of cp laws, cf. Kincaid, 1996 (chapter 3), and Pietrosky and Rey, 1995.

An additional set of problems arises from the correlational character of scientific laws, in the sense that they express correlations between two or more variables. A first problem can be labelled as “the black box problem”; one can know *that* a correlation holds without knowing *how* its variables are connected, but the mere awareness of the existence of a correlation is not an adequate explanation of a phenomenon. There is, for instance, a positive correlation between peace and democracy (democratic countries avoid wars among themselves), and such a correlation can even be mentioned in particular explanations, but it requires itself an explanation, a description of the mechanism that generates it.

A second well-known problem affecting correlational research is the underdetermination of causality by correlation. In order to show how fallible it is to infer a causal relation from a mere correlation, let us consider the following list of causal interpretations of an $x - y$ correlation:

1. x is the cause of y
2. y is the cause of x
3. There is a mutual causal relation between x and y
4. x and y are effects of z (a *common cause*)
5. x causes z (an *intermediate cause*), which causes y
6. The correlation is merely accidental.

There are sophisticated and useful statistical methods that help us to avoid mistakes in inferring causes from correlations, but they are not sufficient to fully guarantee the correctness of causal inferences from statistical data. In fact, statistical analysis may have difficulties in identifying spurious correlations or the so-called confounders, hidden variables that play an unobserved causal role. Equally relevant is the fact that causal inference requires often a local knowledge that it is not available in the form of statistical evidence. For instance, the causal factors that explain crime rates may differ across different cities of the same country as a reflex of (sometimes subtle) contextual differences. The explanation of particular distributions of traits in a population through natural selection is also highly context-sensitive. Furthermore, in inferring causal relations from statistical evidence one must bear in mind that different mechanisms might produce the same correlation. An illustration of this point is provided by Malinowski’s study on the correlation of *wealth*,

social status and *number of wives* in the Trobriand society.¹¹ This statistical correlation can receive many different causal interpretations, but only an ethnographic investigation could discover the underlying social mechanism: in that society, a high number of wives is a cause of wealth (because of the important gifts that men receive on the occasion of a wedding), and wealth is, in turn, a fundamental cause of social influence. Finally, it should be noted that two events may be linked by a causal connection in the absence of a statistical correlation (historical disciplines provide endless examples of this kind).

The limitations of a nomological/correlational approach to science have contributed to a renewed interest in the study of mechanisms, and it seems natural to complement these two approaches to scientific explanation. The need for this complementarity is well expressed in the so-called Russo-Williamson Thesis: “To establish causal claims, scientists need the mutual support of mechanisms and dependencies. (...) The idea is that probabilistic evidence needs to be accounted for by an underlying mechanism before the causal claim is established.”¹²

2. Mechanisms and Their Methodological Virtues

According to the preceding considerations, the study of mechanisms should be considered a key element in scientific explanations. But what is a mechanism? The following definitions are highly influential and often quoted:

“Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up conditions to finish or termination conditions”.¹³

“A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts,

11 Cf. Malinowsky, 1935.

12 Russo and Williamson, 2007, 159.

13 Machamer, Darden and Craver, 2000, 3.

where the interactions between parts can be characterized by direct, invariant, change-relating generalizations”.¹⁴

Decomposition (into “entities and activities” or “parts”) and organization (implicit in words like ‘interaction’ or ‘system’) are the keywords in the study of mechanisms. A mechanistic analysis may start with a functional decomposition,¹⁵ but the ultimate end of decomposition is to identify the entities and activities that constitute a mechanism.¹⁶ The final step in the analysis of a mechanism is the study of the organization of its entities and activities. In practice, however, the study of mechanisms requires what Craver calls “interlevel” integration. A mechanism can have lower-level mechanisms as its parts, and in these cases the decompositional analysis may go deeper. The inverse operation is also important. Sometimes, it is not possible to explain the functioning of a mechanism without integrating it into a higher-level mechanism or sets of mechanisms. Psychology, for instance, must be attentive to the social world and its causal impact on psychological mechanisms, and the study of the biological mechanisms of organisms may require a study of the ecosystem. It should be noted that the hierarchical, nested character of mechanisms provides a useful framework for the organization of interdisciplinary work in science.

Illari and Williamson propose the following account of mechanisms: “A mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon.”¹⁷ This seems a somewhat modest account, but such a modesty is justified by the authors with the intention to offer an account of mecha-

14 Glennan, 2002, S344.

15 In the case of memory, cognitive psychologists were able to decompose functionally the human memory (distinguishing, e.g., between long-term and short-term memory, or between semantic and episodic memory), without knowing the “cogs and wheels” of the workings of memory.

16 The decomposition of mechanisms must stop at some point: either at irreducible regularities, like the basic laws of physics, or where such a decomposition goes beyond the scope of our research interests. The objection that mechanistic explanations face the danger of infinite regress (formulated, for instance, by Norkus, 2005) does not constitute a real challenge.

17 Illari and Williamson, 2012, 120.

nisms that fits not only the biological and psychological domains (where the neo-mechanistic movement has its origins), but science in general. The distinction of entities and activities is preserved (which means that activities cannot be reduced to entities) and references to ‘systems’ or ‘structures’ are avoided in order to accommodate astrophysical mechanisms (like the production of a supernova). Illari and Williamson also drop the reference to “start or set-up conditions” and to “finish or termination conditions”, not only because there are cyclic mechanisms, but also because the identification of these conditions is often guided by pragmatic considerations.

There are strong reasons to accept mechanistic explanations in science as an indispensable complement to nomological explanations. In fact, mechanistic explanations have several relevant virtues. First, they open up the black box of correlations, avoiding the vice of “boxology”.¹⁸ Second, and in close connection with the first point, by exploring the internal structure of correlations, the study of mechanisms introduce new variables or factors, thereby enriching the empirical content of theories (their “explanatory grain”). Third, and precisely because the study of mechanisms offers more fine-grained explanations of the phenomena, they allow causal inferences that are not possible on the basis of macro-level correlations. Conversely, they can avoid hasty causal inferences, as stressed by Elster: “the better we focus the causal story, the easier it is to make sure that we are not dealing with mere correlation.”¹⁹ Fourth, stable mechanisms, like laws, can be used to express regularities. Finally, and as we have already seen, the study of mechanisms can legitimize cp laws by explaining away their exceptions.

In order to illustrate the virtues of mechanistic explanations (and to anticipate some problems that will be discussed in the next section), it is convenient to analyze briefly a particular case, the so-called democratic peace hypothesis, which is based on the already mentioned democracy↔peace correlation and claims, generically, that democracy has a positive causal influence on peace. The simple discovery of a correlation between democracy and peace is in itself insufficient do draw causal

18 Cf Craver and Darden, 2013, 90–91.

19 Elster, 2015, 25.

inferences, because the causal arrow can go, in principle, in both directions: it may be the case that democracy promotes peace, that peace promotes democracy or even that both hypotheses are true (the reciprocal causal arrow hypothesis). But one should also consider the hypothesis that peace and democracy are effects of a common cause, like economic development. In fact, it is not difficult to formulate hypotheses supporting an account of economy as a common cause of both democracy and peace: on the one hand, it is possible to invoke the thesis that a strong middle class is a social structure that facilitates democracy and include it in a mechanistic explanation of how economic development contributes to democracy; on the other hand, economic prosperity strengthens naturally the aversion to war that is characteristic of average citizens. But the introduction of economy as a relevant causal factor in the explanation of democratic peace generates at least two other causal hypotheses: economy and democracy exert a joint causal influence on conflict reduction (1); democracy promotes economic prosperity, which, in turn (as an intermediate cause) promotes peace (2).

As we see, the introduction of a new causal factor to explain a correlation has two relevant consequences. First, it creates more specific, detailed hypotheses that are more easily confronted with statistical evidence or even with experimental tests. Second, it can be a first step in a mechanistic explanation. In the democratic peace example, economic development can be understood as a sketchy reference to a causal mechanism that mediates between democracy and peace or, in a different interpretation, to a mechanism that favours democracy and peace. In fact, higher-level and lower-level variables can cohabitate in a causal model.

Particularly interesting, in this context, is a study from Rasler and Thompson where the authors analyze the democracy↔peace correlation in the light of a causal model containing four variables: external threat, domestic power concentration, democratization and external conflict behaviour. These variables are connected in the following way:²⁰

20 Cf. Rasler and Thompson 2004, 885-6.

- a) external threat increases the inequality of domestic power and the probability of external conflict, but has a negative effect on democratization;
- b) inequality of domestic power increases the probability of external conflict and affects negatively the degree of democratization;
- c) democratization increases the likelihood of peace.

In this particular case, one can consider the new variables (external threats and inequality of domestic power) as lower-level variables,²¹ but they can be analyzed at a still lower level, along the lines of a mechanistic explanation. For instance, the authors explain the relation between external threat and domestic power concentration in the following terms: “External threat creates incentives for fostering power inequalities, either at the chief executive level or at the more general elite level. Decentralizing power in the face of threat would seem inefficient and highly dangerous, perhaps even inviting attack.”²²

The democratic peace hypothesis also supports the thesis that the problem of confounders is more manageable by descending from the macro-level to the micro-level. For instance, Rousseau, Gelpi, Reiter and Huth studied the role of alliances as a possible confounder in the peace↔democracy correlation. Critics of democratic peace may argue that (1) democracies tend to build alliances and (2) allies tend to avoid military conflicts, but by confronting these specific hypotheses with their data set, the authors concluded that the variable “shared alliance ties”, far from being a confounder, is not statistically significant.²³

3. Limits of Mechanistic Explanations

However, the appeal to mechanisms has also a methodological price, which must be taken into account when exercising critical thinking in science. A first source of problems lies in the fact that the interference

21 Cf. Risjord 2014, 230.

22 Rasler and Thompson 2004, 886.

23 Cf. Rousseau *et al.*, 1996, 523.

effects that are invoked to legitimize the existence of cp laws reappear at the level of mechanisms; lower-level, same-level and even higher-level mechanisms may inhibit, or strengthen, the operation of a particular mechanism. One can distinguish two broad types of mechanistic interferences. In *one-way interference*, a mechanism may either strengthen or reduce the effects of another, operating mechanism. An extreme form of this kind of interference is suppression; a mechanism cancels the operation of an opposed mechanism. This possibility is an important reminder of the precautions one must have when defining causes as sufficient conditions of a phenomenon; it is possible that a particular causal mechanism is operating, but its usual effect is neutralized by an opposed mechanism. Sometimes, the triggering of a mechanism entails the immediate deactivation of the opposed mechanism. A typical example of this possibility is the duality of the fight-or-flight response. There are situations where the same stimulus has the capacity to trigger any of these responses, but one of these mechanisms prevails, suppressing the other. A member of a group oppressed by a dictator may feel initially a fear that dissuades her to act and also an anger that moves her to oppose the dictator, but one of these mechanisms will prevail and suppress the other, and it can be extremely difficult to predict which mechanism will be triggered.

A second type of interference is what we can call *two-way interference*: two opposing mechanisms may be both in action, interfering with each other. In some special cases, they might even have a zero net effect. For instance, pollsters know that the results of an electoral poll may trigger the so-called Bandwagon Effect, according to which voters tend to support the frontrunner, but there is also an Underdog Effect, according to which voters might have a preference for the candidate who is expected to lose. If there is a balance between these mechanisms (a zero net effect), the results of a poll do not affect the results of the election. In these cases, there is a surface illusion that the mechanisms are not operating, but they were, in fact, triggered, although without observable effects. Elster is particularly sensitive to the indeterminacy problems raised by the interference of opposing mechanisms in the human sciences, and for this reason he proposes a specific and distinctive account of mechanisms: "Mechanisms are frequently occurring and easily recognizable

causal patterns that are triggered under generally unknown conditions or with indeterminate consequences. They allow us to explain, but not to predict.”²⁴ Because this account is inspired by indeterminacy problems in the activation of mechanisms, and in particular of psychological and social mechanisms, its scope is limited.

A second important problem consists of the fact that mechanism-based explanations face also underdetermination problems, to the extent that the available evidence often allows different interpretations of the underlying structure of a correlation. In other words, the problem lies in the possibility of the use of arbitrary causal stories in the explanation of a correlation; it is often easy to create several *ad hoc* stories to explain a correlation. There are, however, two fundamental ways to restrict the number of these causal stories: they should be evaluated in the light of epistemic values, especially the value of generality (as stressed by Hedström and Swedberg);²⁵ they should be testable.

But the requirement of testability is, in turn, another source of methodological problems in mechanistic explanations. The problem of confounders or hidden variables does not affect only causal inferences from correlations or statistical evidence, but also causal inferences based on mechanisms. However, there are reasons to claim that the problem of confounders is more manageable in the case of mechanisms. Especially in the human sciences, it is often not possible to test or manipulate the macro-level variables of a correlation; and when this is possible such manipulations might raise ethical issues. Things are somewhat different at the micro-level, where the components of a mechanism can be, in favourable circumstances, directly observable, and it is in general easier to test empirically components of a system than the system as a whole and to restrict the list of possible confounders at the micro-level. In the words of Steel, who relates mechanistic analysis with “indirect causal inference” and “process tracing”:

“indirect causal inference attempts to learn the causal relationships among a set of variables by examining the

24 Elster 2015, 26.

25 Cf. Hedström and Swedberg, 1998, 10.

causal relations among a distinct yet related set. In process tracing, the distinct yet related variables represent features of component parts of the larger system of interest. The usefulness of process tracing, then, rests on the possibility that the causal relationships among the components are more directly accessible than those among the macrofeatures of the system”²⁶.

However, one must bear in mind that the testability of component parts of a mechanism can be particularly difficult in the case of non-modular mechanisms. According to the so-called modularity assumption implicit, e.g., in interventionist accounts of causation, it is possible to test, or intervene on, selected components of a mechanism, without affecting other components. Woodward, a leading advocate of the interventionist theory of causation, highlights this point: “the components of a mechanism should be independent in the sense that it should be possible in principle to intervene to change or interfere with the behaviour of one component without necessarily interfering with the behaviour of others.”²⁷ It is not easy to give a general assessment of the modularity thesis. On the one hand, it is often argued that evolution favours modularity, but in the domain of psychological and social mechanisms, where a large number of factors are interwoven, modularity is a highly dubious assumption, and empirical tests in that domain should not rely on it.²⁸

The fact that many explanations in the human sciences use an intentional vocabulary also raises an additional testability problem. There is a significant indeterminacy in the interpretation of intentional states, where trade-offs between the interpretation of beliefs, desires and actions may generate empirically equivalent, but incompatible explanations of the

26 Steel, 2008, 187.

27 Woodward, 2002, S374. For an in-depth and technical critique of the modularity thesis, cf. Cartwright, 2002.

28 Steel (2008, 52–53) advances two reasons to question the modularity of social mechanisms: modularity does not have adaptive value in socially stable environments; policy interventions directed at a specific group may change the behavior of other groups (that perceive, for instance, new opportunities as a result of the new policy). Steel (2008, ch. 8) also offers a detailed analysis of the possibility of “structure-altering interventions” in the social sciences (which violate the modularity assumption).

human action. As a result, it is more difficult to test causal stories in the human sciences and to exclude the arbitrary ones. There remains, however, an important distinction between ungrounded and grounded causal stories, which can play a useful role in theory choice.

A fourth problem that requires critical attention at the level of mechanistic explanations, and again a problem especially salient in the human sciences, lies in the fact that mechanisms are strongly context-dependent; in certain contexts, a particular arrangement of entities and interactions may generate a particular effect or regularity, but not in other contexts. The context-dependence of mechanisms is closely connected with the problem of the individuation of mechanisms. A mechanism can be described as a causal pattern, but this is a vague notion, because causal patterns are often fuzzy or ill-defined. Public policies constitute an illustration of this point: a welfare-to-work program instantiates a pattern, but in an abstract way. The real causal pattern involves contextual factors that vary, sometimes subtly, but significantly, from city to city, which makes it difficult to individuate the causal pattern.

Here we should consider the following suggestion to solve the individuation problem:

“Regularity of operation is crucial for individuating the boundaries of a given mechanism from its environment or context. Individuation of a mechanism involves, among other things, differentiating the entities and activities that constitute the mechanism from those that may occur in close spatiotemporal proximity to the mechanism but without contributing to it. Regularity does just that”²⁹.

The problem with this “regularity criterion” for the individuation of mechanisms lies in the fact that it is often difficult to make Andersen’s distinction between constitutive entities and activities of a mechanism and nonconstitutive collateral factors. There are well-defined mechanisms, but in many cases a mechanism is a causal pattern that contains, as it were, slots or space for contextually varying factors. For this reason

29 Andersen, 2012, 426.

we lack a strong notion of regularity at the level of mechanisms; what we have are “mechanistically fragile generalizations”.³⁰ The individuation problem is a real problem. At most, we can propose a Wittgensteinian solution: a mechanism is, in many cases, a *family* of more or less *similar* combinations of entities and activities. Understood as family-resemblance concepts, causal patterns or mechanisms are vague.

The requirement of “regularity of operation” is also objectionable in the case of what Glennan called “ephemeral” mechanisms (especially important in historical disciplines), as opposed to stable, regular or repeatable mechanisms. Glennan, however, and controversially, tries to preserve, in the case of these mechanisms, the regularity that is in general expected from mechanisms. He characterizes ephemeral mechanisms as a “collection of interacting parts” with the following traits:

1. the interactions between parts can be characterized by direct, invariant, change-relating generalizations
2. the configuration of parts may be the product of chance or exogenous factors
3. the configuration of parts is short-lived and non-stable, and is not an instance of a multiply-realized type.³¹

The first condition, with its Hempelian overtones,³² is stressed in order to make plausible Glennan’s suggestion that narrative explanations can be understood as descriptions of mechanisms. However, this condition simply assumes that human action can be explained with the help of lawlike generalizations. Glennan’s paper does not consider the important epistemological debate on the possibility of laws in the human sciences. Moreover, Glennan claims that “in ephemeral mechanisms, while the manner in which parts come together is chance or unpredictable, how they will interact with each other is not.”³³ This is, at least, a very optimistic claim, as the well-known difficulties of predicting the course of history show. A conjunction of known causal factors that interact for

30 The expression is frequently used by Craver (2007).

31 Cf. Glennan, 2010, 260.

32 Cf. Glennan, 2010, 261.

33 Glennan, 2010, 261.

the first time may have surprising effects; even the repetition of the same conjunction of causal factors can have unexpected effects in novel contexts. Glennan's conception of ephemeral mechanisms is contaminated by a very demanding account of the success of a narrative explanation: "I would argue that to the extent that a narrative fails to show the necessity of the outcome, it fails to explain."³⁴ However, this condition (the necessity of the outcome) is highly controversial from a historiographical standpoint. Surely many historians are willing to accept a narrative explanation if it is plausible (in the light of our knowledge of the psychological and social world), coherent and well grounded in the available evidence.

Finally, and in close connection with the individuation problem, mechanisms raise an extrapolation problem, which is especially acute in medical trials, for example. The fact that a substance is effective in the treatment of malignant tumours in mice does not allow us to infer that it will be also effective in the treatment of humans. The extrapolation problem is equally important in the human sciences. How can we be certain that a crime reduction policy that was successful in a particular city will also succeed in other cities? How can we know whether a welfare program that was successful in a country will also be successful in a different country? The psychological and social domains have a special complexity that includes, for instance, intentional states, values and norms, and this makes extrapolation inferences particularly difficult. At any rate, the awareness of the risks of extrapolations in the human sciences is positive; it can help to prevent serious mistakes. Although there is not a pure methodological solution to the extrapolation problem, there are two points that should be mentioned in this context. In the first place, qualitative methods (like unstructured or semi-structured interviews, case studies or participant observation) can help to identify differences between contexts and thereby reduce extrapolation errors. In the second place, it must be recognized that the human sciences, in particular, require *expert judgment*; a sensitivity to contextual differences and a sense of relevance based on the scientist's cultural, human and scientific *experience*.

34 Glennan, 2010, 262.

Before concluding this section, the connection between mechanistic explanations, *qua* causal explanations, and intentional explanations deserves a reference in this context. According to the standard picture, the latter explain an action by identifying the agent's beliefs and desires (or preferences) that constitute her reasons to act. Davidson famously argued, against anticausalist theories of action, that reason explanations are a species of causal explanations, because only those reasons that effectively cause an action can explain it; reasons that justify the action but do not play a causal role are explanatorily irrelevant. However, Davidson himself recognized that reason explanations should be distinguished from the causal explanations of natural phenomena; reasons are anomic (there are not strict psychological laws correlating beliefs, desires and actions) and normative (according to his Principle of Charity, the interpretation of an agent is based on rationality assumptions, more precisely, on a double assumption of truth and coherence). Understood in these terms, intentional explanations are different from mechanistic explanations. At any rate, mechanistic explanations may make reference to intentional states, and intentional explanations can be complemented by causal explanations of two types: there are causal (and mechanistic) explanations of the origin of beliefs and desires (or preferences), as well as of the unintended results of individual actions.³⁵ Mechanisms studied in the domains of cognitive psychology and social psychology illustrate the first type. Good examples are provided by mechanisms that reduce cognitive dissonance, the discomfort associated with inconsistencies inside one's system of beliefs and preferences. For instance, wishful thinking modifies our beliefs in the light of our preferences, and "adaptive preference formation" adjusts unconsciously our preferences and evaluations to what is possible.³⁶ On the other hand, market mechanisms offer good examples of the second, aforementioned type of causal explanations.

35 Cf. Føllesdal, Walløe and Elster 1988: 154 and ff.

36 The so-called "sour grapes" effect (analyzed by Elster 1983) illustrates the meaning of adaptive preference formation: when, in the Aesop's fable, the fox realizes that he cannot reach the desired grapes, he turns away, devaluating them as being probably unripe and sour.

4. Reductionism and Methodological Individualism?

By explaining higher-level phenomena with lower-level activities and entities, mechanistic explanations seem to have a reductionist character and to vindicate methodological individualism. Elster is a case in point: he combines a mechanistic approach to the social sciences with a methodological individualism. However, the behaviour of a mechanism as a whole cannot be explained by an aggregative view of the parts, because it depends on their organization, which originates higher-level properties that often resist theoretical reduction and have to be described in an irreducible vocabulary. The need to take into account the micro-foundations does not entail by itself a denial of supra-individual causal powers. This insight is behind what is called “structural individualism” in the social sciences:

“Structural individualism is a methodological doctrine according to which social facts should be explained as the intended or unintended outcomes of individuals’ actions. Structural individualism differs from traditional methodological individualism in attributing substantial explanatory importance to the social structures in which individuals are embedded”³⁷.

In contrast with methodological individualism, structural individualism recognizes the “explanatory importance of relations and relational structures”.³⁸ As we see, worries that mechanism-based explanations entail epistemological reductionism are ungrounded. Hedström and Udehn argue that the roots of structural individualism can be found in the famous American sociologist Robert Merton, more precisely in his defence of “middle-range theories” against the prestige of “grand theories” (like Parson’s functionalism). In their interpretation, “middle-range theories” describe social mechanisms, without any commitment to epistemological reductionism.³⁹ However, explanations formulated in holis-

37 Hedström and Bearman, 2009, 4.

38 Hedström and Bearman, 2009, 8.

39 Cf. Hedström and Udehn, 2009.

tic terms are not satisfactory from a mechanistic standpoint. Examples of purely holistic explanations are: “the rise in unemployment led to a higher crime rate” or “the economic depression was the main reason why the war broke out”.⁴⁰ A mechanistic scientist or philosopher would not be satisfied with explanations that only mention macro-level factors without providing a more detailed causal story that specifies the relevant micro-level factors.

It is interesting to notice that the notion of emergence, usually contrasted with the notion of reduction, can be accommodated by mechanistic authors. This is at first glance surprising, and, in fact, ‘emergence’ must be taken in this context in a relatively weak sense. Bechtel is an example of this theoretical position:

“While mechanistic explanations are in part reductionistic, they also accommodate the emergence of higher levels of organization and the need for autonomous inquiry into the regularities found amongst the denizens of these higher levels. So, as mechanists have consistently pointed out, the inherently reductionistic elements of mechanistic explanation need not threaten the explanatory autonomy of higher level psychological explanations – indeed, it depends on them to situate the mechanism in context”⁴¹.

Bechtel, as a philosopher of psychology, gives the instructive example of social psychology, where “environmental contexts often figure centrally in determining the activities of mental mechanisms, and therefore have a nontrivial role in being represented in the explanans of a mechanistic explanation.”⁴²

40 Cf. Zahle, 2016.

41 Bechtel and Wright, 2009, 127. The following passage helps to clarify the meaning of ‘emergence’ according to Bechtel: “Emergence, as I use the term here, simply recognizes that whole systems exhibit behaviors that go beyond the behaviors of their parts. It does have some bite against extreme reductionism, though, in that typically the behavior of the whole system must be studied at its own level with appropriate tools for that level (Bechtel, 2008, 129).

42 Bechtel, 2009, 127.

Bechtel and Craver, in an article with the provocative title “Top-down causation without top-down causes”, propose to replace the reference to interlevel causes in a mechanism with the notion of “mechanistically mediated effects”, which are a mix of constitutive relations⁴³ and (intra-level) causal relations. Putative examples of top-down causes (involving mechanisms) are described in the following terms:

“In this and many similar cases, a change in the activity of the mechanism as a whole just is a change in one or more components of the mechanism which then, through ordinary intra-level causation, causes changes in other components of the mechanism. (...) [T]here is nothing mysterious about appealing to a change at a higher level to explain a change at a lower level. Once we have described the mechanism mediating the effect, the drive to speak of this as a case of top-down causation vanishes, although such language might be useful as shorthand”⁴⁴.

As we can see, the strategy followed by Craver and Bechtel deflates the meaning of ‘top-down causation’. They stress that the notion of level that is implicit in “strongly emergent properties” is different from the mechanistic notion of level: “Levels of mechanisms are constitutive levels; levels of strong emergence are not.”⁴⁵ They also argue that their account “places a burden on the defender of strongly emergent properties to explain why top-down causation from emergent to nonemergent properties is different from mundane causation between two distinct properties.”⁴⁶ In sum, mechanistic explanations are not compatible with the postulation of strongly emergent properties, but they do not entail properly a refutation of strong emergentism.

43 An example of constitutive relation (mentioned by Craver and Bechtel) is the relation between the increase of mean kinetic energy of molecules and heat; it is a mistake to say that the kinetic energy of molecules causes heat.

44 Craver and Bechtel, 2007, 559–560.

45 Craver and Bechtel, 2007, 551.

46 Craver and Bechtel, 2007, 551.

5. Conclusion

The epistemological appeal to mechanisms has benefits and costs. Mechanistic explanations are more fine-grained than nomological explanations, can complement them, enrich the empirical content of explanatory theories (thereby opening new research paths) and offer a genuine understanding of the *explanandum*. On the other hand, and as we have seen, mechanism-based explanations raise several epistemological and methodological problems: they suffer from interference problems, which make it difficult to infer correctly the real causal relations; they are permeable to arbitrary causal stories; they have their own testability problems, resulting, for instance, from the presence of confounders (although this problem seems to be more tractable in the case of mechanisms) or from the impossibility to test in isolation the mechanism's components; they are strongly context-dependent and their individuation as causal patterns can even be controversial; finally, mechanism-based extrapolations are fallible because of the local character of mechanisms, which requires great care in extrapolation inferences.

At any rate, the study of mechanisms is an indispensable part of the human sciences, and the significant epistemological challenges that they raise do not lead to a defeatist, but only to a satisficing (to use a term popularized by Herbert Simon) account of the human sciences, according to which their complexity can be satisfactorily managed with the mobilization of a vast array of resources: quantitative and qualitative methods, mechanistic analysis, an epistemologically informed exercise of critical thinking and a constant cultivation of expert judgment, on the basis of an endless enrichment of the scientist's professional and human experience. This latter point is particularly important, because in addition to its methodological dimension, the human sciences also have an undeniable humanistic character.

References

- Andersen, H., 2012, The case for regularity in mechanistic causal explanation, *Synthese*, 189, 415–432.
- Bechtel, W., 2008, *Mental Mechanisms*, London, Routledge.

Bechtel, W. and Wright, C., 2009. What is psychological explanation? In: J. Symons and P. Calvo (eds.), *The Routledge Companion to Philosophy of Psychology*, London, Routledge, 113–130.

Craver, C. and Darden, L., 2013, *In Search of Mechanisms: Discoveries across the life sciences*, Chicago, The University of Chicago Press.

Craver, C., 2007, *Explaining the Brain: Mechanisms and the mosaic unity of neuroscience*, Oxford, Clarendon Press.

Craver, C. and Bechtel, W., 2007, Top-down causation without top-down causes, *Biology and Philosophy*, 22, 547–563.

Cartwright, N., 2002, Against modularity, the causal Markov Condition and any link between the two, *The British Journal for the Philosophy of Science*, 53, 411–453.

Crick, F., 1988, *What Mad Pursuit*, New York, Basic Books.

Cummins, R., 2000, How does it work? versus what are the laws?: Two conceptions of psychological explanation. In: F. Keil and R. Wilson (eds.), *Explanation and Cognition*, Cambridge, MA, The MIT Press, 117–144.

Davidson, D., 1980, *Essays on Actions and Events*, Oxford, Clarendon.

Earman, J., Roberts, J. and Smith, S., 2002, *Ceteris paribus* lost, *Erkenntnis*, 57, 281–301.

Elster, J., 2015, *Explaining Social Behaviour: More nuts and bolts for the social sciences*, Cambridge, Cambridge University Press.

Elster, J., 1983a, *Sour Grapes: Studies in the subversion of rationality*, Cambridge, Cambridge University Press – Paris, Éditions de la Maison des Sciences de l’Homme.

Elster, J., 1983b, *Explaining Technical Change*, Cambridge, Cambridge University Press.

Glennan, S., 2017, *The New Mechanical Philosophy*, Oxford, Oxford University Press.

Glennan, S., 2010, Ephemeral mechanisms and historical explanation, *Erkenntnis*, 72, 251–266.

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

Roux, S., 2018, From the mechanical philosophy to early modern mechanisms, in S. Glennann, and P. Illari, *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, New York and London, Routledge, 26–45.

Hedström, P. and Bearman, P., 2009, What is analytical sociology all about? In: *The Oxford Handbook of Analytical Sociology*, Oxford, Oxford University Press, 3–24.

Hedström, P. and Swedberg, R., 1998, Introduction. In: *Social Mechanisms*, Cambridge, Cambridge University Press, 1–32.

Hedström, P. and Udehn, L., 2009, Analytical sociology and theories of the middle range. In: P. Hedström and P. Bearman (eds.), *The Oxford Handbook of Analytical Sociology*, Oxford, Oxford University Press, 25–47.

Hempel, C., 1965, *Aspects of Scientific Explanation*. New York, The Free Press.

Illari, P. and Williamson, J., 2012, What is a mechanism? Thinking about mechanisms across the sciences, *European Journal for Philosophy of Science*, 2, 119–135.

Kincaid, H., 1996, *Philosophical Foundations of the Social Sciences*, Cambridge, Cambridge University Press.

Malinowski, B., 1935, *Coral Gardens and Their Magic*, New York, American Book Co.

Norkus, Z., 2005, Mechanisms as miracle makers? The rise and inconsistencies of the “mechanismic approach” in social science and history, *History and Theory*, 44, 348–372.

Pietroski, P. and Rey, G., 1995. “When other things aren’t equal: saving ceteris paribus laws from vacuity”, *The British Journal for the Philosophy of Science*, 46, 81–110.

Rasler, K. and Thompson, W., 2004, The democratic peace and a sequential, reciprocal, causal arrow hypothesis, *Comparative Political Studies*, 37, 879–908.

Risjord, M., 2014, *Philosophy of Social Science*, New York and London, Routledge.

Rousseau, D., Gelpi, C, Reiter, D. and Huth, P. 1996, Assessing the dyadic nature of the democratic peace, 1918–88, *The American Political Science Review*, 90, 512–533.

Russo, F., and Williamson, J., 2007, Interpreting causality in the health sciences, *International Studies in the Philosophy of Science*, 21, 157–170.

Steel, D., 2008, *Across the Boundaries: Extrapolation in biology and social science*, Oxford, Oxford University Press.

Von Wright, G. H., 1971, *Explanation and Understanding*, Ithaca, New York, Cornell University Press.

Woodward, J., 2002, What is a mechanism? A counterfactual account, *Philosophy of Science*, 69, S366– S377.

Wright, C. and Bechtel, W., 2007, Mechanisms and psychological explanations. In: P. Thagard (ed.), *Philosophy of Psychology and Cognitive Science*, Amsterdam, Elsevier, 31–79.

Zahle, J., 2016, Methodological holism in the social sciences. In: E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Available at: <<https://plato.stanford.edu/archives/sum2016/entries/holism-social/>>.