Abstract
Examples from the sciences showing that mechanisms don’t always succeed in producing the phenomena for which they are responsible led some authors to conclude that the regularity requirement can be eliminated from characterizations of mechanisms. In this article, I challenge this conclusion and argue that a minimal form of regularity is inextricably embedded in examples of elucidated mechanisms that have been shown to be causally responsible for phenomena. Examples of mechanistic explanations from the sciences involve mechanisms that have been shown to produce phenomena with a reproducible rate of success. By contrast, if phenomena are infrequent to the point that they amount to irreproducible observations and experimental results, they are indistinguishable from the background noise of accidental happenings. The inability to detect or measure the phenomenon of interest against the background noise of accidental correlations makes it impossible to elucidate a mechanism by experimental means, to demonstrate that a proposed mechanism actually produces the phenomenon, and ultimately to justify why a hypothetical scenario involving an irregular mechanism should be preferred over attributing irreproducible happenings to chance.

1. Introduction
The mechanisms described in scientific explanations are characterized in terms of productivity, organization, and regularity (Craver 2007; Darden 2006; Glennan 1996; 2002;
Productivity and organization are generally accepted defining characteristics of mechanisms, which are viewed as organized systems causally responsible for phenomena (Bechtel and Abrahamsen 2005; McKay Illari and Williamson 2012). The third characteristic - pertaining to whether mechanisms regularly succeed in producing phenomena, function or operate in a regular fashion, and have a fixed structure - is however disputed, most notably because of examples from the sciences showing that mechanisms fail to always produce the phenomena for which they are responsible. Given that the mechanisms described in scientific explanations are not perfectly regular, a debate ensued whether regularity should be replaced by weaker alternatives (Andersen 2012; DesAutels 2011; McKay Illari and Williamson 2012) or dropped altogether (Bogen 2005; Glennan 2011; Machamer 2004).

In this paper, I identify an important obstacle to some of the recent attempts to completely eliminate the requirement for mechanistic regularity. The problem lies in the fact that, if phenomena are infrequent to the point that they amount to irreproducible observations and experimental results, they are indistinguishable from the background noise of accidental happenings, thus making it impossible to distinguish phenomena generated by irregular mechanisms from chance correlations, as well as to interpret the results of experimental interventions required to demonstrate the causal contribution of mechanisms to the phenomena for which they are allegedly responsible. As I will argue in the paper, the mechanistic interpretation can be favored if it is possible to increase the experimental reproducibility of a phenomenon. Since reproducibility plays a fundamental role in the interpretation of the experimental results required to elucidate and support the mechanistic basis of a phenomenon, I conclude that a weak form of regularity is inextricably embedded in examples of elucidated mechanisms that have been shown to be causally responsible for phenomena.

The paper is organized as follows: In section 2, I discuss the different senses in which mechanisms are said to be regular, objections to mechanistic regularity, and alternative solutions. In section 3, I introduce the problem of irreproducible phenomena. In section 4, I discuss possible theoretical workarounds. In section 5, I elaborate the notion of experimental reproducibility, and

---

1 “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, 3). Alternatively, a mechanism is “a complex system that produces that behavior by the interaction of a number of parts, where the interactions among parts can be characterized by direct, invariant, change-relating generalizations” (Glennan 2002).
discuss how it relates to experimental setups, experimental error, chance, and mechanistic regularity. In section 6, I summarize my findings and conclusions.

2. Mechanistic regularity

Mechanisms are said to be ‘regular’ in at least three senses. According to the Machamer-Darden-Craver characterization, mechanisms are “entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions” (2000, 3), where these regular changes amount to the phenomena produced by mechanisms. Thus, mechanisms are regular in the sense (i) that they reliably succeed in producing the phenomena for which they are responsible. Mechanisms may also be said to be regular in the sense (ii) that they operate in the same manner: “Mechanisms are regular in that they work always or for the most part in the same way under the same conditions” (Machamer et al. 2000, 3). This kind of regularity is also present in Stuart Glennan’s characterization of mechanisms as complex systems producing behaviors in virtue of the interaction of their parts, where these interactions are characterized by causal laws (1996, 52) or invariant, change-relating generalizations (2002, 344; 2010, 260). Finally, mechanisms may be said to be ‘regular’ in the sense (iii) that they exhibit stable structural and organizational features: “the parts have spatial (location, size, shape, and motion), temporal (order, rate, and duration), and active (for example, feedback) relations with one another by which they work together to do something” (Craver 2007, 189).

The fact that a mechanism always operates in the same way (ii) and exhibits stable structural features (iii) is likely to contribute to its ability to regularly generate changes from input to output conditions (i). Nevertheless, (ii) and (iii) are not sufficient conditions for (i), and therefore counterexamples to regularity (i) do not automatically entail that regularity (ii) and/or (iii) are false. A mechanism may be unreliable (inefficient at producing a phenomenon) despite the fact that it operates in the same way and has stable structural features. In particular, biological mechanisms may be unreliable due to adaptive pressures, such that less reliable mechanisms are favored, typically because of the costs associated with more reliable mechanisms (e.g., for energy consumption reasons, brain size tends to be proportional with the ¾ power of the mass of an organism); limitations imposed by evolutionary history and developmental constraints (e.g., reti-
nal organization); limitations imposed by their components and mode of functioning (e.g., reset delays, such as refractory period for action potentials in neurons; reliance on tautomeric configurations of organic molecules occurring only in certain percentage of cases); occasional malfunctions (e.g., non-specific binding may occur); and interference from other mechanisms (e.g., competition for shared resources or components, such as metabolic pathways competing for the same substrates).2

Objections have been raised against all three senses in which mechanisms may be said to be regular. The notion that mechanisms function or that their parts interact in a law-like fashion (regularity ii) has been challenged on the grounds that very few interactions and activities in the biological world behave in a fully invariant or fully regular way (DesAutels 2011). In an attempt to provide a mechanistic account of singular events, such as the emergence of life on Earth, Glennan (2010) argues that some mechanisms are ephemeral collections of interacting parts, thus challenging the requirement of structural regularity (iii). However, the most influential and widely accepted objection targets regularity (i) by means of examples from the sciences showing that mechanisms don’t always succeed in producing the phenomena for which they are responsible (Bogen 2005; Glennan 2011; Machamer 2004).

The emerging consensus is that the regularity requirement must be weakened, a trend reflected in recent characterizations of mechanisms. For example, William Bechtel and Adele Abrahamsen (2005, 423) define a mechanism as “a structure performing a function in virtue of its component parts, component operations, and their organization […] responsible for one or more phenomena.” Phyllis McKay Illari and Jon Williamson (2012, 120) propose an even more relaxed and more generally applicable characterization, according to which a “mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon.” Nevertheless, they are careful to point out that “dropping explicit refer-

2 Conversely, the fact that a mechanism reliably succeeds in producing a phenomenon (regularity i) does not automatically entail that the mechanism always functions in the same manner (regularity ii). This might be particularly relevant in molecular biology, where biological phenomena are produced as a result of many copies the same molecular mechanism functioning at the same time (Bogen 2005, 414, note 22). Thus far, mathematical models have shown that the same molecular network may be characterized by more than one stable state, and it has been hypothesized that such states may underlie developmentally differentiated cell types (Kauffman 2004) or physiological cell states [e.g., proliferating vs. apoptotic cells; (Huang 1999)]. According to these hypotheses, the same molecular network may function in two or more significantly different ways, although there is nothing here to suggest that each copy of a molecular mechanism functions in a singularly different way.

4
ence to regularity does not imply that mechanisms in general do not have to exhibit some form of regularity or stability. Some far weaker form of regularity or stability is already present in the idea of mechanisms being responsible for the phenomenon” (2012, 125). Similar views in respect to mechanistic regularity are expressed by Benjamin Barros (2008), Lane DesAutels (2011), and Holly Andersen (2012), who argue that regularity must be maintained, albeit in a weaker, stochastic form, because it plays a crucial role in the ability of mechanisms to support generalizations and predictions, assets which are ultimately required in order to provide better explanations.

In contrast with the above proposals to weaken the regularity requirement, some authors propose to eliminate it altogether (Bogen 2005; Glennan 2011; Machamer 2004). The proposal is motivated by a more ambitious philosophical project aiming to ground causation in mechanistic productivity rather than the contrastive notion of difference-making (Hall 2004; Psillos 2004; Sober 1984). One of the main arguments in favor of elimination is that perfectly deterministic regularity is not necessary for mechanistic explanation:

The mechanisms which initiate electrical activity in post-synaptic neurons by releasing neurotransmitters are [...] numerous enough, and each of them has enough chances to release neurotransmitters to support the functions of the nervous system. But each one fails more often than it succeeds, and so far, no one has found differences among background conditions which account for this (Kandel et al. 2000, 261). No one takes the irregularity of their operation as a reason to deny that on the relatively rare occasions when they do operate successfully these mechanisms release neurotransmitters and exert a causal influence on post synaptic neuronal activity. (Bogen 2005, 400)

Bogen pushes the argument one step further, and extrapolates that “no matter how irregularly the relevant mechanism may operate, in order for a causal explanation to be acceptable, descriptions of its operation must tell us (at least to a good approximation) what happens when it produces the effects it explains. They can tell us that regardless of whether the mechanism satisfies them in any, let alone every other case” (2005, 411). The metaphysical upshot of the argument is that the kind of mechanisms described in scientific explanations are not required to regularly generate a phenomenon or to regularly operate in the same way; thus, regularity (i) and (ii) are explicitly
rejected. Furthermore, the claim that a mechanistic explanation amounts to a description of operations (as opposed to a description of structural features), suggests that the presence of stable structural features (regularity iii) may not be necessary either.

### 3. The dilemma of irreproducible phenomena

Undoubtedly, there are well documented cases of mechanisms that fail to reliably produce their target phenomena. It is probable that perfectly reliable mechanisms don’t exist in biology, engineering, or elsewhere. Intrinsic flaws and limitations, combined with exogenous sources of interference make it such that all mechanisms are likely to fail more or less often. Furthermore, biological mechanisms are subjected to adaptive pressures which may favor less reliable mechanisms over more reliable ones [e.g., brain size (Roth and Dicke 2005)], while evolutionary history and developmental constraints are known to generate limitations hindering the ability of mechanisms to operate optimally or to consistently generate phenomena [e.g., evolution of the eye (Lamb et al. 2007)].

Nevertheless, if there is nothing controversial about the notion that mechanisms are not perfectly regular, it does not follow from here that examples from the sciences support the more ambitious claim that the regularity requirement can be completely eliminated from characterizations of mechanisms. Examples from the sciences refer to elucidated mechanisms shown to be productive of phenomena, where the fact that it was possible to show that these mechanisms produce their target phenomena is a very strong indication that these mechanisms are not completely irregular. In order to show that a proposed mechanism actually generates a phenomenon, there must be a way to demonstrate that interventions on the mechanism and its components have an effect on the phenomenon [e.g., the destruction of the mechanism results in the non-occurrence of the phenomenon (Baetu 2012; Craver 2007)]. The methodology of experimental research in the sciences dictates that the phenomena that can be studied in any given experimental setup are detectable/measurable [ideally as statistically significant deviations] against the background noise, or baseline, of chance happenings such as uncontrolled experimental variation and accidental correlations. Without the ability to draw a distinction between phenomenon and noise, a set of correlated events that may constitute a phenomenon is equally correlated with myriad other
events taking place in the universe, thus making it impossible to carve the sum total of happenings in the universe into individually recognizable phenomena. In turn, the failure to recognize the target phenomenon makes it impossible to demonstrate that the interventions on its mechanism have or don’t have an effect vis-à-vis that phenomenon.

This methodological bottleneck does not preclude the possibility of elucidating irregular mechanisms and demonstrate their causal contribution to rare phenomena. It does, however, impose an important limitation: the causal productivity of a mechanism cannot be empirically demonstrated unless it makes a difference against the background noise of accidental happenings. This limitation becomes obvious if a phenomenon is infrequent to the point that it amounts to an irreproducible observation or experimental result. Such a phenomenon is indistinguishable from the background noise of accidental happenings, thus making it impossible to demonstrate that a proposed mechanism actually produces the phenomenon. This lack of empirical evidence ultimately raises a dilemma whether a hypothetical scenario involving an irregular mechanism should be preferred over attributing irreproducible happenings to chance.

Consider, for example, the following experiment drawn from the case study discussed by Bogen (2005, Fig. 2): “We used [...] Fabs to examine the position of the voltage-sensor paddles when the KvAP channel functions in lipid membranes [...], and to assess whether they change their position when the channel gates open in response to membrane depolarization” (Jiang et al. 2003b, 42). The goal of the experiment is to provide evidence that the production of action potentials involves a specific mechanism of opening and closing (gating) ion channel proteins. The experiments involves the binding of antibodies – or rather their antigen binding domains, known as Fab – to a particular part of the ion channel, known as the voltage-sensor domains. According to the mechanistic model proposed by Jiang et al., the voltage-sensor domains form small ‘paddles’ on the sides of the ion channel; ions bind these ‘paddles’ and are then moved through the cell membrane on the other side, causing the channel to switch between the ‘opened’ and ‘closed’ conformations. Antibodies are huge, bulky proteins, which will not cross the membrane. Thus, when antibodies bind the voltage sensor ‘paddles’, these ‘paddles’ cannot move freely; as a result, ion channels fail to open and close properly, ultimately hindering the ability to generate action potentials.
Bogen (2005, 406) argues that the mechanistic model proposed by Jiang et al. “is controversial in a way that illustrates a second role for laws and lesser generalizations.” Indeed, unlike the paradigmatic case of the Hodgkin-Huxley mathematical model of the action potential, which relies on derivations from the laws of electrochemistry (Hodgkin and Huxley 1952; Weber 2005), the model proposed by Jiang et al. is a typical example of an experimentally elucidated mechanism\(^3\), where the explanation does not rely on any kind of mathematical derivations. However, it does not follow that, because no laws are involved in the explanation, there is no mechanistic regularity involved. Reproducibility is a core requirement of current methodology of experimental research. If action potentials are a reproducible phenomenon, that is, if action potentials can be consistently reproduced with a consistent rate of success per number of experimental trials, then, by conducting an adequate number of trials, one can determine whether the binding of antibodies correlates with a reduction of produced action potentials. Consider, by contrast, a scenario whereby action potentials cannot be experimentally reproduced: there are sporadic reports of them happening, but nobody can consistently reproduce them. How would one figure out if action potentials were not produced because there was an interference with the ion channels; or because action potentials just didn’t happen on the occasion of the experiment; or because ion channels have nothing to do with action potentials; or because there are no such things as action potentials anymore than there are channels on Mars, but only experimental artifacts attributable to an inadequate experimental setup?

The standard response would be to dismiss the dilemma as being a matter of epistemic access: the fact that we cannot detect the rare effects of irregular mechanisms does not mean that there is not mechanism present. One cannot but accept this response as perfectly reasonable. Nevertheless, this does not change in the least the fact that, in as much as it is impossible to distinguish an irreproducible phenomenon from accidental happenings, the mechanistic explanation is empirically equivalent to attributing the phenomenon to chance, thus creating a situation analogous to that of the underdetermination of theory by evidence. Furthermore, unlike a common case of underdetermination, where there is no doubt that there is a phenomenon in need of

\(^{3}\) “We have studied the mechanism of voltage-dependent gating using biochemical, X-ray crystallographic and electrophysiological methods” (Jiang et al. 2003a, 34). All the items in the list are experimental techniques routinely used in neuroscience, and, with the exception of electrophysiological measurements, in molecular biology in general.
an explanation, but it is not clear which explanation should be favored, the alternatives in the case of irreproducible phenomena are much more radical, since they entail fundamentally incompatible metaphysical interpretations, coupled with a deep incertitude about whether there is any need to seek an explanation over and above attributing a phenomenon to chance. Ultimately, mechanistic explanations require real mechanisms. In contrast, coincidences have no mechanistic basis, or any other basis for that matter, be it causal or nomological; strictly speaking, they do not have and do not need an explanation of any kind. Thus, the systematic failure to reproduce a phenomenon not only forfeits the benefits of empirical confirmation by bringing experimental research to a halt, at a more fundamental level it also makes it difficult to justify a research project seeking an explanation in the first place.

On the one hand, the absence of empirical significance, which strips mechanistic explanations of tangible benefits and the possibility of experimental confirmation, combined with the lack of discrimination between two radically different metaphysical interpretations and their implications for the discovery process, is an unhappy result which we would like to avoid. On the other hand, it is also clear that there are no good reasons to assume that mechanisms are perfectly regular. The task, then, is to find out to how far irregularity can be pushed while preserving the empirical relevance of a mechanistic explanation. In the remaining sections of the paper I discuss, first, theoretical workarounds (section 4, ‘Theoretical workarounds’), and then what I take to be a more satisfactory experimental solution to the dilemma (section 5, ‘The experimental solution’).

4. Theoretical workarounds

The dilemma of irreproducible phenomena can be weakened if it is possible to show that there is a principled way of drawing a distinction between phenomena, a subset of which may be explained as effects of productive mechanisms, and accidental happenings. A theoretically grounded principle may not endow explanations empirical significance, but it can provide a scientific justification for preferring a mechanistic explanation over an accidental happenings sce-

---

4 This applies both to the ‘ontic’ and the ‘epistemic’ views of mechanistic explanations (McKay Illari and Williamson 2011; 2012). According to the ontic view, explanations are objective features of the world (Craver 2007; Salmon 1984). On the epistemic view, they are descriptions of the causal-mechanistic structure of the world (Bechtel 2008).
nario. I begin by discussing an intuitive solution based on our cognitive abilities to understand how mechanisms produce phenomena, and then proceed to a solution relying on the notion of theoretical possibility.

Advocates of the complete elimination of regularity suggest a straightforward principle: rely on our capacity to intuitively understand how mechanisms produce phenomena. According to Bogen (2005, 411), descriptions of a mechanism’s operation tell us what happens when the mechanism produces the effects it explains. This view is developed in more detail by Peter Machamer (2004; Machamer et al. 2000): “Intelligibility [...] is provided by descriptions of mechanisms [...] through the elaboration of constituent entities and activities that, by an extension of sensory experience with ways of working, provide an understanding of how some phenomenon is produced” (Machamer et al. 2000, 22). Thus, our current store of mechanistic explanations allows us to distinguish phenomena from accidental happenings based on whether we can imagine or intuitively understand how a mechanism might produce a phenomenon.

There are arguments to support the view that our intuitions are reliable (e.g., evolutionary arguments), which I will not discuss here. The main concern is that, in order to provide a workaround for our dilemma, the principle of intuitive understanding must be able to handle irreproducible phenomena. This is not something that can be clearly established. The little that we know about how our mental capacities of imagination and intuitive understanding suggests that they rely on associative learning and generalizations from previous experience, in which case a certain amount of regularity is implied. Furthermore, intuitive understanding seems to work in a primarily qualitative way. Nothing seems to indicate that we are particularly good at mentally simulating quantitative and dynamic aspects of mechanisms and the phenomena they produce, yet quantitative-dynamic aspects are crucial for determining whether a mechanism is, at least in principle, causally relevant for producing a phenomenon, especially when the mechanism in question is an irregular one and the phenomenon it produces is extremely unfrequent.\footnote{For instance, mathematical models revealed a wealth of surprising properties of molecular mechanisms which previously escaped our intuitions (Baetu forthcoming).}

Another way around the dilemma is to argue that, given a suitable theoretical background, it is possible to hypothesize a putative mechanism and show by means of mathematical derivations or computer simulations that, in principle, the mechanism can generate or contribute to the
generation of a rare, improbable, or singular phenomenon. To avoid charges of circularity, the theoretical background in question must be confirmed in respect to other predictions, and therefore constitute an accepted piece of scientific knowledge. In addition, in order for this solution to work, the theoretical background must also be rich enough to allow for detailed descriptions of mechanisms.

These expectations are met in some fields of investigation, most notably physics, but not without some limitations. One important concern is that, even in physics, there is always a lingering concern that a thoroughly confirmed theory may fail in limit cases (Cartwright 1999; Mitchell 1997); for this reason, physicists are particularly keen on testing exotic predictions rather than just accepting them as true. This uncertainty creates a universal drive in science for preferring experimental confirmation over theoretical justification.

A second limitation stems from the fact that, in many fields, including ones in which mechanistic explanations are particularly prevalent, there is no theoretical background rich enough to allow for detailed descriptions of mechanisms. This is often the case in the life sciences: even though we have good reasons to believe that all living organisms are the product of evolution and all biological activity has a physico-chemical basis, these broad theoretical considerations are not sufficient to develop in any reasonable detail mechanistic explanations of specific biological phenomena, such as inheritance, development, immunity, or cancer. Such theoretical considerations must be generously complemented by additional empirical constraints before they become stringent enough to single out a preferred set of possible mechanisms. This suggests that scientists must have access to at least some of the empirical details before they can hypothesize a possible mechanistic explanation.

A third limitation stems from a strong demand for identifying the actual causes of a phenomenon. A mechanism that can produce a phenomenon – be it in principle, or as demonstrated by past experience – yet, as a matter of fact, does not produce the phenomenon in question, can-

---

6 Theoretical approaches proved extremely fruitful, and there cannot be any doubt that they play an important role in scientific discovery. For example, it seems highly unlikely that entities that are not directly observable, such as atoms or DNA, would have ever been discovered have they not been previously postulated by theories aiming to explain puzzling phenomena, systematize our knowledge, or probe the limits of that which is theoretically possible.

7 E.g., possible mechanisms of DNA replication were devised only after a detailed knowledge of the chemical structure of the DNA double-helix became available.
not be said to be a satisfactory causal explanation of that phenomenon. There is a wealth of examples from the life sciences showing that a mechanism possible according to some theoretical background, or even a mechanism that has been shown to produce similar phenomena, is not necessarily the mechanism responsible for the phenomenon under investigation. For instance, while all cancers resemble each other, they are produced by many distinct mechanisms, and one cannot claim to explain, and certainly not treat, a particular type of cancer by vaguely pointing to possible mechanisms. Likewise, even if there are many conserved features of molecular mechanisms shared by across many phyla, subtle differences are equally ubiquitous, from slight variations of the genetic code, to significant dissimilarities in the sequence, biochemical structure and mode of operation of the molecular machinery responsible for replicating and expressing the genome. To give just an example, differences between the mechanisms of translation in eukaryotes and prokaryotes explain why antibiotic treatments work (antibiotics kill bacteria, but not their eukaryotic hosts); thus, even though the mechanistic explanation of translation in prokaryotes provided valuable insights on how translation is likely to proceed in eukaryotes, the mechanism of prokaryotic translation does not explain eukaryotic translation for the very simple reason that the former is not the cause of the latter. This kind of considerations supports the view that a successful mechanistic explanation should point to the actual mechanism(s) producing a given phenomenon (Baetu 2012; Craver 2006; McKay Illari and Williamson 2011). The satisfaction of this requirement translates in pragmatic benefits, since knowledge of actual and specific mechanisms provides the basis for developing technologies that allow us to control phenomena (trigger them at will, prevent them from occurring, or alter them in ways that benefit us).

Given the above limitations, it seems unlikely that a strictly theoretical approach can provide a conclusive solution to the dilemma of irreproducible phenomena. If practicable, theoretical considerations may tilt the balance in favor of a mechanistic interpretation, or at very least justify a ‘quest for mechanisms’ research program. Nevertheless, it should be noted that a theoretical approach is in fact impracticable in many fields of investigation dealing with mechanistic explanations. Moreover, no amount of theorizing can demonstrate the empirical significance of a mechanism where there is none to be observed, and, in the long run, a systematic failure to demonstrate the empirical significance of the postulated mechanisms is likely to lead to the demise of
the research program. This view is reflected in most fields of investigation: an empirically significant scientific explanation is always preferred over a scientific explanation involving theoretically grounded, but untestable possibility. In many cases, the latter are explicitly treated as hypotheses guiding research and denied the status of accepted explanations.

5. The experimental solution

In this section, I introduce an experimental solution to the dilemma of irreproducible phenomena inspired from the experimental practice of the life sciences. I take this solution to be more satisfactory because it preserves the empirical significance of mechanistic explanations while remaining compatible with the view that mechanisms can be irregular. I begin by discussing the particular form the dilemma takes in an experimental context and how irreproducibility can be attributed to a poorly circumscribed experimental setup, after which I proceed to elaborate a solution based on the notion of experimental reproducibility and how this solution impacts on the requirement for mechanistic regularity.

5.1 The dilemma of irreproducible phenomena in experimental practice

Mechanisms can be, and often are elucidated experimentally, by intervening on putative mechanistic components and organizational features. In a typical study, interventions on potential mechanistic components provide evidence for their causal relevance vis-à-vis the phenomenon under investigation (Craver 2007, ch. 4; Woodward 2003; Woodward and Hitchcock 2003), while interventions on multiple variables, usually the initial conditions and a putative component, further demonstrate that the putative component is indeed part of the mechanism responsible for producing changes from input to output conditions (Baetu 2012).

However, if a phenomenon is undistinguishable from the background noise of chance happenings, it is impossible to determine whether experimental interventions have an effect on the target phenomenon or not; as a result, the mechanism cannot be elucidated and shown to produce the phenomenon. A solution is nevertheless possible if we take into account the fact that, from an experimental standpoint, three (rather than two) metaphysical scenarios may account for an irreproducible phenomenon:
• the phenomenon is irreproducible because it is generated by a highly irregular mechanism;
• there is no phenomenon to be explained because we are dealing with a strictly accidental happening, such as a coincidence devoid of any mechanistic underpinnings;
• a phenomenon is in fact produced by a mechanism, but the phenomenon is masked by chance variation and accidental correlations associated with a poorly circumscribed experimental setup.

As I will argue in the remaining of this section of the paper, empirical evidence for a phenomenon and its mechanistic basis can be gathered, thus favoring the mechanistic interpretation, if it is possible to increase the reproducibility of an event by operating changes in an experimental setup; that is, if it is possible to show that the third option is the case.

5.2 Poorly circumscribed experimental setups

A poorly circumscribed experimental setup may be too variable and/or not specific enough. In the first scenario, the experimental setup varies from one trial to the next, most notably by sometimes including and sometimes excluding a component of the mechanism generating the phenomenon. The lower the control over the various aspects of the experimental setup, the higher the probability of uncontrolled variation of experimental variables. For example, in 1667, Jean Baptiste Denis, performed the first documented transfusion of blood in man. While successful in some cases, the practice of blood transfusion remained a highly risky medical procedure for almost four centuries due to a seemingly random phenomenon known as transfusion reaction (hemagglutination, hemolysis and renal failure). It is only after the discovery of the blood groups by Carl Landsteiner in the first decades of the 20th century that it became clear that differences in blood type, and the antigens associated with them, play a crucial role in the mechanism responsible for transfusion reactions.

In the second scenario, the spatial/temporal scale is inadequate and the experimental setup fails to specifically reunite only the causally relevant factors and mechanistic components. The more global the setup, the higher the probability of observing accidental correlations; as a result, researchers have to deal with a cacophony of correlations that not only have nothing to do with the phenomena under investigation and its putative mechanism, but also mask the relevant
mechanistically-underpinned correlations. For example, a poorly circumscribed experimental setup for studying immune responses\(^8\) may consist of whole blood extracts rather than purified leukocytes. In jawed vertebrates, including humans, leukocytes (white blood cells) play a crucial role in initiating, mediating and regulating immune responses. Studying purified leukocytes rather than whole blood facilitated the discovery of the various chemicals synthesized and secreted by leukocytes (e.g., mRNA, cell surface receptors, cytokines, antibodies), which are sometimes present in very minute amounts, as well as the complex sequences of cell-cell interactions required for orchestrating immune responses, ultimately paving the way towards the elucidation of the molecular and cellular mechanisms responsible for immune responses. By contrast, studying whole blood extracts, which consist primarily of erythrocytes (red blood cells) results in the minimization of the relevant molecular and cellular correlates of immunity, which become a lot harder to detect, while increasing the overall background noise of molecular and cellular processes that accidentally correlate with immune responses without being mechanistically related to it.

In both scenarios, artifacts are generated by the choice of a particular experimental setup. As the combined background noise generated by experimental uncertainty and accidental correlations increases, it becomes more and more difficult to distinguish between chance correlations and correlations related to the mechanistic basis of the phenomenon of interest. It is worth pointing out that errors stemming from poorly circumscribed experimental setups constitute a distinct category, and can occur independently of other common kinds of experimental error, such as errors due to faulty instruments/techniques and errors related to quantitative measurements. The fact that an instrument/technique has been shown to be reliable and all the internal control checkpoints have been passed eliminates an important source of experimental errors, namely those related with the functioning of the instrument/technique. Nevertheless, this cannot guarantee that the experimental setup is properly circumscribed. For example, one may use exactly the same instruments to study immunity in an experimental setup consisting of whole blood extract and in that of purified leukocytes. Even though the instruments/techniques have been thoroughly checked, it will be problematic to identify the relevant cell surface markers associated with an

---

\(^8\) An organism’s defense reactions to pathogens and potentially dangerous chemicals. The first evidence for the cellular basis of immunity comes from the work of Elie Metchnikoff on phagocytosis.
immune response when using whole blood extracts simply because a whole blood extract is made mainly of erythrocytes, which play no role in immunity. Studying immune responses in this setup is a lot like pointing a telescope to the moon in broad daylight: there is nothing wrong with the telescope; the problem lies in the background noise associated with the experimental setup.

5.3 Experimental reproducibility

When dealing with a phenomenon infrequent to the point that it blends in the background noise of chance correlations, the goal is to increase the reproducibility of the phenomenon by manipulating the experimental setup in which the phenomenon is studied; these manipulations may be conducted in light of relevant theoretical considerations, or by trial and error. Researchers alter the experimental setup and observe whether this has any impact on the reproducibility of the phenomenon. If the degree of reproducibility increases, the following inference is granted: a phenomenon which occurs more frequently in an experimental setup than another is more likely to require an explanation because a chance happening would have been just as infrequent irrespective of the experimental setup in which it is observed.

As reproducibility is increased, standardized experimental setups emerge, providing operational characterizations of the experimental protocols and of the initial conditions maximizing the reproducibility of the phenomenon under investigation (i.e., the kinds of details one finds in the ‘Methods and Protocols' section of a scientific paper; such an experimental setup is often referred to as ‘an experimental model of a phenomenon’). Standardization allows different research groups to exchange data and findings about the same object of inquiry (Ankeny 2001; Bowker and Star 1999; Clarke and Fujimura 1992; Müller-Wille 2007). Standardization also means that the target phenomenon is consistently produced with a fixed rate of success in a given finite number of attempts to produce it. In turn, frequency of reproducibility determines the number of trials required to interpret a series of unsuccessful attempts as likely to constitute a negative test result (non-occurrence of the phenomenon); the higher the frequency, the lesser the number of trials. Once a phenomenon can be reproduced consistently, it becomes possible to interpret interventions on the various physical aspects of the experimental setup consistently resulting in
changes in the phenomenon of interest as likely to target causally relevant factors, such as components of the mechanism producing the phenomenon.\textsuperscript{9}

If it is further assumed as a working hypothesis that the putative phenomenon under investigation has a mechanistic basis, reproducibility supports a second inference: the fact that the phenomenon can be consistently reproduced in a series of trials reduces the possibility that the required mechanistic components just happen to be reunited within the physical boundaries of the experimental setup by chance. The probability that all the necessary mechanistic components for the production of the phenomenon are reunited in a series of consecutive trials by chance alone gets smaller and smaller as the rate of reproducibility increases. The more consistently reproducible a phenomenon is, the more likely it is that the mechanism sought by researchers is located within the spatial-temporal boundaries of the experimental setup in which the phenomenon is documented.

Thus, an increase in reproducibility supports two crucial inferences guiding the discovery process: (A) the experimental setup is likely to be circumscribed precisely enough to distinguish correlations related to the phenomenon of interest from the background noise of chance correlations; and (B) the experimental setup is likely to contain all the relevant factors and mechanistic components. Without (A), researchers have no experimental means to determine whether experimental interventions have an effect on a phenomenon or not. Without (B), scientists have no reasons to assume that the experimental setup in which a phenomenon is documented circumscribes within its physical boundaries a mechanism producing that phenomenon and no clues where to look for a mechanism and its components. Without reproducibility, researchers have no reasons to look for a mechanism, and no experimental means to demonstrate that there exists a mechanism as postulated by some theory or to elucidate one experimentally.

To illustrate the above, let us consider the blood transfusion example. Due to experimental difficulties and medical risks associated with the procedure, blood transfusion between living organisms was, until recently, an infrequent procedure, meaning that transfusion reaction was a phenomenon researchers rarely had the occasion to witness and study. Furthermore, on the rare

\textsuperscript{9} Conversely, interventions that do not result in changes may be interpreted as possibly targeting causally irrelevant factors, although it should be noted that negative results do not preclude the possibility of non-linear causal contributions or non-modular causal interactions in which factors that must act in concert with other factors in order to contribute to a change in the phenomenon of interest.
occasions transfusions were performed in animals and humans, the phenomenon was documented as a highly unpredictable, seemingly random adverse reaction. The first breakthrough, which made possible the systematic investigation of the phenomenon, was the realization that two of the key features of the transfusion reaction, hemagglutination and hemolysis, can be documented in vitro, by mixing blood samples. By circumscribing more precisely the experimental setup, it became possible to drastically increase the reproducibility of the phenomenon, while at the same time eliminating errors and confounding side-effects related to the transfusion procedure, but not to the immunological reaction itself. This net gain in reproducibility allowed Karl Landsteiner to identify the A, B and O blood types based on the observation of consistently reproducible patterns of reaction. At the same time, by carefully circumscribing the minimal experimental setup in which hemagglutination can be observed, Landsteiner was able to infer that something in the blood plasma of the recipient interacts with the blood cells of the donor, causing them to agglutinate. This observation paved the way to the eventual elucidation of the molecular mechanism responsible for transfusion reactions.10

5.4 From experimental reproducibility to the elucidation of mechanisms

Experimental reproducibility does not guarantee that reproducible phenomena are generated by mechanisms, just as lack of reproducibility does not guarantee that irreproducible phenomena are accidental happenings. Ultimately, it is impossible to distinguish with absolute certitude between phenomena produced by irregular mechanisms and accidental happenings on the basis of reproducibility alone. In order to proceed to the next step of the investigation, researchers must eventually assume that what is likely to be the case, is actually the case. That is, they must assume that (A∗) reproducible findings constitute a phenomenon in need of an explanation, and (B∗) the experimental setup circumscribes a hypothetical mechanism producing the phenomenon. Assumption (A∗) states that the observed events are not part of background noise. This assumption is embedded in the more ambitious assumption (B∗), stating that the experimental

---

10 In parallel, another breakthrough resulting in the elaboration of a more precisely circumscribed experimental setup and the elucidation of the genetic basis of transfusion reactions, was the realization that immune rejection is less likely to happen if donor and recipient are genetically related, followed by the discovery that the probability of immune reactions associated with transfusions follows consistently reproducible probabilistic patterns that vary depending on the degree of genetic relatedness between donor and recipient. For a brief overview of the discovery of the ABO blood types and their genetic basis, see (Crow 1993).
setup amounts to a ‘black-box’ containing all or most of the inner workings required by a yet to be elucidated mechanism to produce the phenomenon under investigation; in other words, it is assumed that a mechanism exists within the physical boundaries of the experimental setup. In as much as the above combo of assumptions is not inferred to be true on the basis of their likelihood alone, but is intended as a preliminary framework eventually subjected to testing, we may just as well call it the ‘mechanistic framework hypothesis.’

The mechanistic framework hypothesis does not have to specify putative components and the mode of operation of the mechanism. Rather, it is a more fundamental hypothesis about where in space-time, and which [kind of] physical system or experimental setup researchers should investigate in order to elucidate the mechanism responsible for what is assumed to be a phenomenon (e.g., cell, organism, in vitro blood reaction model). Given this hypothesis, researchers can begin to characterize the phenomenon as an extended web of correlations, formulate more specific hypotheses about the details of the mechanism, test putative components for causal and constitutive relevance (determine whether they belong to the causal pathway connecting input and output conditions), and ultimately elucidate the various components, stages, organizational features composing the mechanism.11

It is the elucidation of a mechanism which ultimately confirms the mechanistic framework hypothesis, and vindicates the underlying assumptions that reproducible events constitute a phenomenon and that the mechanism responsible for generating this phenomenon are to be found within the physical boundaries of the experimental setup within which the phenomenon can be reproduced. Alternatively, if a mechanism fails to be elucidated, then the question remains open as to whether a mechanism exists, whether a statistical model provides a more fruitful description, or maybe an explanation of the putative phenomenon, or whether there is no phenomenon, but only a coincidence that needs no further explanation (Figure 1).

11 Examples of mechanisms and their discovery are discussed in (Baetu 2012; Bechtel 2006; Bechtel and Richardson 2010; Craver 2007; Darden 2006).
5.5 The relationship between experimental reproducibility and mechanistic regularity

Reproducibility is needed in order to determine whether a phenomenon can be detected or measured against a background of chance correlations, and whether the interventions required for elucidating mechanisms have a detectable/measurable effect distinguishable from accidental happenings. Since the requirement for experimental reproducibility entails that a mechanism must produce a phenomenon with a measurable degree of regularity, it follows that a minimal form of mechanistic regularity (i) characterizes mechanisms that have been shown to contribute in an empirically measurable way to a phenomenon or mechanisms that have been elucidated by means of experimentation. Furthermore, the elucidation of mechanisms by experimental means requires interventions targeting organizational features in order to demonstrate their causal relevance. Since the results of these interventions must be reproducible in order to determine whether or not they have an effect on the phenomenon produced by the mechanism, a minimal form of mechanistic regularity (iii) is also implied: organizational features must be present often enough in order for repeated interventions on them to make a measurable difference on the phenomenon. Finally, a point can be made that many mechanistic explanations rely on experimen-
tally demonstrated properties of mechanistic components (e.g., ability to bind, diffuse, catalyze, or, in more general terms, perform a certain mechanistic activity). Just like in the case of organizational features, interventions must be conducted (e.g., mixing, changing concentrations, adding substrates, inducing mutations and other chemical modifications) and reproducibility is required in order to determine whether these interventions have an effect or not. The fact that mechanistic components behave in a minimally regular fashion and that these behaviors are required for the functioning of the mechanisms of which they are parts lends some indirect support to the claim that, to some extent, mechanisms work in the same way in order to produce a phenomenon (regularity ii).

This said, if experimental reproducibility relies on a certain degree of mechanistic regularity, the latter need not amount to a law-like generalization. Reproducibility is not a direct measure of mechanistic regularity, but something determined in part by the intrinsic regularity of the mechanism under investigation, and in part by the researchers’ ability and motivation to conduct large-scale experimentation. In principle, the degree of reproducibility can decrease on condition that researchers have the experimental resources to repeat the same experiment a large number of times. Thus, at least up to a certain point, an increase in the willingness and ability to conduct large-scale experimentation is expected to result in the elucidation of less and less regular mechanisms.\(^\text{12}\)

6. Conclusion

The proposal that an explanation may amount to the description of completely irregular mechanisms raises a dilemma: in as much as it is impossible to distinguish the phenomena produced by such mechanisms from the background of accidental happenings, the mechanistic explanation is empirically equivalent to treating the phenomena in question as accidental happenings. Thus, the complete elimination of regularity is achieved at the cost of the loss of empirical significance. I discuss several ways to solve or circumvent the dilemma, focusing on an experimental solution according to which irregular mechanisms can be elucidated to the extent that the

\(^{12}\) At the same time, it is seems also reasonable to assume that the ability to conduct experiments cannot increase indefinitely and that, in the end, no amount of experimentation can compensate for a complete lack of mechanistic regularity.
phenomena for which they are responsible can be consistently reproduced in a given experimental setup. According to the solution proposed in this paper, experimental reproducibility is necessary in order to distinguish the effects of interventions from the background noise of chance correlations, and to circumscribe putative mechanisms within the physical boundaries of experimental setups. By the same token, however, this solution entails that mechanisms that have been shown to be causally responsible for phenomena are always regular to the extent that they are productive of reproducible events.

Acknowledgments

This work was supported by a generous fellowship from the Konrad Lorenz Institute for Evolution and Cognition Research. In particular, I would like to thank Stuart Glennan for useful discussion and comments on earlier drafts of this paper. I would also like to thank the editor, James W. McAllister, and two unnamed referees for their excellent comments and suggestions.

Bibliography


8885 words