Université Paris 1 Panthéon - Sorbonne, Ecole Doctorale de Philosophie

Sapienza - Università di Roma, Dipartimento di Filosofia

Marco Casali

This dissertation is submitted for the degree of

Doctor of Philosophy

23rd June 2022

Chance Underlying Evolution. Stochastic Explanation in Molecular and Cellular Biology

French Supervisors: Francesca Merlin, Philippe Huneman

Italian Supervisors: Simone Pollo, Nunzio Allocca

Dissertation Committee

André Ariew (University of Missouri)
Marta Bertolaso (Università Campus Bio-Medico)
Elisabetta Affabris (Università Roma Tre)
Thomas Heams (AgroParisTech, Paris)

“Or mentre in guisa tal fera tenzone
è tra ’l fedel essercito e ’l pagano,
salse in cima a la torre ad un balcone
e mirò, benché lunge, il fer Soldano;
mirò, quasi in teatro od in agone,
l’aspra tragedia de lo stato umano:
i vari assalti e ’l fero orror di morte,
e i gran giochi del caso e de la sorte”

T. Tasso, *Liberata*, XX, LXXIII

"It rises: flashes, there's expansion
In a moment more it will be done.
Great aims seem foolish at the outset:
But we'll laugh at Chance itself, yet"
W. Goethe, *Faust,* 6865

Acknowledgements

Although the work of writing a PhD dissertation must take place in creative and exploratory solitude, it is also true that what precedes it, or intersperses it, is social and cooperative work. I thank Francesca Merlin for her help, her firm belief in me, and her beautiful way of doing philosophy that remains a continuous source of inspiration. Her analysis on chance was the starting point and basis of my dissertation, and her constant availability and kindness has allowed me to carry out my doctoral work in the best possible manner. I thank Michel Morange, who at the beginning of my PhD journey motivated me in the development of my work. I thank Marie Kaiser who has been essential in, among other things, helping me establish the methodological barycenter of my analysis. I thank Philippe Huneman, for his valuable comments and suggestions, often at the limits of space and time. I thank my tutors in Rome: Simone Pollo, Elena Gagliasso and Nunzio Allocca, who, like a basket team, supported and followed me with confidence. I thank Elisabetta Affabris, Thomas Heams, Marta Bertolaso and André Ariew who had the patience and kindness to be part of my dissertation jury. I thank Antonine Nicoglou, Laurent Loison, Matteo Mossio and Lucie Laplane, for the precious meetings we had at the start of my journey. A special mention goes to the PhD students and post-docs at IHPST with whom I have shared this difficult and beautiful path. They have offered mutual and unconditional help and support during these years: Marie Michon, Caroline Angleraux, Gaëlle Pontarotti, Matías Osta Vélez and Perceval Pillon. I would also like to thank the PhD students in Rome who, even though I have not been very present in the faculty, have shown me rare kindness and availability: Andrea Olmo Viola, Margherita Bianchi, Leonor Reis and Ryoji Sugawara. A special mention goes to Katherine Waters, who from afar, has accompanied me with kindness and competence in improving my understanding of the beautiful language that English is, helping me create increasingly colourful and consistent linguistic landscapes. I would like to thank Alberto Vianelli who, in the course of our many thought-provoking meetings, revealed to me so much of the fantastic world of molecular biology as well as the challenges and achievements that this branch of science is currently experiencing. I would like to thank Francesca Merlin, Livio Riboli-Sasco, Guillaume Achaz and Perie Leila for welcoming me as a member of their reading group on inheritance. The group's frequent, stimulating meetings greatly enriched the biological content of my work. I would also like to thank everyone I met on this long journey who, in different ways, have contributed to the realization of my dissertation: Roberta Millstein, Charles Pence, Alberto Naibo, Giuseppe Longo, Daniel Nicholson, Derek Turner, Jan Baedke, Francesco Genco, María José Ferreira

Ruiz and Antonio Piccolomini d'Aragona. A special thanks to Ilaria, who supported me as I wrote passages while floating somewhere in the middle of the sea between Civitavecchia and Sardinia.

I also want to thank my brother Christian who has supported me with his sense of calm and amazing music and whose effervescence has always encouraged me. My mother, Adriana, who never doubted me and is ever there, ready for a coffee break. My grandfather, Giulio, perpetually curious to understand what a researcher's job consists of. My grandmother, somewhere in the universe.

Finally, I want to thank Cecilia who first gave me the desire to study philosophy. Without her I would not be here today.

Contents

Introduction

This thesis in philosophy of biology concerns the role of chance in molecular and cellular explanations. It also discusses developmental explanation. In what follows, I will introduce some of the elements necessary to understanding the kind of work the reader will be venturing into: the philosophical and biological context of my research, the main subject of the dissertation, the specific problem I address, the nature of the questions and answers that I propose, and the peculiarities and limits of my analysis. In the first part of the introduction, I do not give a precise definition of the notions that I will use more frequently during my work, that is to say chance, stochasticity, and random(ness). The reason for this is to keep the reader's mind as free as possible from any specific definition while I introduce problems and scenarios that can be understood as prolegomena to my work, that is, prior to asking the question that concerns chance, stochasticity, and random(ness). Only in section 1.6 will I begin to circumscribe the meanings that such notions will have for the development of my dissertation. For now, I can say that chance can intuitively be conceived as an umbrella concept which spans many different meanings. Millstein (2010) demonstrates this point very well, highlighting through four examples how chance already has a rich semantics in everyday speech:

"I'd give that horse a 50–50 chance (i.e. degree of belief) of winning.
I hope I get the chance (i.e. opportunity) to see you.
I found this great restaurant just by chance (i.e. accident).
I'm sorry, but I just can't take the chance (i.e. risk)" (p. 426).

Likewise, we will see that even in biology the meanings ascribed to "chance" are very disparate. However, its semantic pluralism[1] won't be the main topic of my work. The core of my work will rather be chance as an element of our descriptions and explanations of the biological world. In other words, chance as an epistemological property rather than an intrinsic feature of extant biological systems.

[1] On this point see for example Diacoins and Skyrms (2017).

1. First turning point: from evolution to biomolecules and cells

Since Darwin first articulated the theory of evolution by natural selection ([1859] 1964), chance in biology has been considered as intimately connected with the source of heritable variation.[2] In line with Darwin's thinking, the Modern Synthesis, developed in the 1940s and 1950s, invoked chance to characterize genetic variation via random mutations, as well as the evolutionary processes of genetic drift. Philosophy of biology as an autonomous research field within philosophy of science emerged more or less ten years later, that is, in the 1960s and 1970s. When it began addressing chance in biology, around the 1980s and 1990s, it inherited this interest in reflecting on evolution. Indeed, researchers focused on philosophical issues raised by chance in evolutionary biology and, more specifically, in evolutionary theory (ET). From that point on, a set of philosophical and historical reflections on the study of chance in evolution followed. For example, in the 1990s and 2000s, a key debate erupted between philosophers asking why ET is a probabilistic theory (Horan 1994; Millstein 2000a). Shortly after, the debate shifted to more specific metaphysical questions: whether the evolutionary process itself is deterministic or indeterministic (Brandon and Carson 1996; Graves, Horan, Rosenberg, 1999), and whether natural selection and random genetic drift can be conceptually and/or empirically distinguished (Beatty 1984; Millstein 2002). From there onwards, many other reflections on chance have been proposed through the analysis of important key concepts in evolutionary biology. These include random mutation (Millstein 1997; Merlin 2010; 2016a), fitness (Drouet and Merlin 2015; Rosenberg and Bouchard 2015), probability and propensity (Cerredo 2020; DesAutels 2014), evolutionary chance (Eble 1999; Shanahan 1989, 1991, 1992, 2003), and contingency (Beatty 1984, 1992, 2006a, 2006b; Cerredo 2020; Gould 1989; Losos 2017; Wong 2019).

Even though this list is not exhaustive, it is nonetheless representative in showing that philosophical debates on chance in biology have been mainly focused on evolutionary theory, and more generally, on the framework of evolutionary biology. But what about chance *underlying* evolution? Let's look at what I mean by "underlying". This work focuses on domains in which chance is not necessarily reducible to its evolutionary meanings, or at least, its meaning is richer than its definitions from an evolutionary point of view (e.g., as "source of

[2] The study of what Darwin meant by chance is rich and far from providing a unequivocal answer to this question. It has often been argued that Darwin referred to different notions of chance to characterize the origin of variation (see, e.g. Millstein 2011). But there are also researchers who aim to show the other way around. For example, in "Darwin's Dice" Johnson (2015) stresses that, throughout all his writings, Darwin never changed, his belief in the chancy nature of variation (see also Merlin 2016b). Instead, Hodge (2016) defends the idea that Darwin always conceived of chance in the same way, namely as an ignorance of causes.

undirected variation" or as "indiscriminate sampling process", cf. Millstein 2011). First, by "underlying" I do not refer to the dismissal of evolution, but to a mode of enquiry that examines what chance could mean at a finer-grained spatial-temporal scale. Evolutionary biology typically, but not exclusively, focuses at the level of evolving natural populations. This project aims to explore the physiological individual level[3] where molecular and cellular processes, which are underlie the main evolutionary exploration, take place.[4] Second, "underlying" in my framework also indicates metaphorically the willingness to make a study that frames chance with epistemological tools other than the classical evolutionary ones. Looking "under" evolution means analyzing chance from an epistemological point of view which, even starting from the evolutionary framework, tries to propose something new.

I cannot but agree on the importance of continuing to pursue philosophical work that investigates and explores the various notions of chance in evolution. However, I also think that the study of chance should not be limited to its strictly evolutionary analysis because this could lead to an underestimation of the richness of philosophical issues, not always directly related to evolution, that chance can raise in other biological fields. Sarkar and Plutynski have already denounced this tendency in philosophy of biology to underestimate biological research fields that are not directly concerned with evolution:

''Traditionally, evolution has been the focus of most philosophical attention. While it surely remains true that 'nothing in biology makes sense except in light of evolution' (Dobzhansky 1973), this tradition within the philosophy of biology is myopic insofar as it ignores much—if not most—of the work in contemporary biology'' Sarkar and Plutynski (2008, p. xviii).

That contemporary philosophers typically do not find the molecular aspect interesting has also been argued by Pradeu (2017). He analyzed all the papers published by the *Biology and Philosophy* Journal from 2003 to 2015[5] and all the scientific papers published by the *Proceedings of the National Academy of Science of the USA* (PNAS) during the same period of time. Comparing these data, he found a significant mismatch between the biological domains

[3] Physiology can refer both to the processes of respiration, digestion and blood circulation in animals and to the processes of membrane transport, photosynthesis and water exchange in plants. However, there is a physiology specifically called cellular physiology that deals with all intra- and inter-cellular processes. I refer mainly to this last type of physiology, unless explicitly stated otherwise.

[4] Of course, there are important evolutionary works that focus attention on evolution at the molecular and cellular level (e.g., Nei and Kumar, 2000). But what I am specifying here is that I want to provide a philosophical analysis of molecular and cellular biology that is not centred on evolution, that is to say from the perspective that is usally called funcitonal biology (Mayr 1961).

[5] A similar work was made by Gayon (2009) for the period between 2003 and 2015. Pradeu also reports Gayon's results (see Pradeu 2017, p. 153).

explored by philosophers of biology publishing in *Biology and Philosophy*, and those covered by biologists[6] publishing in PNAS (Pradeu 2017, pp. 150-151).[7] More specifically, he found that 62 % of philosophical papers are devoted to evolution (i.e. *Biology and Philosophy*) but only 5% of biological papers (i.e. PNAS journal). He also found that molecular and cellular biology are at best marginally mentioned, and at worst completely absent, in the literature in philosophy of biology[8].

These data might suggest that biological fields, other than the evolutionary ones, might *in fact* be less interesting from a philosophical point of view. There are fewer works *because* topics other than evolutionary ones are not really interesting from a philosophical perspective. But this reasoning does not hold if we focus on recent studies which show the exact inverse, for example, Merlin's (2016a) philosophical reflection on the stochasticity of genetic mutations at the molecular level, Baedke's (2018) work on epigenetics, Pradeu's (2011) problematization of the concept of self in immunology, Théry's (2016) analysis of non-coding RNAs, Laplane's (2016) reworking of the properties of stem cells, etc. Although this list is far from complete, it still allows us to counter that non-evolutionary topics are fertile ground for philosophical reflection.

Still, we might now ask whether it could be interesting to elaborate a philosophical analysis about chance underlie evolutionary biology. Just looking around, it is clear that this kind of analysis is very rare in philosophical reflections, and leaves an important conceptual gap unexplored that biologists themselves often underline. For example, Heams (2012) writes about a "schizophrenic biology" (p. 9). If in evolutionary biology chance is generally conceived as a source of new variation, in molecular biology it is by contrast often conceived as "error"[9] or "noise", namely a source of nuisance for (biological) systems. But this paradoxical situation does not exclusively concern the comparison between evolutionary and molecular biology. In immunology, for instance, chance plays a central role in explaining the synthesis of antibodies. But the entities with which this domain deals are biomolecules, the same entities approached by molecular biology! More generally, how is it possible that chance is so differently conceived,

[6] There have been attempts to bridge this gap between science and philosophy, even with interdisciplinary PhD programs (see e.g. Boniolo and Campaner 2020).

[7] However, I must point out that one of the main limitations in Prader's work is the fact that the *Biology & Philosophy* journal alone may not be representative of all the work in philosophy of biology done between the years 2003 and 2017. This could be true also for the fact that its founders, Ruse and Sterelny, were specifically interested in evolution, which may have influenced the general tendency of the journal. Nonetheless, this limitation does not exempt us from being able to say that, beyond precise statistics, there was indeed a certain tendency to prefer evolution in the philosophy of biology which, to some extent, still persists.

[8] For a similar bibliometric work on the topics of papers published in *Biology and Philosophy* between 1986 and 2017, see Malaterre et al (2020).

[9] Monod (1971) was one the first to clearly claim that evolution is possible thanks to errors in DNA replication.

with contradictory (negative and positive) connotations, throughout different biological research domains? How can we explain this conceptual heterogeneity?

In this context it might be useful to distinguish two points that could be pursued in parallel, namely 1) the meaning of chance and 2) its role:[10]

1. There are few philosophical analyses of the meaning of chance in molecular and cellular biology. What can be said about this? Can we reduce it only to the evolutionary meanings that Millstein (2006, 2011) proposed to us?[11] Or, by contrast, **does chance have specific meanings in cellular and molecular biology?**
2. With the exception of immunology, chance in molecular and cellular biology is often conceived as error or noise and therefore not as part of the explanation of biological systems (the reverse is true in evolutionary biology). This raises the question: **can chance have an explanatory role in molecular and cellular biology?** And if yes, to what extent and in what ways can chance be explanatory?

These two points nicely summarize the focus of the whole dissertation. Chapters 1, 2, 3 and 4 will mostly focus on answering the second question, regarding the explanatory role of chance and which corresponds to the core of my research. The first question concerning the meaning of chance will be the focus of Chapter 6. In it, I will suggest a definition of chance through a specific case study. Chapters 5 and 7 will address both questions.

1.2 Second turning point: from theories to explanation

Significant philosophical literature on chance between the 1990s and 2000s focused on answering why evolutionary *theory* is probabilistic. Two main positions were discussed: ET is probabilistic since the evolutionary process is inherently indeterministic, or, since we are ignorant of the details of the process that, if known, would enable us to account for deterministic phenomena, ET is probabilistic[12]. In order to understand why these philosophical reflections are focused on the *theory* of evolution, rather than on evolutionary models and/or explanations, let us take a step back and look at the context in which philosophy of biology emerged during the 1960s-70s as an autonomous research field.

[10] Interestingly, Millstein (2016) focuses on three different roles that chance can play in evolutionary biology: "an explanatory role, an instrumental role, a representational ("realist") role, and a justificatory role" (p. 685). As will become apparent, I focus only on the explanatory role of chance in molecular and cellular biology.
[11] On this point see also Coffman (2014), Lenormand *et al* (2009) and Wolfe (2012).
[12] Millstein proposes and agnosticism position regarding this point. See Millstein (2003).

Philosophy of science has often been interested in physics, in particular during the first half of the 20th century, when physics was considered the paradigmatic science. This interest culminated in the development of philosophy of physics that typically reflects on theories since there are plenty in this domain of study. Quantum mechanics and general and special relativity are two exemplary pillars of philosophical reflection on physical theories.

In biology, there are fewer theories. Actually, there are three (Hoquet and Merlin 2014): evolutionary theory (ET), which has been the center of attention since Darwin proposed it in 1859; cell theory; and the theoretical framework from which molecular biology was founded (i.e., the central dogma, if we consider it as a "theory"). Notwithstanding this, Nicholson and Gawne (2015) have shown that logical empiricism, a philosophical current born in the first half of the 20th century within the Vienna Circle, was an attempt "to make biological *theories* more like their physical counterparts by forcibly fitting them into rigorous deductive systems" (Nicholson and Gawne 2015, p. 347; emphasis added; see also Callebaut 1993). More specifically, echoing Wolters (1999, p. 195), Nicholson and Gawne specify that this "forced manner" of doing philosophy of biology that was pursued by logical empiricists consisted in applying three theses that are actually misguided when applied to biology rather than physics: "antimetaphysics" (the thought that any metaphysics is to be avoided in philosophical reflection), "reduction" (the belief that biology is reducible to chemical-physical explanation) and "physics as model science" (the conception that physics is the model science that biology should try to emulate) (Nicholson and Gawne 2015, p. 350). This last thesis is the one that most concerns us in the present section. Without getting into the fascinating historical debate regarding the roots of the philosophy of biology (Nicholson and Gawne 2014, 2015; Pradeu 2018), what I want to underline here is that 1) the tradition in philosophy of science may have influenced the focus of philosophy of biology on theory – the logical empiricist third thesis is an argument for that, and 2) the theory of evolution is THE theory in biology and many philosophical works have focused on the various issues it raises.

In this sense, there was then, in the 1960s, a methodological break proposed by Hull (1969) who instigated a move away from this way of doing philosophy of science when dealing with biology. In order to develop philosophical reflection relevant for science, he suggested looking at what real biologists do. The much-quoted text is one in which Hull claims there is much to do from a philosophical point of view with respect to issues integral to biology but that

at the time of writing, 1969, nothing had yet been done[13]. Pradeu (2018) enthusiastically states that Hull's desire was in fact fulfilled in the following decades: "Hull's wishes have been answered, better than he could have hoped" (p. 454). Today philosophy of biology is a well-formed and rich branch of "bottom-up" work, that is, work which looks closely at scientific practice to extrapolate interesting philosophical questions that are potentially relevant to science (cf. Chapter 1). But what intrigues me at this point is another less famous quote which directly rails against the obsession of the logical empiricists to axiomatize all sciences, including biology:

"Too often the applications of mathematical logic to problems in biology give the impression that more or less commonplace ideas have been expressed in tiresome exactitude when they could have been conveyed more easily and more directly in a few sentences of plain English" (Hull, 1969, p. 173).

Here, Hull is clear: he critiques the logical empiricists' attempt to muscle biology into formal and mathematical axioms. This was a major breakthrough that led philosophy of biology to open up to different methodologies that were more attentive to biological practice and not channeled by the narrow perspective of working only on theories (cf. Chapter 1).

From what we have said so far, could we affirm that the daily practice of biology is not about formulating/constructing theories, but models and explanations instead? (cf. Waters 2014). This would be too strong a statement, but something less straightforward can be said about it. Taking any biological textbooks, we can see that what matters most in biology is explaining processes: a textbook of molecular biology aims to explain, for instance, DNA duplication, the process leading to gene expression, post-translational modifications, etc. A textbook of cellular biology aims to explain, for example, mitoses, meiosis, cell differentiation, and so on. In the same vein, one of the primary aims of articles on biology is to explain the last empirical observations in biology (though of course we can also find descriptive papers). Titles such as *Explaining Gene Expression Using Twenty-One MicroRNAs* (Asiaee *et al* 2020) or *Can Genes Explain Biological Complexity?* (Szathmáry *et al 2001*) stand testament to this tendency.

But an objection arises: biology provides several theories instead! As already mentioned, there are at least three: the theory of evolution, cell theory and the central dogma. Moreover, saying that biology focuses on explanations and models and not on theories could

[13] "From what has been said thus far, the reader might infer that philosophers have very little to contribute to biology. On the contrary, there are many aspects of the scientific endeavor to which philosophers could contribute. My complaint is that by and large they have not" (Hull 1969, p. 171).

stimulate the curiosity of the reader concerning what I mean by these three categories. But taking the path of defining a model and a theory would take too much time, and is not be essential for this dissertation (even if, on the other hand, I will spend a lot of time defining stochastic explanation). What I am interested in is only specifying that, to the extent that we talk about chance in molecular and cellular biology, it is *de facto* present mainly in explanations and models, whereas by "explanation" I simply refer to any sort of answer to a why question, and by "model", any kind of diagram and/or qualitative/quantitative representation that aims to describe and explain biological processes.[14] By noting that chance is most of the time mentioned in explanations and models allows me to take away the burden of saying that in biology *in general* there are only explanations and models.[15] In the present work, I focus mainly on explanations postulating the assumption that often models are built *to provide* explanations. For instance, we will see that a model for cell wall expansion is proposed to explain the spherical shape of certain plant cells. But in order to avoid always mentioning explanations *and models*, from now on I will write mainly about explanations. Discussions of models, if relevant, will eventually follow.

Last but not least, my focus on explanations (rather than theories) seems particularly relevant and fruitful with respect to the kind of philosophy of biology characterizing my research. Love (2008) writes: "[the] relation between philosophy and science can be beneficial but is inherently precarious because it requires maintaining a *tension* between conceptual proximity to scientific practice and interpretative distance from philosophical reflections" (pp. 66-67; my emphasis). For at least two reasons, explanation presents a perfect topic for finding a good balance for this tension between philosophy and biology. The first is that explanation enables me to work in proximity to scientific practice, since explanation is one of the most widely developed and used categories in biological practice discussing chance. This might seem trivial but it is not, since it traces a precise position with respect to the type of philosophy that I propose with the present work. As will become clearer in Chapter 1, a philosophical project focused on science has to take what is going on in biology seriously and seek to be as close as possible to its practice. That position, which is already codified in the universe of philosophy of science, is called philosophy of science in practice (PSP; Ankeny 2011). This way of doing philosophy is a clear positioning against a "top-down" philosophy that is less sensitive to

[14] See Vorms (2018) for a discussion on the nature of theories and models. See Braillard and Malaterre (2015) for a discussion of biological explanation.

[15] Flipping the argument around, another objection might be that physics also develops models and explanations. This is undoubtedly true. However I find this objection less strong since, as I have already argued, philosophy of physics and the attitude of the logical empiricists is to rely on theories anyway.

contemporary biology (or science), and instead takes its greater strength from *a priori* philosophical arguments (cf. Sterelny and Griffiths 1999, xi; Turner 2011a, p. 12). The second reason is that, focusing on explanation, I have a very wide background of philosophical reference. For more than 50 years, scientific explanation has been subject to careful philosophical critiques and discussions: from the first model of Hempel and Oppenheim (1948) up to the very final proposals by Strevens (2008) and Woodward (2003), scientific explanation has been carefully analyzed with a heterogeneous set of philosophical accounts. What I am trying to suggest is that the topic of explanation in the study of chance could be seen as an excellent meeting ground between biology (which works all the time with explanations) and philosophy (which has been studying explanations for more than 50 years).

Having specified that my subject is chance in molecular and cellular biology from a non-evolutionary perspective, and that its frame is scientific explanation, I would like to paint a more complete picture of the kind of epistemological work that I will carry out, starting from these two elements. The next section focuses just on that.

1.3 The epistemological nature of my work

In cellular and molecular works, chance is often conceived as error or noise, and remains most of the time on the sidelines, seen as an epistemological obstacle,[16] that is to say, a problem for research progress.[17] More specifically, the very general question often addressed in biological literature is: what is the meaning and role of chance in biological *phenomena*? The works that emerge from this swing as a pendulum between cognitive limitation and ontological reductionism. In biological literature, we can find two main meanings for the label "chancy" when it is used to characterize a biological phenomenon. The first refers to the fact that it is simply too hard to unpack all the specific dynamics involved in the phenomenon under examination. In other words, when figuring out what is going on in phenomena, or how they produce certain outcomes becomes problematic, it is often said that "this process happens by chance". In this signification, chance refers to our cognitive limitation of the knowledge, explanation and understanding of certain phenomena. On the other hand, in cellular and molecular biology, chance is often also described as a physical process (e.g., thermal agitation of molecules) that influences the behaviors of biological phenomena. This second signification

[16]Bachelard (1934, p. 105) also writes about an "epistemological obstacle", but the meaning is different from that proposed in the present discussion.

[17] Although in recent years biologists are shifting paradigms by asserting that noise may play a role in molecular and cellular biological processes (cf. Chapter 5).

of chance is committed to a sort of ontological reductionism, whereby chance at the biological level is nothing but the thermal agitation of molecules at the physical level. I would like to escape this pendulum swing. In the present work, I will propose an idea of chance as an essential element for biological explanation rather than using it to emphasize our cognitive limitations or to reduce it to the physical phenomena underlying biological processes (cf. Chapters 5 and 6).

The philosophical challenge of this present work is to overcome this conception of chance as the "enemy" of knowledge in cellular and molecular biology. I am not the first to try to restore this sort of "dignity" to chance. In the interface between physics and biology, among others, Bravi and Longo (2015) and Buiatti and Longo (2013) try to give back a meaning to chance in terms of functionality which can extend to all levels of biological organization.[18] From a purely biological point of view, Kupiec (2009a, 2009b) proposes extending Darwinism as far as inside cells, endowing chance with a core role in the functionality of most biological phenomena. One of the books he edited is even titled *Le hasard au cœur de la cellule*. All of these works represent important contributions to the debate and I will have the opportunity to engage with some of them in Chapter 4, where I will examine certain accounts already proposed and discussed in the philosophical and biological literature.

But the thing I would like to specify at this point of the introduction is that my contribution is more modest and circumscribed compared to the works just mentioned. It is not my intention to carry on both an ontological and an epistemological argument about chance in molecular and cellular biology (cf. section 1.4). My contribution is to focus on biological explanations in these disciplinary domains, and thus focus on the epistemological side, 1) arguing that chance can have a central role in the explanation of, at least, some molecular and cellular processes, and 2) showing *why* chance can have such fundamental role. Even if the first point is already addressed by researchers such as the ones mentioned above, nonetheless they often oscillate between ontological and epistemological arguments, namely between the functional and explanatory role of chance (cf. section 1). By contrast, my work aims to remain mainly at an epistemological level, addressing the question of the role of chance in explanation, that is to say its explanatory power (see after). The second point, which concerns why chance can have a fundamental role in the explanation of some biological processes (in particular at the molecular and cellular level), is seldom addressed. To do so requires openness to a kind of philosophical exploration that seeks to go deeper by challenging trivial answers. Indeed, one

[18] A development of this idea can be found in Montévil *et al* (2019; see Chapter 4)

could make this second question trivial by answering that chance is essential because the explanation in which it is present just works. Although I would agree with this answer because it presupposes that chance has a role in some biological explanations (which is nonetheless important part of this work), the answer is not philosophically satisfying. Indeed, the question remains: *why is chance essential* in some biological explanations? There is a subtler philosophical terrain here that has not yet been fully discussed. What does chance possess to make it so effective in explanation? In other words, from where does the explanatory power of chance stem?

The philosophical exercise I am proposing on this issue reminds me the classic game of "Why?" that often takes place between parents and children. "Why is the tree green?" "Because there are leaves." "Why are the leaves green?" "Because there is chlorophyll." "Why is chlorophyll green?" "Because it is a molecule that reflects green light and absorbs red and blue light." "Why do the molecules reflect green light and absorb red and blue light instead?" "Because the electrons inside the molecules are excited and go to higher energy levels," and so on. In developing this project, I felt like the kid who keeps asking why. I think this is a good strategy for exploring the philosophical implications hidden in the ways whereby biologists handle the natural world.

Another specification that I have to make concerning the epistemological nature of my work is that I am not proposing a new account of scientific explanation to add to the already long list proposed in the literature of philosophy of science (see Braillard and Malaterre 2015; Brigandt 2013; Newton-Smith 2000). In addition, I want to remain – as much as possible – neutral on the question of a *unique* theory of explanation. As well as many other authors (e.g. see Brigandt 2013; Kellert, Longino, and Waters 2006; Ruphy 2016), because epistemic interests are varied and each researcher/team of researchers may base their inquiries in different ways. There is no reason – and the philosophers just mentioned argue this point very well – to think that biological processes can be explainable by using a single theory of explanation. In this vein, I am sympathetic with Morrison (2009) when she says: "[w]e simply cannot specify in advance what qualifies as an explanation or what form it will take" (p. 124). I read in this line the message that, in light of the heterogeneous ways in which explanations can be provided, it is simply impossible to give an *a priori* standard that could fit all of these heterogeneities

What I want to propose is, on the one hand, a reflection that aims to make the reader able to recognize the value of chance in certain explanations and to reevaluate the epistemological judgment about it. On the other hand, I would like to suggest that, at least in certain circumstances, it could be useful to reframe certain explanations and propose new ones, using the notion of chance without fear.

In this section, I have been explicit about the kind of work I will be pursuing, namely an epistemological enquiry that centers its attention on the role of chance in scientific explanations. Although I have specified that I will not pursue work on the ontology of chance, why did I make this decision? Why do I not focus on ontology too? In the next section, I justify why, in my work, I want to keep the epistemological and ontological dimensions as separate as possible.

1.4 The ontological-epistemological separation

My analysis is not a metaphysical/ontological[19] discussion of chance. I do not want to ask about the extent to which chance exists and the forms it takes in the biological world at the molecular and cellular level. Nor do I wish to examine the ontological definitions we can give to chance in order to make sense of the biological processes in which it operates, or to propose another positioning with respect to the debate about the deterministic vs indeterministic nature of biological phenomena. Of course, biological systems are physical systems which are inevitably influenced by chance at a physical level in terms of thermal agitation and quantum phenomena. Moreover, it is now well known that biological systems handle, and sometime even exploit,[20] physical chance depending on the circumstances and reference phenomenon. Showing how chance can be exploited by biological systems is an empirical question, not a philosophical one. Nonetheless, as Kaiser (2015) points out "[t]here is more to be said about ontology than "We're all ontological reductionists. Case closed" (pp. 92 – 93). Translating to our context: there is more to be said about the ontology of chance in molecular and cellular biology than the mere fact that it is due to thermal agitation and quantum effects. I am completely in line with this thought. Thus, I trust that future work can make richer ontological contributions on chance in biology.

[19] For a discussion of the relation between metaphysics and ontology, see Chakravartty (2017); between ontology and realism, see Fine (1984).

[20] In this context, Bravi and Longo (2015, p. 10) and Buiatti and Longo (2013, p. 20) write about "canalized randomness".

But actually, looking around, one realizes that chance as thermal agitation and quantum phenomena are not the only ways in which an ontological discussion of chance has been proposed. There is, at least, another cluster of publications in the philosophical literature that put aside investigation into the nature of chance to ask instead how the probabilities that are in fact used to describe and measure it should be interpreted. I identify this area of study as metaphysical/ontological reflection because it questions the very nature of probability, asking "What are probabilities?" If chance is out there, what do probabilities describe and measure? By contrast, if chance is not out there, can we say that we use probability only because we are limited human beings, unable "to carve nature at its joint" (see Campbell *et al* 2011; Khalidi 1993; Velasco 2012)? Many answers have been given with respect to what probabilities are; the classical, logical, subjective, frequentist, Bayesian and propensities interpretations (see Galavotti 2005; Gilles 2000; Hacking 2006; Mellor 2005) are the main (but not the only ones) examples of this richness. All in all, no single account has tied up the debate. To the extent that the debate is still open, why does my project not take a position? An initial answer could be that the rich work that has already been done discourages me from attempting any sort of placement – but this is a superficial reason. The deeper, real reason concerns the aim of my research. In order to argue for the explanatory role of chance in biological explanation, I find a more – I hope – convincing argumentative strategy by not closing myself off in the debate on the interpretation of probability. I will show that the notion of chance can have explanatory power for some biological explanations. Therefore, my focus will be on the *notion* of chance and not on the *way in which it is described and measured* (i.e. probability; see e.g. Feller 1968).

But one might ask whether probability might also have some explanatory power. As will become clear in the course of this dissertation, my focus will be primarily on certain explanations that we can find in biological textbooks and scientific articles in contemporary molecular and cellular biology. These explanations do not use probability. Indeed, they are developed with the aim of qualitatively explaining the development of a certain process (and not to measuring their chancy character). For example, the explanation of the spherical development of cell walls mentioned above does not use probability. It only specifies that microfibrils (which are the main components cell wall is made of) are randomly oriented (cf. Chapter 3). This reference to "random" orientation of course proves that a notion of chance is engaged in this explanation. It will be on these kinds of notion of chance (without probability) that I will focus in the chapters which follow.[21]

[21] Nonetheless, nothing prevents the explanations that I propose in the present dissertation from being eventually formulated in probabilistic terms. But for the reasons specified above, that is not what I am interested in.

Of course, my study of chance does not exclude the possibility of studying the role of probability in certain explanations – which doubtless would be of great interest.[22] But, simply put, this would be *another kind of* study, different from the one I set out for this dissertation. It is one thing to ask about the role of chance in the explanation of cell wall development and quite another to ask what measure of probability can be attributed to the possible orientation of the microfibrils and how this probability can be explanatory, if at all. These two questions highlight two different research agenda. This work aims to focus on the first.

So, now we have ascertained that I am interested in developing an idea of explanatory chance that does not incorporate probability. How then do I intend to develop this work? Is it an exclusively epistemological work? If so, what types of questions could be developed? I have already found my position with respect to explanatory chance, could I now be more precise and connect this with my methodology for developing a project in philosophy of biology?

My idea is to shift from a question concerning the meaning and role of chance in biological *phenomena* to a question on the meaning and role of chance in biological *explanation* by reflecting on *explanans-explanandum* schema. This shift is crucial to understanding my research. But even from such a quick presentation of my project, yet another complaint arises right away with respect to ontology. Does saying that chance can have a central role in biological explanation mean that it could be central for biological phenomena as well? In other words, is the epistemological-ontological separation tenable or will it inevitably collapse at some moment in the argument? For example, if I say that chance has a central role in DNA replication, am I saying that, out there (by which I mean not only in the explanation *but in the phenomena*) chance plays a functional role in this molecular process? In other words, through my epistemological analysis, am I also proposing a new ontology of chance? This is a very tricky aspect of my research since many philosophers do not keep the epistemological and ontological dimensions separate. They argue that to do so is exhausting as well as unnecessary. For example, Rosenberg writes that combining an *ontological* reductionism with an *explanatory* anti-reductionism leads to an "unstable equilibrium" (Rosenberg 2006, p. 7; For a discussion see also Brandon 1996). This then leads me to wonder whether this idea of keeping ontology and epistemology separate inevitably leads to a balance that is difficult to maintain. Despite my interest in epistemological analysis, do epistemology and ontology *in the end* need to go together? At the end of the day, is Rosenberg right?

[22] Many studies have already been carried out in this direction (e.g. DesAutels 2014, 2015; Millstein 2000a; Popper 1959; Suárez 2020).

I do not have a convincing counter argument and I am actually sympathetic with this general skepticism of the separation. It could be reasonable to say that the way in which we understand and account for the world goes parallel with how the world actually is. I do not want to develop a sophisticated account of scientific realism but I can say that I am a realist concerning science. I take a positive attitude to believing that the best theories, explanations and models biologists provide contain information that advances our understanding of what the biological world looks like.[23] For example, when we refer to causes, we do not (only) refer to the fact that they can be useful for explaining something, but that these causes are actually out there and by explaining them, we add knowledge to the way we conceive of the world. I guess that we all more or less agree that if we provide any sort of causal explanation it is because we believe that causes, or at least causal histories,[24] are in the world, *pace* to Hume. I think that for the most part explanations are as they are because, to some extent, they try to make sense of how the world actually is. What I have just sketched here is of course a naïve form of realism. But if we take this realism as granted, it is enough to raise an important objection to my stance: from a realist (even naïve) standpoint, why am I still convinced in proposing a separation between epistemology and ontology?

An argument for why I focus on the epistemological versus ontological aspect concerning chance is a kind of *temporal* priority in the development of philosophical research. We must first understand how we describe and explain the world in order to be able to say what the world is like (because our descriptions and explanations are our only way of accessing it). It is reasonable to argue then that it is necessary, *first of all*, to study and investigate the way in which biologists work, structure their work, think and organize their daily practice, develop a certain problem etc. and *only after that* to address the problem of the nature of what they study, which is to say the ontology of chance.[25] For this reason, when I started thinking about developing an analysis on chance in molecular and cellular biology, the questions that came up were (mostly) epistemological. But what are the articulations of these epistemological questions? The following section aims to propose this articulation by dividing my dissertation

[23] The Stanford Encyclopedia of Philosophy, defines scientific realism as "a positive epistemic attitude toward the content of our best theories and models, recommending belief in both observable and unobservable aspects of the world described by the sciences" (Chakravarty). The author adds that scientific realism is heterogeneous but "[w]hat all of these approaches have in common is a commitment to the idea that our best theories have a certain epistemic status: they yield knowledge of aspects of the world, including unobservable aspects".

[24] "Whatever causation may be, there are still causal histories" (Lewis 1986, p. 217).

[25] A note of caution is needed here. That an epistemological work precedes an ontological one is not to say that the former is only preparatory and finalized in order to open the road to the latter (in this case an ontological-based reflection). Underestimating the role of the epistemological work in this case would be a mistake since this kind of work has its value and richness regardless of whether an ontological work follows or not.

into a critical part and a positive thesis.[26] These questions are obviously developments of the two basic questions I proposed in section 1. In cellular and molecular biology, does the concept of chance have a specific meaning? And, can chance have an explanatory role in molecular and cellular biology?

1.5 Articulating epistemological questions about chance: the critical part and the positive thesis of my work

Then, my contribution consists of a critical part and a positive thesis.

The critical part is focused on the following questions. Why is chance conceived of in different ways in different biological domains? Do these different meanings influence the role the concept of chance plays in these various research contexts? A philosophically interesting issue is that heterogeneity in the conception of chance is found not only between evolutionary biology and non-evolutionary biologies, but can also be found *within* non-evolutionary-biologies, as the aforementioned example of molecular biology vs immunologists testifies (cf. section 1). I propose to reflect on molecular and cellular biology because in the last years the role of chance in these kinds of explanations has increasingly come under investigation, thereby enriching the ground of philosophical reflection. More specifically, from an epistemological point of view, I will analyze the ways in which biologists conceive of chance and the ways in which these conceptions actually influence their research. From a historical point of view, addressing specifically molecular biology, I will go over the motivations that led biologists to talk about chance as noise, namely as a source of nuisance for (biological) systems, trying to figure out why today the notion of noise is still used in a negative epistemological fashion.

Concerning my positive thesis, I will explore the two main epistemological questions I already presented above. Does chance have a role *in explaining* the proper function of certain processes? And if so, by virtue of what? In other words, what philosophical arguments can be formulated in favour of chance as explanatory? My aim is to convince that reader that in certain explanations chance can, in fact, have an essential explanatory role. In order to do that, I will provide a philosophical argument that is based on the idea that chance is an abstractor which enables the *explanans* to abstract and synthetize numerous details that could undermine the effectiveness of explanation. What I mean by the "effectiveness" of an explanation (I will use also the term "goodness") will be specified in Chapter 3, where I introduce the notion of

[26] These parts should not be understood in a didactic sense. I will not divide the chapters of my thesis according to this division. However, traces of these two can be recognized in each chapter (some more marked the first, and others in the second, part).

abstraction. The arguments that I will propose as justification are purely philosophical. By contrast, to test this idea, I will use examples that come from molecular, cellular and developmental biology. Then, the question of whether the account I propose in this dissertation could be extended to domains different to those in which I tested it (developmental, cellular and molecular biology) is a question that can be carried into future work (cf. conclusion of the dissertation).

Assuming that these epistemological questions are more urgent in a project on chance in molecular and cellular biology does not preclude that, a more ontologically-centered analysis could follow in future work. Indeed, it seems to me that my epistemological work could give rise to ontological questions different to those mentioned so far (chance as a property of physical systems, the interpretation of probability). Throughout my dissertation I will assert that the explanatory power of chance can be found in the explanation of *at least certain* phenomena and not others. Thus, it appears that some phenomena possess properties that allow them to be explained in terms of chance *whereas others do not*. So, if chance has a role in the *explanation* of certain phenomena, what can we say about them? Does chance explain certain biological phenomena because these phenomena have specific properties? If yes, of what kind? And how can we articulate the relationship between the role of chance in explanation and its ontology in phenomena? These ontological questions are interesting and original. However, they are beyond the scope of this work. But they are certainly good material for future research (cf. conclusion of the dissertation).

1.6 Stochasticity, randomness, and chance: two nominalist and one philosophical characterization

I have extensively lingered on some important issues regarding my project on chance, but said little with respect to the object of my study, which is, to be precise, chance in cellular and molecular biology. In the last part of this introduction, I develop the definition of three notions that will be central for the development of my work, that is to say stochasticity, random(ness) and chance. I would like to circumscribe the notions of stochasticity and random(ness) by directly looking at the biological literature. In other words, I want to see how biologists use these two notions, trace their meaning, and use them in my research context. By contrast, I will provide an original notion of chance.

But before moving on, three disclaimers. First of all, the ultimate goal to this dissertation is not to provide a unified definition of chance[27] but rather to highlight what makes it explanatory with respect to certain molecular and cellular phenomena. The notions that I will propose will help for this main aim. Second, as I mentioned in the previous section, I focus on chance in the explanation of biological processes (vs. their biological outcomes). Third, the original part of my work is not only found in the answer (of why chance can be explanatory) but also in the question itself. In fact, asking if chance does have a role in cellular and molecular explanations is a question that, to my knowledge, has never been formulated so explicitly in a philosophical context.

1.6.1. Stochastic processes

Even if it is not always the case, when we can find the notion of stochasticity in biological literature, it usually refers to a property of processes. For example, we can find statements such as: "[w]e find that start site selection is largely *stochastic*" (Boersma *et al* 2019, p. 459, emphasis added). Here, "start site selection" is a process by which ribosome attaches to an RNA to initiate translation and "stochastic" is the property attributed to this process. Let's take a look to another example:

> "The immune system exploits the inherent *stochasticity* of small numbers in the 'VDJ recombination *process*' which gives rise to T-cell receptor diversity" (Honegger and de Bivort, 2018, p. R10, emphasis added).

VDJ gene recombination is a process that happens during the development of B-Lymphocytes. VDJ recombination involves the part of genes that, once the lymphocyte is mature, will code for anti-bodies. Having VDJ random recombination permits the production of a huge variety of anti-bodies. But aside from these technicalities, that is interesting also in this case is that "stochasticity" refers to the property of a process, i.e., VDJ recombination.

Let us now try to give a more formal definition of what stochasticity means with respect to (biological) processes. For this definition I borrow the lexicon given to the stochastic process from a technical book on probability theory:

[27] This holds even though a particular definition of chance will be proposed for a specific case study in Chapter 6.

"*[S]tochastic* processes [are] systems which evolve probabilistically in time or more precisely, systems in which a certain time-dependent random variable X(t) exists [:]
P (x1,t1; x2t2; x3t3…)" (Gadiner 2004, p. 42; emphasis in original)

A process is stochastic if the variables used to describe it do not vary regularly over time but randomly over a given space of probability in which its multiple possible values are distributed. This kind of variables is depicted in the proposed definition as "X(t)", that is to say a variable "X" that randomly varies according to/with time "(t)". "P(x1,t1; x2, t2; x3, t3;…)" is a further formalization of this idea: "P" represents the probability that at t1 "X" is "x1" and that a t2 "X" is "x2", and so on.

There is an immediate objection to this definition. In section 1.4, I specified that it is not my intention to pursue an account of chance in terms of probability. By contrast, developing this definition of stochasticity in terms of probability might lead to the presumption that I am succumbing to contradiction. It is a legitimate objection. However, it is not the case since this notion of stochasticity will not be used to develop my account of chance in cellular and molecular biology. The function of this definition is only to make clear to the reader what I have in mind when I use notions such as "stochastic process", "stochastic phenomena", "stochastic gene expression", and so on. In a nutshell, I can say that this characterization of stochasticity has a nominalist function: it will be used to identify processes that biologists describe as stochastic (see after, Table 3).

1.6.2 Random results

Let's now talk about the notion of randomness. With respect to the literature, we find that "random" is used in different ways, and to characterize both processes and outcomes.[28] Following Merlin (2009), I choose to use "random" and "randomness" with respect to outcomes. As for "stochasticity", I give the notion of "randomness" a nominalist function: it will be used to identify outcomes that are described by biologists as random, more precisely, results that are most of the time unpredictable, or do not follow regular patterns, but that nonetheless can (eventually) be described with a probability law or distribution (see Table 1).

[28] For example, it can be used *interchangeably with stochasticity*: "[probabilistic cell differentiation] has often been termed 'random' or 'stochastic,'" (Symmons and Raj, 2016, p. 788, emphasis added). But more often, it is used *as a property of a phenomena.* More specifically, "random" is also often used as an adjective before a noun specifying something "[h]aving no definite aim or purpose; not sent or guided in a particular direction; made, done, occurring, etc., without method or conscious choice; haphazard." (OED, Oxford English Dictionary). This is the way in which the term is used when we read "random inactivation of X chromosomes", "random sampling", "random mutation", "random walk" and so on.

Thus, as with the results of the throw of a dice or of a coin, the inactivation of the X chromosome or the onset of mutations are said to be random.

The two notions of stochasticity and randomness have been proposed to respectively characterize processes and outcomes in light of their most widespread description. In the next section, I will follow a different path with the notion of chance.

1.7 Chance is a property of a description of a process

Let me underline again the purpose of my thesis. With a philosophical argument, I want to convince the reader that chance is essential for certain explanations in molecular and cellular biology. To do that, I will propose what I call "stochastic explanation" (SE), a philosophical category that refers to any kind of explanation that includes an explanatory role for chance. But the task of explaining in detail what I mean by this explanation, what it applies to, and what implications it has with respect to chance will be developed in Chapters 2 and 3. The present section aims only to provide a first general definition of the notion of chance that will be at the core of my account of stochastic explanation. It goes as follows:

Chance is a property of a description of a process. Given a fixed set of initial conditions, this property specifies that the described process could be otherwise.

Let's unpack this definition step by step.

1) "*Chance is a property of a description of a process*". The development of this part of the definition is primarily motivated by the fact that, since my work is an epistemological analysis, I do not want to develop it thorough ontological inferences. Therefore, in my framework, chance is not a property of the process itself but a property with respect to how it is described. To this end, I have tried to construct a definition that does not refer directly to the phenomenon, but to the way in which it is described. I postulate that chance is not a property of the phenomenon but rather a property of a description of a phenomenon. "Of a description of" creates an epistemological space in which a work sheltered from any strong inference with respect to metaphysics might be built.[29]

[29] Although I do admit that some metaphysical inferences might pop up during the development of the present work. In fact, how metaphysics and epistemology interact and how they are structured in a philosphical project is an interesting research subject (see Chakravartty 2017) that I cannot discuss in this dissertation but that could provide input on possible future works.

2) "*Given a fixed set of initial conditions* […]". This first part of the definition is essential because if there were no fixed initial conditions, chance would be nowhere: all events would be causal, dependent upon their previous conditions. If the initial condition X were not the case, event Y would not follow. Rather, if the initial condition had been Z, the event would have been Q. This argument is summarized in Table 1.

No fixed initial conditions
Initial conditions X → Event Y Initial conditions Z → Event Q Chance is nowhere: If the initial conditions X were different (e.g. Z), the successive event Y would also be different (e.g. Q)

Table 1

The great advantage of fixing the same initial conditions lies in the fact that we go beyond the notion of causal dependence. In this way in fact we set the scenario to build a notion of case in terms of descriptions of processes that may be different starting from the same initial fixed conditions. Moreover, we must specify that the set of initial conditions that we take are local and not universal.[30] See the Table 2 below.

Fixed initial conditions
Initial conditions A → Event B Initial conditions A → Event C Chance specifies that, given (a set of) initial conditions (A), more than one event is possible.

Table 2

[30] This specification is necessary to avoid to falling into an indeterminist account. In fact, if we took all the past and present conditions of the whole universe as fixed initial conditions and postulated that by "chance" we mean processes that can be otherwise, we would have fallen into a metaphysical position that sees chance as something uncaused, or at least, beyond causal history. I want to firmly reject this metaphysical account.

3) "[…] *this property specifies that the described process could be otherwise*". As different philosophers (Beatty 2006, Desjardins 2011, Lewis 2018; Wong 2020) reminds us, we can all agree that evolution is contingent at least because *evolutionary outcomes could be otherwise*. For example, the hypothesis that evolutionary outcomes could be otherwise is normally associated with counterfactuals "if we had A then we would have B". This counterfactual is based on the fact that history happened in a certain way, and we question what would have happened if things were different (even if they actually were not).[31] This part of my definition of chance takes some elements from the debate on contingency (which took place between the 1980s with Gould, e.g. 1989, up to Turner's latest work 2011a, 2011b) but introduces some modifications. First, since I am focused on repeatable processes, I replace "could" with "can". For example, I will argue that the explanation of DNA replication *can* be described and explained through an important, and hitherto neglected, chance step (see Chapter 3, section 3.3). I use "can" since the process under consideration is DNA replication, namely a process that repeats itself continuously in the course of cellular life. Second, I do not use "outcome" in my definition, as philosophers of contingency do, but "process". I am interested in biological processes and the ways in which we can explain them (cf. section 1.5). I do not want to only explain why the result of DNA replication might be otherwise, for example due to some "error" in duplication (in Monod terms), but I want to be able to explain how the process of DNA replication itself can be otherwise and then produce, as a consequence, different outcomes. For example, I am interested in *how we can explain* that different nucleotides can be added to the newly synthetized DNA sequence through various chemical-physical interactions (process) and not *only* in the fact that these nucleotides are, in the end, added (outcome). Let's summarize the definitions just proposed in Table 3 below:

[31] There exists a whole branch of research called counterfactual history of science that tries to investigate hypothetical effects, had certain scientific events been different (see e.g. Jamieson A. and Radick 2017).

Terms	Definitions
Stochasticity (Nominalist function)	General term that I use to refer to processes/events. I define "stochastic" using the following formal definition: "[S]tochastic processes [are] systems which evolve probabilistically in time or more precisely, systems in which a certain time-dependent random variable X(t) exists […]" (Gadiner 2004, p. 42; emphasis in original).
Random and Randomness (Nominalist function)	General terms that I use to refer to the results of processes/events: they are most of the time unpredictable or do not a follow regular patterns but, nonetheless, can (eventually) be described with a probability law or distribution.
Chance (Philosophical function)	Philosophical terms that refers to the following definition: Chance is a property of a description of a process. Given a fixed set of initial contidions, this property specifies that the described process could be otherwise.

Table 3. Definitions of the notions useful in developing my epistemological account of chance in molecular and cellular biology

These notions are the bases from which I will begin to develop my idea of stochastic explanation (SE) in cellular and molecular biology. This will be elaborated in detail from Chapter 2 onwards.

1.8 The symmetrical structure of the present work

Having specified the topic and purpose of this work, what will be the attitude towards the notion of chance carried out, and the specific meaning of each term and notion at the core of the dissertation's argument, I want to illustrate here that my work has a symmetrical structure. It is composed of seven chapters, the fourth of which is a sort of "interlude" that allows the reader to absorb what has gone before and prepare for what comes after. Chapter 4 also marks a division with respect to the *type* of work proposed. The first three chapters have a strongly philosophical nature, in the sense that they lay out the meta-philosophical preliminaries from which I develop my account (Chapter 1), discuss philosophical literature relevant for my

work (Chapter 2), and present my account of stochastic explanation in detail (Chapter 3). However, this does not mean there will be no biological content in them. On the contrary, the biological content will help draw out new philosophical questions. By comparison, the last three chapters are more biologically-centered. I will describe and analyze in detail case studies from molecular, cellular and developmental biology which further strengthen my account of chance. I will provide an analysis of gene expression which specifically concerns transcription (Chapter 5), alternative splicing (Chapter 6), and translation (Chapter 7). This does not mean, however, that a philosophical component will not be present. Through the biological examples that we will consider, I will show they are perfectly understood as specific instances of stochastic explanation (SE).

Thus, this thesis has a symmetrical structure: the first three philosophical chapters are balanced by three chapters more biological-centered ones. The interlude acts as a fulcrum and holds the six chapters in equilibrium and suspension (see Figure 1).

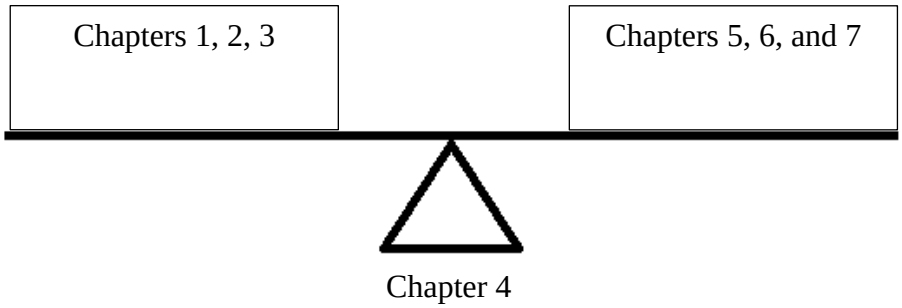


Figure 1. A qualitative representation of the organization of my dissertation.

Part I – Philosophical-centered analysis

Introduction to the first part of the dissertation

1. The structure of the first three chapters

In the first three chapters of the dissertation, I will develop the basis of my original philosophical proposal on chance in molecular and cellular biology (including also some reference to developmental biology) that I will call "stochastic explanation" (SE). The proposal aims to re-evaluate the role of chance in biological explanations.

Chapter 1 reflects on the philosophical preliminaries essential to building a project in philosophy of science in general and philosophy of biology in particular. I will pay special attention to the dimensions of normativity that my project engages with. I think that any project in philosophy needs this kind of clarification. What am I proposing with respect to chance? Which parts of my project are descriptive and which are normative? What kinds of descriptive and normative parts are we talking about?

In Chapter 2, I provide a critique of the notions of mechanism and mechanistic explanation. I will show that the mechanistic framework tries unsuccessfully to annex the explanation of stochastic phenomena. This critical part argues that it is time to go beyond the mechanistic accounts which supposedly explain certain phenomena described as stochastic. Even though this chapter is mainly critical, I nonetheless make a few steps towards my own proposal by introducing the philosophical category of stochastic explanation (SE). The main purpose of this category is to give the reader the philosophical tools to recognize how chance can have an important explanatory role in some biological explanations. I will argue that there are three main criteria of adequacy that specify when a biological explanation can be labelled as stochastic explanation. In this chapter, I will introduce and discuss two of these criteria that I will call "negative", because they specify what stochastic explanation *is not*.

Chapter 3 is the philosophical core of the dissertation. In this context, I will provide the third, "positive" criterium of adequacy which specifies the necessary feature by which an explanation might take the label "stochastic explanation" (SE). In this explanation, chance acts as an abstractor. I will analyze explanations from molecular, cellular and developmental biology that seem to meet this positive criterium as well as the two other, negative criteria.

1.2 Some notes on the genesis of this first part of the dissertation

In Chapter 1, it will become clear to the reader that the type of project I would like to develop is a "bottom-up" project, that is, a project which develops philosophical issues from a close and in-depth comparison to contemporary biological practice. This approach was born in opposition to the "top-down" philosophy of science that proceeds from general, *a priori* philosophical arguments which attempt substantiation through ad doc examples taken from science. The methodology of this "bottom-up" approach is therefore to start from biology and gradually approach more general, philosophical questions. However, in leafing through these first chapters, the reader will find that the order of exposition does not seem to reflect the methodology just proposed. In fact, Chapters 1 and 2 are two strongly philosophical chapters in which, if biological studies on chance are discussed, it is only to confirm the philosophical argumentation. Only at the end of Chapter 3, after developing my account of stochastic explanation (SE), do I propose three biological case studies to test my account. In light of that, it could then be objected that, even though a "bottom-up" methodology (from biology to philosophy) is being preached, in fact a "top-down" work (from philosophy to biology) is being proposed. How can this (appropriate) objection be answered? Is it actually the case? At the end of the day, am I providing a top-down philosophy of science project? No. My work, despite appearances, is a bottom-up project in philosophy of biology. Let's see why.

The key is underlining that the sequence by which this PhD was written does not correspond to the order of its development. The preliminary development up of this work (i.e. prior to its first draft) consisted of reading publications (research articles and textbooks) in molecular, cellular and developmental biology. This immersion lasted more or less a year. Indeed, my idea of stochastic explanation (SE) came from careful analysis of biological explanations, especially from textbooks in which notions such as "stochastic", "stochasticity", "random", "randomness", "chance", "chancy", etc. are found. From this careful analysis, I extracted an interesting philosophical question: could it be that these explanations are "good" (demonstrated by their presence in articles and/or textbooks), not in *spite* of but *thanks to* chance? And, more generally, could the notion of chance play a role in some biological explanations? As should be clear, my philosophical idea was born *as a result of* this immersive analysis of scientific literature.

But if this was the order of development (biology → philosophy), why did the writing progress the opposite way around (philosophy → biology)? The answer must be sought with respect to a simple methodological reason. Although this project is inspired and brought to light by careful biological analysis, it remains a philosophy of biology project. The philosophical

idea that has arisen from reading scientific works must be refined, grown and matured into a robust philosophical account. In order for the ideas "bottomed-up" from science to be credible, they must be strengthened by a solid theoretical structure. This is why, before elaborating and showing reflections that are bottomed-up from biology, I must first create a robust philosophical scaffolding that can accommodate, in the right way, these reflections. This is so important to me that not only do I start my work in developing this structure (which I hope will be solid), but I even problematize the meta-structure of a philosophical project itself. What does it mean to do a project in philosophy of biology? How should such a project be undertaken? What is biological what is philosophical? How normative is philosophy with respect to biology and what role does the latter play with respect to the former? So, let's start with these meta-philosophical questions.

Chapter 1: Meta-philosophical preliminaries

Introduction

In order to develop my project on chance in cellular and molecular biology, I need a clear methodology, and to make explicit the factors that govern the normativity I implement. I provide the entirety of the present chapter for these two aims. The main reason I make this point so extensively is that my project is interdisciplinary and spans the research domains of philosophy and biology.[32] I will analyze biological works, using philosophical tools to look for conceptual issues.[33] A particular difficulty for this kind of project lies in giving a strong and informed biological background while at the same time providing a philosophically-relevant analysis. As this work is intended to be read by both philosophers and scientists, it is necessary to provide a work that could be of interest for both. In short, this particular kind of interdisciplinary work is delicate and it is important to be explicit about the meta-philosophical preliminaries.

Philosophy of biology deals, clearly, with both philosophy and biology. But how they should be combined, what they have in common, and what their differences is not so clear-cut. My project comprises a descriptive and a normative dimension. In the former, I describe biological phenomena; the latter is philosophical and deals with normative and self-reflective statements. One could ask in what sense and to what extent my project could be useful and of interest for both philosophers and biologists. To answer the question, I think it is more important to ask what kind of philosophy of biology I do and, more generally, what philosophy of biology is.

The aim of this chapter is to make explicit the methodology that I use and highlight the normative dimension I embrace. Such specifications are usually spelled out over a few lines or paragraphs in the introduction of a book. However, I follow Kaiser (2015) in providing an entire chapter on the meta-philosophical preliminaries, as they are essential to strengthening my analysis.[34] Actually, this chapter is built on her proposals on normativity. I go through all of the categories she developed to consider which one my research project is concerned with.

---

[32] More specficially, interdisciplinarity means "of or pertaining to two or more disciplines or branches of learning; contributing to or benefiting from two or more disciplines." Oxford English Dictionary

[33] One of the aims of the present chapter is to unpack and make clear this statement.

[34] In this regard Kaiser (2015) writes that "[a]lthough there is a long-standing debate about how philosophy in general is to be carried out (called meta-philosophy), *the specific question of how to do philosophy of science is rarely extensively debated or pursued in its own right*. Rather, the meta-philosophical remarks one can find are mostly located in introductions of monographs, in which the prior aim of the author is to argue for certain theses in first-order philosophy of science" (p. 7; emphasis added).

The structure of the chapter is as follows. Section 1 introduces a non-exhaustive list of the ways a research project in philosophy of biology can be described and pursued. In section 2, reflecting on these approaches and in line with Kaiser (2015), I outline the main meta-philosophical preliminaries guiding my own project. In section 3, I go deeper and explain its different normative dimensions. In section 4, I try to resolve some issues that these dimensions can give rise to.

1. What is philosophy of biology?

What is philosophy of biology?[35] Since this domain has its own history (albeit short), a rich bibliography, and its own masterpieces the question could seem trivial.[36] The immediate answer is that philosophy of biology reflects on the principles, methods and practices of the biological sciences.[37] While this answer is satisfactory, it is general enough to be of little help in grasping the domain's essence, and the plurality of ways in which philosophy of biology can be pursued does not help with a general definition. In the following sections, I claim there is no unique way to do philosophy of biology, but rather different methodologies for approaching it.[38] Although this plurality is well known, it is sometimes not appropriately explained. I will not be exhaustive of course; my aim is more modest. I simply want to stress that from a methodological point of view, the domain's plurality is evident. This is useful to the extent that it triggers self-reflection regarding which methodology and approach to choose, given an interdisciplinary project.

1.1 Philosophy of biology methodologies

In this section, I describe some ways of pursuing philosophy of biology which are not mutually exclusive and that, as we will see, can also overlap. The first feature that comes to mind when considering philosophy of biology is its role in clarifying key concepts in biology.

[35] Often in the literature the question at stake is, rather, what philosophy of biology "*should* look like" (e.g. Callebaut 2005; Hull 1969; Ruse 1989; cf. introduction of the thesis). This second question is strongly normative. It assumes that there is a *good* way to do philosophy of biology that should be followed. For example, Hull (1969) criticizes certain approaches to philosophy of biology and shows "what can happen when a philosopher does not have sufficient understanding of the view he criticizes" (p. 159). In general, he writes that philosophers could contribute a lot to biology but his "complaint is that by and large they [do] not" (p. 171; see also Ruse 1989). All in all, we will see that I address normativity in a different fashion. I do not ask what philosophy of biology should be in general, but rather what kinds of normativity a project in philosophy of biology could have (see after).

[36] For an overview about these masterpieces refer to Griffiths (2007).

[37] In a similar vein, Pradeu (2018) writes: "[p]hilosophy of biology refers to the critical examination of the conceptual, theoretical and methodological foundations of today's life sciences" (p. 430).

[38] In the following section, I delineate philosophy of biology through its methodological approaches. For a more historically-based reconstruction, see Griffiths (2007).

For example, some scholars try to provide a satisfactory definition of species while others counter that the concept of "gene" is not so unproblematic as it initially appears. Others still focus their efforts on defining drift and natural selection. All of these examples highlight a first role for philosophy of biology: it analyses, reworks and clarifies problematic concepts in biology. In this regard, echoing Richardson and Reichenbach (1938), Laplane *et al* (2019) write that one of the contributions of philosophy to science concerns the "clarification of scientific concepts" (p. 3948). Most of the time, this philosophical clarification aims mainly to improve scientific works. I qualify this way of doing philosophy of biology as the "philosophy of biology of concepts" (hereafter, PBC). Note, however, that this way of seeing philosophy of biology is quite problematic. Indeed, some philosophers would disagree with the definition of philosophy of biology as *just* PBC because they conceive of their work as something more than merely clarification for biological purposes. Often, the disagreement centers on the fact that the value of the philosophical work *is* (or *is not*) measurable only (or also) by scientific utility. In this vein Kaiser (2015) writes: "it is implausible to regard actual utility [for scientists] as a *measure of the quality* of a philosophical account" (p. 9; my emphasis) even if "this does not preclude that a philosophical account should be potentially useful to scientific practice" (p. 9). I am in line with Kaiser in saying that PBC could help to clarify scientific concepts but that the quality of the philosophical analysis is not predicated upon the degree to which it helps in this way.

Just as biologists use concepts to develop the map of the living being, philosophers of biology draw and redraw the conceptual topography of the biological sciences. Livingstone Smith writes that philosophy of biology can be characterized as a "higher-order biological theorizing" (Livingstone Smith 2017, p. 1).[39] Kaiser (2015, p. 12, referring to Carrier 2007, p. 15) asserts that the philosophy of science in general and the philosophy of biology in particular are "*second-order* discipline[s] […] and the phenomena [they try] to account for are also located on a 'second level'" (emphasis in original). Therefore, philosophy of biology is a discipline that studies second order information (e.g. Kaiser 2019, p. 45) coming from science (e.g. biology). Let me call this way of doing philosophy as "philosophy of biology as second order discipline" (here after PBSOD). Note that PBSOD could present a further stratification. Some projects in PBSOD can directly reflect on biological concepts. For example, Keller (2000a), Keller and Harel (2007), Falk (1986, 2000, 2010), Griffiths and Stotz (2007) and Portin (1993, 2009)

[39] Fagot-Largeault (2009) puts it in a similar fashion: "[w]hat do [philosophy and science] share? It may be a common desire for truth. How do they differ? One may hypothesize that science works at conquering new pieces of knowledge, and philosophy retrospectively studies how science did the work" (p. 32).

reflects on the concept of genes in evolution and development.[40] Others can *indirectly* reflect on biology through philosophical epistemic norms (e.g. the problem of laws, the structure of explanations, the role of models, etc.). For example, Kaiser (2015) provides an analysis of reduction and reductionism (i.e. epistemic norms) in biology. My project is similar to Kaiser's to the extent that I work with epistemic norms, namely explanation and abstraction with regard to chance in biology. In this way, my project is both PBSOD and PBC.

A third and a fourth way of doing philosophy of biology refers to two different *approaches* often cited in literature on how to conduct philosophy of biology: "top-down" and "bottom-up". The first, also called the "philosophy-first approach" (e.g. Sterelny and Griffiths, 1999; xi; Turner 2011a, p. 12), asks philosophical questions about science in general. (For example: What is science? How can we separate it from pseudo-science? What is its aim?) Often, philosophers who embrace this approach use "biological science as a *testing ground* for claims in [the] general philosophy of science" (Griffiths, 2007, p. 69).[41] One could call this approach a "top-down philosophy of biology" (hereafter TDPB). The second approach is the "science-first approach" (e.g. Turner 2011a, p. 12) and consists of two stages: the study of biological works, followed by the extrapolation of philosophical questions. Philosophers don't have to be "afraid of getting their hands dirty by grappling with the empirical domain [of] scientific practice" (Livingstone-Smith 2017, p. 7; i.e. first part) and extrapolating relevant, philosophical issues from these practice (i.e. second part). More explicitly: "[t]he challenge is to begin with the science and then gradually work one's way up into philosophical territory, usually by following up on conceptual or normative questions that arise during the course of scientific research" (Turner 2011a, p. 13). Kaiser falls into this methodological camp, writing that: "philosophy of science should be 'bottom-up' […] that is, it should emerge from a detailed investigation of contemporary scientific practice" (Kaiser 2015, p. 12). I want to add that this way of doing philosophy could be considered a gamble. Indeed, it could be the case that from a deep analysis of biological work, no "bottom-up" philosophical questions arise. Turner, proposing his project in the philosophy of paleontology, clearly shows that the adequacy of this approach depends on the results that it can, *maybe*, give. He asserts: "I won't try to offer an

[40] The reader can see here the first overlap between the ways of doing philosophy of biology. For example, Keller provides an example of a PBSOD and PBC project because she analyzes works by biologists to provide conceptual clarifications. Generally speaking, the fact that the categories that I am proposing can overlap each other is not a problem for my analysis but rather a further feature of complexity in the study of philosophy of biology (see after).
[41] Note that this approach is often problematic because it could lead to an arbitrary normative paradigm on science. For example, "a top-down procedure can tempt one to distort the empirical basis for example by selecting not the phenomena that are paradigmatic and important but that support one's philosophical theory" (see Kaiser 2015, p. 13).

elaborate defense of this [bottom-up] approach here, in part because this kind of methodological decision can only be justified by its fruits" (Turner 2011a, p. 13). Let's call this methodology the "bottom-up philosophy of biology" (hereafter BUPB).

BUPB permits reflection upon an important aspect in developing philosophical projects, namely their justification. Normally, the justification of why, *in the first instance*, a project is interesting and relevant for the philosophical community is provided in its first pages. But with BUPB, since justification resides in the very results (if any) to which the approach gives rise, a BUPB project does not have to provide justification at the beginning of the dissertation. Justification can come at the end of the work. But this could lead to a view that is too radical with respect to the justification of this type of project. A vision that reflects a compromise might be desired. Indeed, if on the one hand I can agree that part of the value of a BUPB project (of which this project is one) can be related to its results, on the other hand, I nonetheless do recognize it is still necessary, at the beginning of any philosophical project, to justify the *approach* taken to arrive at hypothetical results. This chapter aims to provide exactly this kind of justification.

Griffiths (2007) sums up the general idea behind the TDBP and BUPB approaches thus:

"[p]hilosophers have engaged with biological science in [two] quite distinct ways. Some have looked to biology to test general theses in philosophy of science [i.e. TDPB]. Others have engaged with conceptual puzzles that arise within biology itself [i.e. BUPB]" (pp. 67-68).[42]

If Griffiths' quotation simply and effectively explains the difference between TDPB and BUPB, Kaiser describes historical and contemporary projects by adopting these two approaches. She proposes seeing TDPB as the "first" philosophy of biology, dating back to the 60s and 70s. An emblematic debate of this "first" philosophy is the effort to apply Nagel's model of reductionism (1961) to biology (e.g. Schaffner 1967; Hull 1974), or to deny this possibility. Famously, Kitcher (1984) developed one of the first anti-reductionist arguments in favor of the irreducibility of Mendelian genetics to molecular biology. By contrast, bottom-up philosophy (i.e. BUPB) could instead be traced to the latest trends in philosophy of biology. For example, Kaiser (2015) provides an understanding of the reductive explanation in biology though deep

[42] To be fair, Griffiths gives a third way in which philosophers are engaged with science: "[f]inally, philosophers have looked to biological science for answers to distinctively philosophical questions in such fields as ethics, the philosophy of mind, and epistemology" (pp. 67-68). I have the impression that here Griffiths is referring to the discipline called bio-philosophy (Livingstone Smith 2017). Bio-philosophy is a separate field which uses biology as a source for elaborating new reflections on philosophy. Nevertheless, in the present context I cannot enter into details.

analyses of actual (non)reductive biological explanations (e.g. see Kaiser 2015, chapter 6); work by Baedke (2018) explores explanations in epigenetics – among other things – starting from the work by stem cells biologists (e.g. Baedke 2018, chapter 2); analyzing data from paleontology and paleobiology, Turner (2011a, 2011b) explores the meaning and philosophical implications of concepts such as species selection, trends, and progress.

Philosophy of biology methodologies	**Definitions**
Philosophy of biology of concepts (**PBC**)	PBC is a methodology that aims to analyse, rework and clarify problematic concepts in biology (e.g. the concept of gene, natural selection, fitness, etc.).
Philosophy of biology as second order discipline (**PBSOD**).	PBSOD is a methodology that refers to the study of second order information. Philosophers of biology draw and redraw the conceptual topography of the biological sciences.
Top-down philosophy of biology (**TDPB**).	TDPB is a methodology that uses biological sciences as a testing ground for claims in philosophy of biology
Bottom-up philosophy of biology (**BUPB**)	BUPB is a two-step methodology. Firstly, comes the immersion of biological practice in in-depth studies of scientific work and practice. Secondly, relevant, philosophical questions are extrapolated. From out of these questions the philosophical work is developed.

Table 1 List of the different methodologies of philosophy of biology described in the present section

PBC, PBSOD, TDPB, and BUPB are all artificial categories, comprising a non-exhaustive list of the different ways in which philosophy of biology can be pursued (see Table 1). For example, Kaiser's project on reductive explanation in biology (2015) could be seen variously as BUPB, because it is based on the understating of biology in practice (cf. see section 2.3); as PBSOD, because it takes on biological works from a philosophical perspective; and as PBC, because part of her work is to understand what "reduction" means for biological practice

(Kaiser 2015, p. 44). My own project can be seen as PBC, because it will provide an analysis of the concept of chance in biology; as PBSOD because it aims to investigate works by biologists and the ways in which they think about chance; and also as BUPB because my philosophical reflections stem from the analysis of biological works. However, my work is not TDPB, because one of the key assumptions is it has to be as close as possible to practical science.

2 - Kaiser's meta-philosophical preliminaries as a guideline for our project

In the previous section, I provided a general outline of a few ways of undertaking philosophy of biology. In this section, I enter more deeply into the criteria of adequacy that make my project a good example of three ways of doing philosophy of biology: philosophy of biology of concepts (PBC), philosophy of biology as a second order discipline (PBSOD), and bottom-up philosophy of biology (BUPB). These criteria constitute methodological naturalism (section 2.1), and the three steps involved in properly describing biological phenomena in a philosophical project, are: focusing on relevant data, explicating underlying assumption and establishing coherence (section 2.2). In order to explain these points in detail, let us first of all briefly describe the original framework by Kaiser from which I have borrowed and developed these points.

In the second chapter of her book, Kaiser (2015) makes explicit a series of meta-philosophical preliminaries that are pivotal to understanding the project she develops throughout the rest of the book. She writes that these meta-preliminaries could be useful not only for her specific project but more generally, when developing any project in philosophy of science. She writes:

"Even though this chapter primarily serves to illuminate the goals and methodology of my analysis of reduction, the views I develop and, in particular, the four *criteria of adequacy* that I propose point beyond my specific project of analyzing reductive explanation in biology. These criteria can easily be adapted such that *they apply more broadly* to philosophy of science in practice" (Kaiser 2015, p. 41; my emphasis).

These "criteria of adequacy" could thus be a reference to different projects in philosophy of science in general and philosophy of biology in particular. I take this advice seriously, embracing some (but not all) of them as a good foundation for my project. In the following two sections (2.1 and 2.2), I rebuild some of Kaiser's meta-philosophical preliminaries that are

relevant for my work. In section 3, I specify the normative dimensions of the project, using the conceptual tools she provided in her book.

2.1. First general criterion: methodological naturalism

Often, philosophers specify that the boundaries between science and philosophy are not discrete but a *continuum* (e.g., Laplane *et al* 2019, p. 3950).[43] Kaiser's first criterion of her philosophical analysis, which she calls methodological naturalism,[44] corresponds to this idea of continuity between philosophy and science (2015, p. 10). But what sort of continuity it consists of is still hotly debated; methodological naturalism has been defined in different ways (e.g. Papineau 2009; Plantinga 1996; Kornblith 2007).[45] Kaiser (2015) writes:

"Although the details of how methodological naturalism is characterized vary […], the general statement about philosophy remains largely the same: a methodological naturalist claims *that with respect to their methods (and aims) there exists no principled difference between philosophy and the natural sciences*. In other words, philosophy can be pursued by applying methods that are similar to those successfully employed in the natural sciences (and vice versa)" (p. 11; emphasis added).

For Kaiser, the continuity between science and philosophy concerns *aims* and *methodology*. Passing from the first- to the second-order discipline (i.e. from science to philosophy; cf. Table 1, PBSOD) we see no difference with regard to their methods and aims. This definition of methodological naturalism can be called the *broad* notion of methodological naturalism.[46]

[43] As to other authors, Love writes that: "[this] continuum can be conceptualized in terms of the degree of abstraction with philosophy occupying the more abstract end of the continuum and science spread out toward the less abstract. The relationship between philosophy and science is then understood *as the movement back and forth along this abstraction continuum*" (Love 2008, p. 68; my emphasis; on the metaphore of "going back and forth" cf. section 3.2). In a similar fashion, Callebaut (2005) writes that "[i]n its self-understanding […] philosophy of biology is on a par and *continuous* with biology" (p. 96; emphasis in original). In addition, Rosenberg and McShea (2008) write: "the difference between philosophy and theoretical science is not a matter of kind but of *degree*" (p. 5; my emphasis).

[44] Naturalism is a very complex topic that has a long history (for an overview see Rosenberg 1996). I do not enter into this complexity; I only develop a specific concept of "methodological naturalism" (see the main text).

[45] The origin of this debate can be fixed at the publication of Quine's article entitled "Epistemology Naturalized" in 1969 in which for the first time the method of translational epistemology of linking observation and theories through deductive logic is seriously questioned. Quine states that it is more interesting to understand how the process of knowledge takes shape, that is to say "an empirical psychology study of our cognitive processes" (Kim 1988, p. 387), rather than continuing to create unjustified, since *a priori,* speculations. He famously states: "Why not settle for psychology?" (Quine, 1969, p. 92) instead of remaining unsatisfactorily "to *translate* science into logic" (ibid., emphasis in original)?

[46] Chakravartty (2017) adds, to "aims" and "methods", the "criteria of evaluation for theories and hypothesis [...] such as simplicity, internal consistency, coherence with other knowledge, and the capacity to unify otherwise disparate phenomena" (p. 79).

An objection to this broad definition is that between philosophy and science there are *nonetheless* important differences. *A priori* intuitions are central in developing philosophical theses. See, for example, the metaphysics of science that works with *a priori* reflections (e.g. Lewis 1986). By contrast, in science, empirical investigations, such as experiments, characterize the research (Kaiser 2015, p. 11). Kaiser responds to this objection by flipping the argument, noting that scientists also deal with *a priori* intuitions[47] and that philosophers also make use of empirical data. But in order to firmly resolve the issue that this objection raises, Kaiser provides a more sophisticated definition of methodological naturalism, which I will briefly rebuild now.

Kaiser develops a narrow notion of methodological naturalism, starting from the differentiation between descriptive and normative philosophical projects. A descriptive project in philosophy of science is "an account […] that aims at capturing and understanding actual biological practice and that proceeds by 'describing' this practice" (Kaiser 2015, p. 15). On the other hand, a normative project specifies the way in which, for example, a concept should be understood, or the way in which research *should* be pursued. She takes reduction as an example, noting that in normative projects philosophers establish how reduction should be done "without caring much about what cases of reduction[…] actually performed in biological practice look like" (p. 15). Furthermore, she stresses, there are no purely descriptive or purely normative projects in philosophy of science: the majority are descriptive-normative (for further clarification on normativity c.f. section 2.3 of this chapter). In the present context, Kaiser is interested in the relation between the descriptive part of a philosophical project and the notion of methodological naturalism:

"What I want to call attention to […] is that there exists in fact a similarity between conducting philosophy of science *in a descriptive fashion* and performing research in the natural sciences. Exactly *this* is the reason why many philosophers of science characterize their descriptive account as naturalistic. The similarity between descriptive philosophy of science and the natural sciences is that in both cases we have 'empirical data' or 'phenomena' that need to be captured (or saved) by the account or theory that is developed" (Kaiser, 2015, p. 11, emphasis in original; see also Kaiser 2019, p. 11).

[47] The fact that science deals with *a priori* reflections is also one of the main problems for certain authors interested in developing a project of naturalized metaphysics (see e.g. Chakravartty 2017). What could be a good methodological distinction between an *a priori* metaphysics and naturalized metaphysics is that the latter takes seriously the empirical evidence of science. But the fact that science consists not only of empirical evidence but also of a *priori* reflections could be used as an objection that could jeopardize the effectiveness of the *a priori/a posteriori* discriminant underlying the distinction between *a priori* vs naturalized metaphysics (.Chakravartty 2013).

For Kaiser, methodological naturalism refers to the descriptive part of a descriptive-normative project in philosophy of science (see Figure 1, the blue set on the left). More specifically, the descriptive part of a philosophical project follows the same *methodology* that scientists embrace when performing research in the natural sciences (see Figure 1, the green set on the right): both work with data and phenomena which have to be captured and elaborated upon to yield a theory/explanation/prediction (see Figure 1, the intersection of the two sets). This is what I call the *narrow* definition of methodological naturalism provided by Kaiser. The main difference between Kaiser's broad and narrow definitions of methodological naturalism is that by pointing at the descriptive part of a philosophical project that shares methodology and aims with science, the narrow definition is clearer.

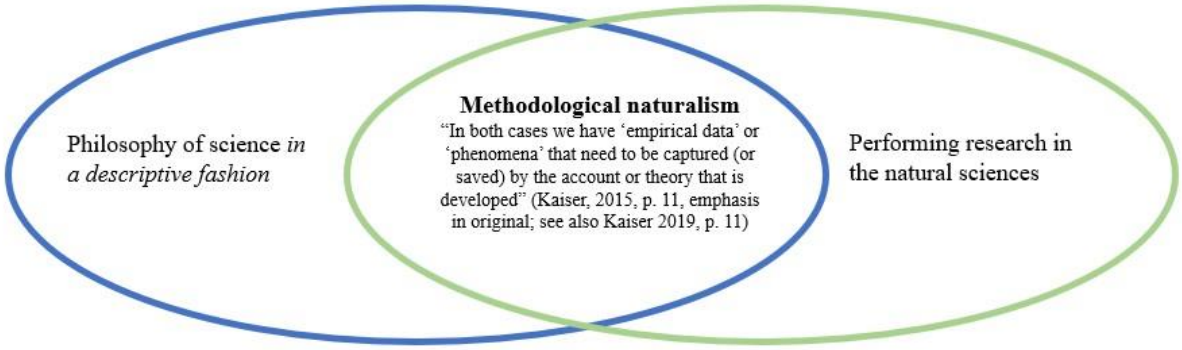


Figure 1 A representation of the narrow definition of methodological naturalism provided by Kaiser (2015) depicting the intersection between the descriptive part of a descriptive-normative philosophical project (the blue set on the left) and research performed in the natural sciences (the green set on the right). For further detail, see main text.

I am sympathetic to this narrow definition and my project shares this criterium of adequacy since it presents very clearly one of the ways in which philosophy and biology can interact and overlap.

Since philosophical projects are mainly descriptive-normative, it is essential to satisfactorily clarify what it actually means (and implies) for a project to be jointly descriptive and normative. In the following section (2.2) the focus will be on the other three criteria of adequacy developed by Kaiser. In section 2.3, I will build briefly on her reflections on normativity. Both of these discussions will be useful for me to clarify how to develop a good work in philosophy of biology, and in delineating the descriptive and normative parts of my philosophical project on chance in molecular and cellular biology (see section 3).

2.2 Kaiser's other three criteria of adequacy

We have just seen that, according to Kaiser, the aim of the descriptive part of a philosophical project is "capturing and understanding biological practice" (2015, p. 15). But describing biology is not straightforward. Which areas of biological literature should I consider in my project? What stance should I take when even biologists cannot agree on certain things? What should be done when the things of interest in a philosophical project are only briefly hinted at in biological discussions? Having asked similar questions, Kaiser states that "the process of developing a philosophical account of science is to be characterized as an *active, critical reconstruction*" (Kaiser 2015, p. 19; emphasis in original). In saying this, Kaiser stresses the idea that a *pure* and *passive* description of biological practice is, in the end, not possible.[48] She characterizes this active reconstruction in three stages that correspond, with methodological naturalism, to the criteria of adequacy of her own philosophical analysis:

- *Focusing on relevant data.* "[T]he amount of available information about scientific practice is [so] immense and overwhelming […] that philosophers must find a means to select […] information that [is] relevant, for instance, because they represent paradigmatic, typical or important cases" (Kaiser 2015, p. 19). She also underlines limitations faced by all philosophers, in that "it is simply impossible to analyze all available cases from each scientific field" (p. 20). Philosophers can focus on cases particularly debated in the scientific community, or cases that provide successful explanations, but the point is that they have to choose – namely *critically* select – the work of their interest. For example, in Chapters 5 and 6 I focus on the case study of noise in gene expression because the topic has been central to molecular biological research on stochasticity in the last 20 years.
- *Explicating underlying assumptions*. This deals with the idea that it is often necessary "to *make explicit* the assumptions that are only implicitly present in scientific practice" (Kaiser 2015, p. 20; emphasis in original). Kaiser takes the example of reductive explanation. Very rarely do biologists specify whether their explanation is reductive or not, and when they do, it is usually because the reductive character of an explanation is

[48] More explicitly, Kaiser (2015) asserts that it is not possible to develop a philosophical account that is completely descriptive, namely one that has no normative commitments. The simple act of reconstruction is, as she writes, an active act, and as such, it implies some sort of (even minimal) normativity. On the other hand, a completely normative account would not be of interest to philosophy of science in practice because it would be too distant from biology in practice (p. 18).

inadequate. This aspect of reconstruction is concerned with making explicit the philosophical commitments of biological studies.

- *Establishing (or just discussing) coherence*. Within the vast fields of biological work and data could exist conflicting concepts, stances and principles or, at the very least, heterogeneous situations. Kaiser quotes a series of philosophers who have already underlined this issue (e.g. Waters 2004, p. 38; Woodward 2003, p. 7). The work of philosophers is, then, to establish coherence between heterogeneous biological frameworks. For example, if biologists use the word "reduction" in different fashions, it becomes a philosophical task to unravel and order that conceptual heterogeneity. Note that making "order" does not necessarily mean establishing homogeneity from heterogeneity. For example, Keller (2000a) provides an analysis of the various concepts of gene. She does not question the heterogeneity of its conceptualization in scientific literature, rather, she provides a pluralistic framework whereby these positions might co-exist. In this case, it is about discussing (and not establishing) coherence.

It should be noted that these three criteria of adequacy could shed some light on what it means to do bottom-up philosophy of biology (BUPB), with philosophers "getting their hands dirty by grappling with the empirical domain" (Livingstone Smith 2017, p. 7-8). What bottom-up philosophers do is to focus on relevant data, explicate underlying assumptions, and establish (or just discuss) coherence.

2.3 Multifaced normativity in philosophy of science by Kaiser

Let's now look at the normative dimension provided by Kaiser (2105, 2019). The aim of this section and of those that follow (sections 2.3 and 2.4) is to briefly build (in a non-exhaustive way) upon her analysis in order to make explicit the normativity of my own research (that will be explicitly elucidated in section 3).

Recent years have seen a general tendency in philosophy of science to shift "toward scientific practice", rather than focusing on abstract theories detached from science (e.g. TDPB). This shift is called "practice turn" (e.g. Ankeny *et al* 2011; Boumans and Leonelli 2013; Soler *et al* 2014) and consists of the idea that it is time to move away from normative and scientifically inaccurate projects in favor of empirically informed, descriptive projects in philosophy of science. Nowadays, this new attitude is often termed philosophy of science in

practice (PSP). One of the main aims of PSP philosophers is to reflect on how scientists do what they do to develop scientifically informed projects.[49]

Advocates of PSP often stress that the normative aspects of philosophical projects must be rejected. Nonetheless, Kaiser (2019) points out that in practice, not all kinds of normativity are actually rejected. Rather, the normative parts rejected are the non-empirically informed ones. For example, an advocate of PSP could reject Schaffner's model of general reduction-replacement (GRR, 1974) because it is a non-scientifically informed normative project that would tell biologists how they should pursue research (see also Kaiser 2015, p. 15). By contrast, Kaiser argues that normativity in PSP is not totally eradicated, rather it is reframed in different and more formal ways.

Generally speaking, Kaiser writes that normativity is not "a general feature that a philosophical theory either has or lacks" (Kaiser 2019, p. 36). Rather, it is a complex "multifaceted phenomenon" (p. 36). The first kind of normativity that Kaiser proposes in her 2015 and 2019 works is what she calls "metanormativity", whereby a "philosophical theory (T) about a feature or element of science (E) is metanormative if T contains normative claims about E" (Kaiser 2019, p. 41). In other words, metanormativity refers to any philosophical project or theory that, to some extent, dictates how science should proceed and what it should look like. An example that Kaiser provides for metanormativity is the aforementioned GRR that Schaffner (1974) proposed in response to evidence that Nagel's formal model of reduction (1961) did not work for biology (Kaiser 2015, p.17-23; 2019, p. 40). Indeed, with this model, Schaffner "posits how reductions ideally *should* look" (Kaiser 2019, p. 17; my emphasis). The GRR could be seen as T in Kaiser's definition and suggests a specific way in which reduction (i.e. "a feature or element of science E") should be pursued.[50]

Kaiser adds a further level of complexity to the category of "metanormativity" by developing two subcategories: 1) "ex-cathedra metanormativity"; and 2) a kind of metanormativity that could be present (in an explicit or implicit way) in PSP projects. For the sake of argument, let us call this second subtype of metanormativity "in-practice

[49] A PSP project can be transversal concerning the philosophy of biology methodologies sketched in section 1.1. It can be a PBC project since it can be aimed to reflect on concepts actually used by scientists (e.g. the concept of noise, see Chapters 5 and 6). It is of course a PBSOD project to the extent that it provides philosophical reflections on scientific topics. And it is necessarily a BUPB project since starting the analysis from scientific practice, as PSP suggests doing, coincides with the principle of BUPB to bottom-up philosophical questions from the study of scientific materials (e.g. explanations, models, methods, etc.).

[50] Kaiser also provides other examples of this kind of normativity, e.g., Popper's (1959; 2002 [1932]) idea of the falsifiability of science, Hempel's (1965) theory of the structure of explanation. Kaiser specifies that these are examples of philosophical works that tell how science should look like, both from a prescriptive and evaluative point of view (see Kaiser 2019, p. 42).

metanormativity". Kaiser proposes that ex-cathedra metanormativity can be found in the metaphysics of science literature (she cites the example of Lewis 1999; see Kaiser 2019, p. 43).[51] In general, "ex cathedra metanormative claims about a certain feature or element of science are *not informed by* and *cannot fail* in light of the empirical reality of scientific practice" (Kaiser 2019, p. 43; my emphasis). This first kind of metanormativity is incompatible with PSP because it clearly excludes the development of philosophical projects in an empirically informed way. On the other hand, in-practice metanormativity is compatible with PSP. She writes: "a philosophical theory about a certain feature or element of science E contains only such normative claims about E that take into account, are drawn from, or *are informed* by factual claims about E" (Kaiser 2019, p. 44; emphasis added).

In short, normative claims of ex-cathedra metanormativity are not empirically informed, while normative claims of in-practice metanormativity are. To substantiate this point, Kaiser gives examples of in-practice metanormativity: Craver's account (2007) of mechanistic explanations in neuroscience, which distinguish between good and bad explanations through scientifically-informed and up-to-date arguments[52] and Woodward's theory (2005) of causal explanation that, starting from accurate descriptions of scientific scenarios, elaborates on philosophical reflections on explanation and causal relation. To this list might also be added Baedke's work (2018) which develops a philosophical account of epigenetic explanations in light of the numerous available biological explanations (e.g. Baedke 2018, Chapter 5).

Let us now look at the second kind of normativity discussed by Kaiser, which she calls "methodological normativity". In section 2.2, I wrote about how the descriptive part of a project in philosophy is not passive, but is rather an active reconstruction of biological theories, data, explanations and models. I sketched the three different ways in which this active reconstruction could takes place: by focusing on relevant data, explicating underlying assumptions, and establishing coherence. These three criteria of adequacy also have a normative dimension to the extent that the philosopher chooses the relevant data, the assumptions to be made explicit, and the subjects by means of which they establish coherence (see Figure 2). In line with Kaiser, I call this dimension "methodological" normativity because it refers to the way in which philosophers pursue a PSP project.

[51] The discourse is however open with respect to whether naturalized metaphysics uses this type of metanormativity and if not, which normativity can be attributed to it.

[52] But one could argue that the very fact of assuming a certain kind of explanation could be understood as a return to an ex-cathedra metanormativity. I leave this question open because it is not the focus of this section. Undoubtedly it is difficult to find completely uncontroversial examples of in-practice metanormativity (see after).

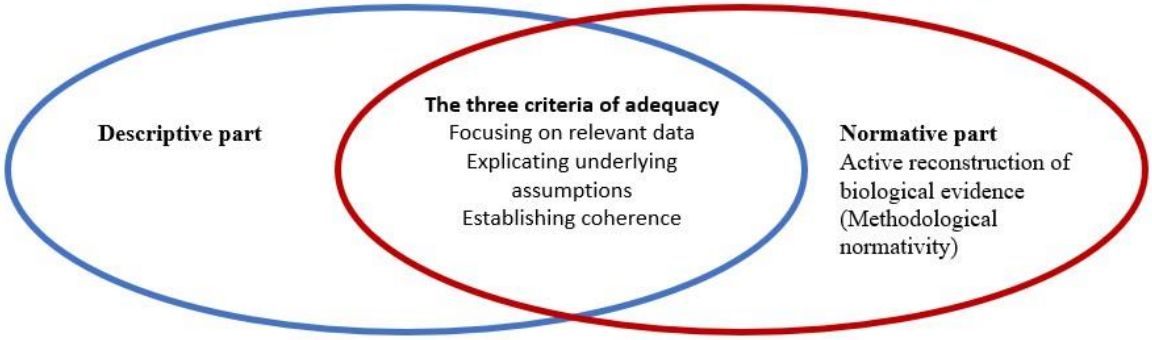


Figure 2 A representation of the descriptive and normative dimensions of the three criteria of adequacy proposed in section 2.2. For further details, see main text.

At this point in the argument, a reader might complain of a certain conceptual confusion between methodological normativity and methodological naturalism. Methodological naturalism sees philosophy of science in a kind of continuum with science. In this context, it is enough to say that a philosophical project that deals with general methodological naturalism "must be empirically adequate, that is, it must capture and find evidential support in the available empirical data" (Kaiser 2019, p. 45). Its difference to a scientific project resides only in the fact that it deals with empirical data or information from science rather than the natural world itself. Therefore, what is the relation between methodological naturalism and methodological normativity?

Methodological naturalism by Kaiser (2015) Broad definition: continuum between science and philosophy: they share the same methods and aims
Methodological normativity by Kaiser (2019)
Active reconstruction of biological theories, data, theories, models etc. (i.e. focusing on relevant data, explicating underlying assumptions and establishing coherence).

Table 2 The relation between methodological naturalism and methodological normativity. For further details, see main text.

To answer the question, I propose focusing on Table 2 above. Methodological normativity and naturalism are two linked but different concepts. Methodological naturalism (green) is strictly tied to methodological normativity (red) because the former is the theoretical framework from which the second can be established. Only from a scenario that allows the idea of continuity

(cf. broad and narrow methodological naturalism, section 2.1) between science and philosophy is it possible to elaborate a PSP project on biology which focuses on relevant data, explicates underlying assumptions, and establishes coherence.

Let's now turn to the third kind of normativity that Kaiser (2019) calls "object normativity" in which the object of the analysis is itself normative. In this case, the focus is not on species, genes or fitness but on epistemic and/or social norms. As an example, consider a PSP project focused on explanatory power (i.e. the epistemic norm) that aims to investigate the extent to which one can say that a scientific explanation is a good explanation (cf. "theorizing about norms", see after). This is a perfect case of object normativity. Kaiser proposes dividing this kind of normativity in subcategories:

- She distinguishes between philosophical accounts that describe norms that *are* accepted in science, and norms that *should* be accepted in science.
    - She calls norms which are accepted "describing norms" and gives two examples. The first is Lloyd's (2006) work in which she shows "how gender biases influence the development of adaptive explanations of female orgasm" (Kaiser 2019, p. 53). Lloyd's aim is to describe a norm (i.e. gender biases) and how it influences a science practice (i.e. the development of adaptative explanation of female orgasm). As a second example, Kaiser takes her own project on reductive explanation (Kaiser 2015) in which she analyses "how biologists evaluate the strengths and limits of reductive strategies" (p. 14). Even in this case, what is under investigation are the norms that are accepted by biologists concerning reductive explanations. The aim of the philosopher is to describe these norms.
    - She calls norms which should be accepted "positing norms". These she defines as "philosophical theories that posit which epistemic or social norms should apply to science and justify why these norms should be accepted" (Kaiser 2019, p. 54). She gives the example of Craver's (2007) work on how mechanistic explanations in neuroscience *should* be evaluated[53].

[53] Concerning explanation and neuroscience see also Bickle (2003).

- Kaiser also distinguishes philosophical theories that theorize about norms, and theories that are only *related* to norms.

    - She calls the former "theorizing about norms". This category refers to philosophical projects that are *about* norms, i.e. they investigate the norms themselves. She gives the example of Thagard (1988) who proposes a theory concerning the value of the epistemic norm of simplicity "considering how we *should* understand it and why it is justified" (p. 54; my emphasis);
    - She calls the latter "theorizing relating to norms" in which the object *E* of the philosophical theory "is not epistemic norms but another nonnormative element of science (for example, causal inference, reduction, or the concept of gene), and the philosophical theory includes claims about how *E* is related to certain epistemic norms (that either are in fact or should be accepted in science)" (Kaiser 2019, p. 54). This second subtype of object normativity plays a big part in my philosophical normative project, in which I aim to investigate chance (i.e. which is not an epistemic norm) in relation to the epistemic norm of explanatory power (see section 3.1.3).

Then, if describing norms consists of a philosophical analysis of what is already accepted in science, positioning norms is proposing what should be accepted in science. In other words, the former is an evaluative project and the latter a prescriptive one. Moreover, if theorizing about norms consists of elaborating a theory *about* epistemic norms (e.g. cognitive salience, simplicity, precision, explanatory power, likeness, etc.), theorizing related to norms consists of elaborating a theory about nonnormative elements of science that could be related to epistemic norms. In a word, in the former, the object E of the philosophical theory *is itself* an epistemic norm (e.g. a theory about simplicity). By contrast, in the latter, the nonnormative object E *is related* to an epistemic norm (e.g. the relation between the concepts of gene with explanatory power).

We could say that these paired subtypes of object normativity (i.e. describing and positing norms, and theorizing and relating to norms) are not mutually exclusive but could overlap. For example, Kaiser recognizes that her project (Kaiser 2015) both describes norms (i.e. "how biologists evaluate reductive explanations as adequate" p. 53), and theorizes in relation to norms because the object of her analysis is (the adequacy of) reductive explanations that are, thus, related to the epistemic norm of explanatory power.

I have briefly presented the four kinds of normativity elaborated by Kaiser (2019). Table 3 below gives a clear idea of all the categories and a brief summary of their descriptions:

Kinds of normativity by Kaiser (2019)	**Subcategories**	**Brief definitions**
Metanormativity (Refers to any philosophical projects or theories that to some extent dictate how science should proceed and what it should look like).	Ex-cathedra metanormativity	"[E]x cathedra metanormative claims about a certain feature or element of science *are not informed* by and cannot fail in light of the empirical reality of scientific practice" (Kaiser 2019, p. 43; emphasis added)
	In practice metanormativity	"A philosophical theory about a certain feature or element of science E contains only such normative claims about E that take into account, are drawn from, or *are informed* by factual claims about E" (Kaiser 2019, p. 44; emphasis added)
Methodological normativity	Three criteria of adequacy: 1. Focusing on relevant data; 2. Explicating underlying assumptions; 3. Establishing coherence	The part of a philosophical project that consists in an *active reconstruction* of biological theories, data, theories, models etc.
Object normativity ("[t]he object of philosophical theorizing itself can be normative", Kaiser 2019, p. 51).	Describing norms	Philosophical account that describes norms that are accepted in science
	Positioning norms	Philosophical account that describe norms that *should* be accepted in science
	Theorizing about norms	Philosophical account that are *about* norms, i.e. they investigate the norms themselves
	Relating to norms	The object of a philosophical account is not an epistemic norm but the project in its argument aims to *relate* this non-epistemic norm with an epistemic one

Table 3 Summary of the kinds of normativity proposed by Kaiser (2019).

Kaiser's analysis of normativity is more sophisticated than this roughly-sketched version which, nevertheless, will suffice for our needs. This section lays the foundations of the following sections which develop and make explicit the normative parts of our work on chance. More specifically, by using the conceptual tools provided by Kaiser, I will point out the multidimensional normativity of my work.

3 Normativity and our project about chance

The conceptual tools provided by Kaiser (section 2.3) show us that normativity can be seen as "a multifaceted phenomenon rather than a general feature that a philosophical theory either has or lacks" (Kaiser 2019, p. 36). The challenge in this section is arguing that my project contains (at least)[54] certain aspects of the multifaced normativity developed by Kaiser. Recognizing this multifaced normativity enables me to better calibrate my analysis (e.g. to have clearly in mind its descriptive and normative dimensions), to slow down when I get too far out of line (e.g. avoiding certain kinds of normativity, see section 2.3 and hereafter), and to recognize the boundaries of academic domains (e.g. not transgressing the fine line between philosophy of biology and biology).

3.1. In practice metanormativity, methodological normativity, and object normativity

In my research project on chance in molecular and cellular biology, I recognize a kind of *in practice metanormativity* (see section 2.3). More specifically, in the present dissertation I make at least two kinds of in practice metanormative claim:

- I make evaluative claims, namely I affirm that certain biological explanations are *good* explanations because they use chance as a central explanatory element. For example, in Chapter 5 I describe the explanation of antibody synthesis provided by immunologists. In this context, I argue that this explanation is a *good* explanation and its "goodness" stems from the fact that chance here has a key explanatory power[55] (see Chapter 5, section 3.1);

[54] I write "at least" because I do not exclude the possibility that further philosophical analysis could shed light on new facets of normativity in PSP projects.
[55] The philosophical reason for *why* chance has such a key explanatory element is elaborated in Chapter 3.

- I make prescriptive claims, namely I make claims about the approach biologists should adopt in specific situations. More specifically, this prescriptive part of in practice metanormativity characterizing my own research consists of two subtypes of claims:

    1. Biologists *should* recognize that chance can have a role in (certain) biological explanations. In Chapter 5, I argue that chance in gene expression is not only "noise", but (in certain specific situations) should be conceived as a central stochastic element for understanding some features of cell behavior;
    2. Biologists *should* sometimes favor stochastic explanations (i.e. a specific kind of explanation that I develop in Chapter 2) over other explanations (i.e. non-stochastic, mechanistic, and more fine-grained explanations). For example, what I propose in Chapter 5 (section 3.3, Figure 4) is that, if we have to explain the variability in cell fate and cellular differentiation, then we should prefer using an *explanans* that contains an explanatory chancy element (i.e. "stochasticity in gene expression") rather than a non-stochastic and finer grained *explanans.*[56]

My work deals also with the dimension of *methodological normativity*. I select, evaluate, and interpret biological claims (e.g. data/evidence/theories/explanations, etc.) for philosophical purposes. As already mentioned, this is a point where descriptive and normative dimensions overlap (see section 2.3). The descriptive dimension resides in the fact that I actually describe biological scenarios. The normative part resides in the fact that, in rebuilding biological scenarios I in fact make *active choices* about which data to focus on, which assumptions need to be explicated, and which parts of biological works need coherence (cf. section 2.2). One example is my analysis of noise in Chapters 5 and 6. Noise is a term often associated with stochasticity in gene expression. There, I select important biological works about noise, make explicit some underlying assumptions (i.e. the explanatory role of chance), and I discuss about coherence between often heterogeneous definitions.

[56] Furthermore, these prescriptive claims, (1) and (2), are linked. I will show that in order to figure out how chance can have a role in biological explanations (1), we must focus on why a biological explanation with stochastic parameters (i.e. stochastic explanation; see next two chapters and Chapter 5) might be preferred over a fine-grained non-stochastic one (2).

Finally, the third dimension of normativity that I recognize in my work is *object normativity* in terms of theorizing-relating-to-norms. The example that Kaiser provides for this kind of object normativity is a philosophical theory focused on the concept of gene (i.e. "the nonnormative element of science") that also aims to "clarif[y] how the gene concept is related […] to the epistemic norm of explanatory power (for example, by discussing the limitations of gene-based explanations)" (p. 56). In the same vein, I am not dealing with explanation in itself but I deal with chance and its explanatory power. My work is about exploring the role of chance in biological explanations. It therefore links a nonnormative element of science E (i.e. chance) with a normative one (i.e. explanatory power). In what follows, inspired by the works and tables one can find in Kaiser (2019), I propose a figure that visually sums up the multifaced normativity that my project deals with, and their relations.

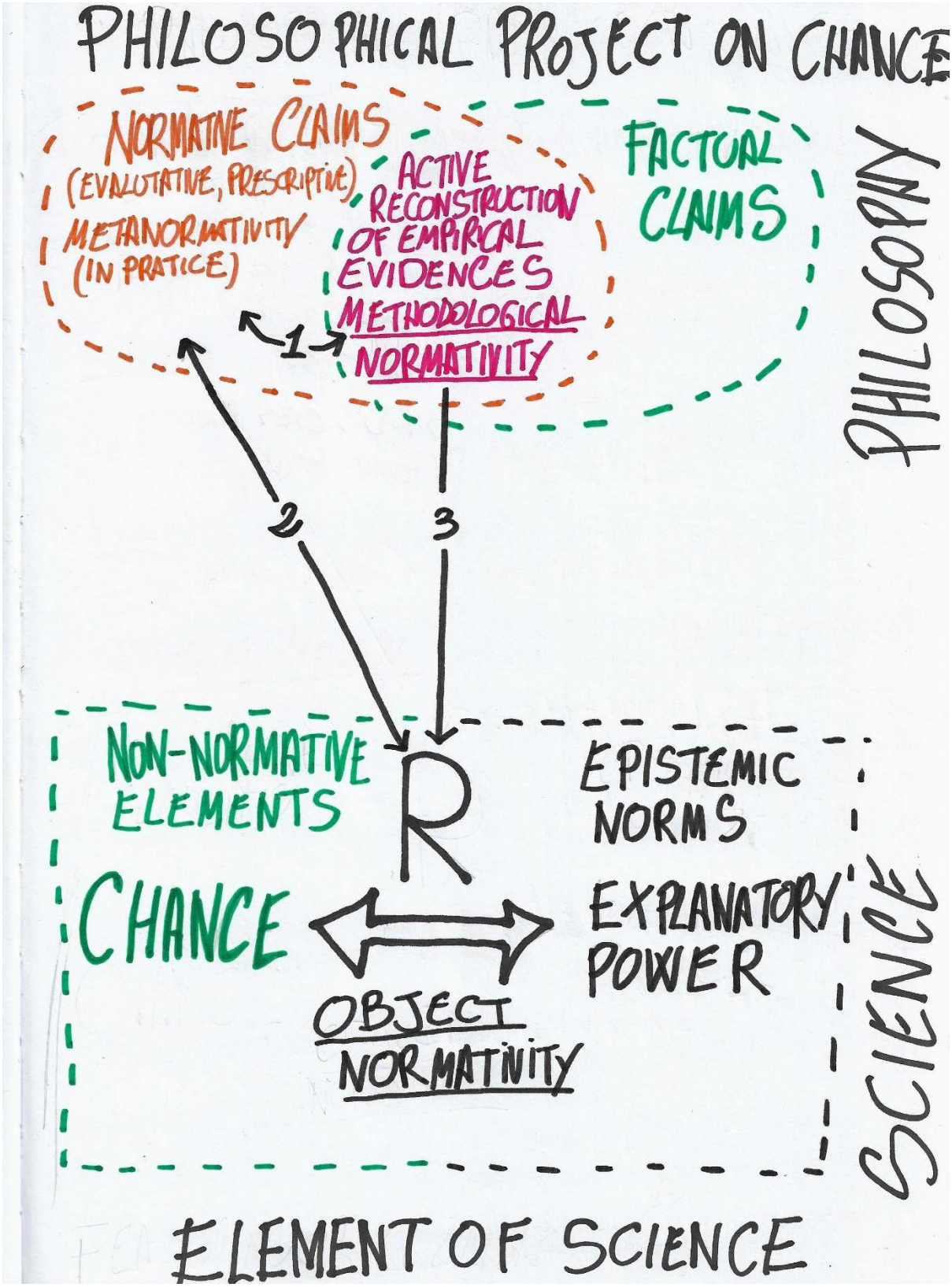


Figure 3 A representation of my project on chance and its kinds of normativity. For further details, see main text.

In the upper part of Figure 3, we can see the philosophical dimension of my project. My project is represented by two parts. The first consists of normative claims (on the left, in orange), and the second in factual claims (on the right, in green). In-practice metanormativity – that concerns claims about how biologists should work (prescriptive claims) and claims about the evaluation of the "goodness" or "badness" of explanations (evaluative claims) – figures in the "normative claims" dimension. I left empty the case of "factual claims" because I agree with Kaiser that a purely descriptive philosophical project is "not possible or not philosophical" (Kaiser 2015, p. 18; see section 2.2). In the original work by Kaiser (2019), normative and factual claims are separated by a dotted line with no overlapping possibilities (see Kaiser 2019, p. 41, Figure 2, p. 49; Figure 3, p. 53; Figure 4; p. 55, Figure 5). By contrast, my representation of "normative claims" and "factual claims" as Venn diagrams[57] enables the creation of an intersection. In this intersection is found the active reconstruction of biological works. In this way, I show where the descriptive and normative dimensions of this project overlap. Indeed, the active reconstruction is both descriptive because it rebuilds biological scenarios, and normative to the extent that the active act of reconstruction implies in itself a degree of normativity that Kaiser called methodological normativity.

In the lower left-hand part of Figure 3 is the nonnormative element of science E of our interest (i.e. chance) and on the right is the normative one (i.e. explanatory power). The double arrow between "chance" and "explanatory power" labelled "R" represents my objective-normative-effort to link chance in biology with some sort of explanatory power. This part of the project is challenging because it is about giving the conception of chance – often associated with disorder, chaos, unpredictability, ignorance – an explanatory dimension. The next two chapters aims to develop and argue in favor of R (cf. conclusion).

Let's now focus on the arrows (i.e. 1, 2, and 3) presented in Figure 3. Generally speaking, these arrows highlight relations between the different normative parts of my project. Arrow 1 points in both directions between "normative claims" and "active reconstruction of biological evidence". With this link I represent the fact that my normative claims (i.e. evaluative and prescriptive) are not ex-cathedra statements but are *informed by and modeled upon* empirical evidence (i.e. in practice metanormativity). This arrow points both ways because the two sets are closely tied together. For example, I make normative claims about which kind of explanation biologists should provide *in light* of empirical evidence, as well as rebuilding

[57] Venn diagrams illustrate possible relationships between elements inside different sets. The elements are usually represented as points in closed curved lines called sets. When two sets overlap, and there are elements in the overlapping space. This shows that these elements have features in common.

biological scenarios which enable me to provide metanormative claims.[58] Arrow 2 represents a link between metanormativity and my object normativity, R. In this case, the double points of the arrow represent the fact that these statements are closely linked. For example, because I argue[59] in favor of the explanatory power of chance in certain biological explanations (i.e. object normativity claims), I thereby suggest that biologists should recognize this role and develop appropriate explanations in light of this argument (i.e. evaluative in practice normativity claims). Arrow 3 goes from methodological normativity to object normativity R and is the only unidirectional arrow. Why so? This is to stress that my final goal is to link chance with the epistemic value of explanatory power, and that all active reconstructions of biological scenarios aim to establish exactly this link. Mostly, the other way around does not work. Indeed, I do not provide R for methodological normativity since methodology is prior to any other steps of the analysis. I develop a PSP project through methodological normativity and then try to achieve R.

With the upper and lower part of Figure 3 I aim to make explicit and to posit the kinds of multifaced normativity at play in my work. I am confident that the reader, with the conceptual tools provided here, will be able to recognize when I make, in the present work, a normative claim, what its dimensions are and of what kind it is. Nonetheless, at this point a final concern arises. Since with Figure 3 I give important insights concerning the relation between the normative and the descriptive parts of my work (i.e. the arrow 1, 2, and 3), nevertheless the status of the normative part is not yet fully clear. If one asks for the source of the descriptive part, I can answer: biological work. But if one asks where the normative part comes from, I am more embarrassed. Focusing on the arrows in Figure 3, one could answer that the source of the normative part is the virtuous circle between the two parts: the normative can influence the descriptive, and the descriptive the normative. But the virtuous circle specifies the relation with the parts, not the source of the normative one. Thus, focusing only on Figure 1 we cannot see where the normative part stems from. In the next section, I widen the discussion to provide a more satisfactory answer to this question.

[58] A problem lurks here. If I base my metanormativity only on what biologists already do, I risk elaborating a work that is just a confirmation of what is already done in science. I will elaborate this point in the next section.
[59] In the next chapter, I will specify the extent to which a concept of chance could have an explanatory role. In this context I only gesture at this role.

4. On the origin of normativity

4.1. "Is-ought" and "back-to-theoretical-assumption" fallacies

Where does the normative part of a normative-descriptive work stem from? How can this normativity be justified? On the one hand, one could affirm that "the normative part of [philosophers'] project[s] is modeled or informed by the descriptive part" (Kaiser 2015, p. 35). Philosophers advocating this position could affirm "that the epistemic norms that *should* hold are simply the ones that are *in fact* accepted in science" (p. 36). But there is an inconsistency here which Kaiser calls the "is-ought fallacy" (p. 35; Kaiser 2019, p. Kaiser 2019, p. 44). She points out that, in this case, "the normative project degenerates to the descriptive project *plus a stamping of one's foot* and an insistence that the norms actually accepted in science are also the ones that should be accepted" (Kaiser 2015, p. 36, my emphasis).[60] This is a kind of fallacy in which the philosopher could be seen as nothing but an agent who confirms what biologists have already done. Craver (2007) tries to escape it by stressing that his account of mechanistic explanations does not only endorse the norms already present in the neurobiological works, but also *justifies* them, explaining why these norms are the good ones for these kinds of explanations. Still, how can he justify this goodness? It could be justified by the fact that biologists already use it – or an a priori argument could be provided. But with either of these, we find ourselves again facing the is-ought fallacy.

On the other hand, the normative part could be independent from the descriptive part. Let us say that it comes from *a priori* intuitions, or from theoretical reflections/commitments/assumptions. In this case, the is-ought fallacy is avoided because philosophers *actually* say "something different" compared with purely descriptive claims. But if this is the case, problems arise with respect to the PSP assumption. If the source of normativity as a theoretical reflection/commitment/assumption is compatible with methodological naturalism,[61] it cannot be true that this source is compatible with an informed PSP project. Indeed, if our normativity has *a priori* origins, we return to the ex-cathedra metanormative dimension, which PSP denies (cf. section 2.3). For the sake of argument, I will call this second fallacy the "back-to-theoretical-assumption". Let's reframe these issues via the following schema:

[60] In the same vein, Bechtel and Richardson (2010) state: "[i]t is uncontroversial to point out that we cannot simply decide what it is right to do by noting what is actually done" (p. 10).
[61] Indeed, saying that the normative part stems from, e.g., *a priori* intuitions does not preclude postulating a kind continuity between science and philosophy.

If we say that the normative part stems from the descriptive part → then we have the is-ought fallacy
If we say that the normative part stems from *a priori* assumptions → then we have the back-to-theoretical-assumption fallacy

We must then be careful not to fall into these two fallacies. How can I escape from this philosophical impasse? Kaiser proposes that philosophers, after the "turn to practice" (Soler *et al* 2014), should find new ways to link normative claims "close[r] – but not too close[r]" (Kaiser 2019, p. 44) to factual (i.e. descriptive) claims. In the same vein as other authors, she proposes the idea of "*reflective equilibrium* to specify how normative conclusions can be drawn from descriptive matters" (p. 44; my emphasis). More specifically, Kaiser writes about carrying out a "mutual adjustment" between normative and descriptive parts (Kaiser 2019, p. 46) until a reflective equilibrium "is reached" (p. 46). I am sympathetic with this suggestion but what does "mutual adjustment" refer to? And when can we say that a "reflective equilibrium" is reached? In the next subsection, I propose a solution to the is-ought and the back-to-theoretical-assumption fallacies by clarifying possible meanings of the concepts of "mutually adjustment" and of "reflective equilibrium".

4.2. The "mutually adjustment" and "reflective equilibrium" argument

As already specified in section 2.3, Kaiser (2015) proposes three tasks in order to make explicit the "act of active reconstruction" of empirical evidences: focusing on relevant data, explicating underlying assumptions, and establishing coherence. In her 2019 paper, Kaiser adds a fourth task: "*mutually adjusting* philosophical claims and empirical information until a reflective equilibrium is reached" (Kaiser 2019, p. 46; emphasis in original). In other words, "the process of developing a philosophical theory by taking into account empirical information from and about scientific practice is not a one-way process but involves a repeated mutual adjustment and moving back and forth between philosophical theory and empirical information" (Kaiser 2019, p. 48). This fourth task could be seen as the process that leads to a "reflective equilibrium" between abstract theory and concrete data. Kaiser specifies (Kaiser 2015, p. 36) that this idea of reflective equilibrium comes from moral philosophy (e.g. Cath 2016; van Thiel and van Delden 2010)[62] and refers to an "exchange" between philosophy and

[62] This idea of reflective equilibrium, namely having pre-concepts and then adjusting them in different steps (i.e. mutual adjustment) through experience, sounds quite similar to the Bayesian idea (and theorem) of probability in

science, philosophical theory and empirical information, pre-concepts and evidence (whatever one wants to call them).

"The process of mutual adjustment often starts with provisional philosophical claims and preconceptions that are brought together with, sharpened, and modified in the light of provisional selections, interpretations, and evaluations of empirical information" (Kaiser 2019, p. 48).

When Kaiser specifies the meaning of "mutual adjustment", it seems to fall into circularity. Here Kaiser specifies that provisional philosophical claims are modified in "light of provisional selection, interpretations, and evaluations of empirical information" and that this modification is what "mutual adjustment" refers to. But selection, interpretation and evaluation are very similar to the first three tasks that she developed (focusing on relevant data, explicating assumptions, and establishing coherence, cf. section 2.2). It seems then that the definition of the fourth task contains elements of the definition of the first three task such as "selection, interpretation and evaluation". At this point, two solutions could be chosen: 1) We admit that the fourth task, being *defined* by the three others, is not at the same level; 2) We recognize the circularity and instead develop a definition that is detached from the other three tasks. I want to follow solution 2) to attempt to give to the fourth task a sort of independence from the others.[63]

4.3 The bird's-eye process

My strategy is to describe a possible chronological order in which the four tasks could be pursued. The underlying assumption of this strategy is that exemplifying a chronological ordering for our tasks could help in clarifying the fourth, vague one. I want to stress that the order I am proposing is not to be conceived of as exclusive or correct. Indeed, different PSP projects can be developed in different ways, and they can also have different methodological time-ordering tasks. My example is only an exemplification of a PSP project. That exemplification concerns the analysis that I will develop in Chapters 5, 6 and 7. For the present context, it suffices to sketch its methodological structure.

which you start the analysis with an initial probability distribution (e.g. credence) and then corroborate it (i.e. adjusting that probability) with empirical data (e.g. see Hacking 2001).

[63] In a personal communication, Kaiser said that she would instead favor the first option, that is, to acknowledge the fact that the fourth task is present in all the other tasks and that it would be interesting to untagle how it might manifest itself in them. Despite this, I would instead like to explore the second option – without denying the fact that future work may synthesize the two approaches, e.g., giving to the four tasks a independent definition that can, nonetheless, be applied to all of the first three tasks.

One of the biological case studies that I will analyze in these chapters is (the use of) the notion of noise in the literature on gene expression. My main aim there is to provide a philosophical argument proposing that chance is "more than just noise" (Huang 2009), i.e. chance is an epistemic element having a specific role in biological explanations. More precisely, I realize that the analysis of recent biological papers on noise in gene expression opens interesting philosophical questions (cf. doing BUPB (bottom-up philosophy of biology) is a gamble, recall section 1.1). For the sake of the argument, I limit myself to mentioning only one: when biologists refer to noise, what do they actually have in mind?[64] This question arises because different papers give different definitions of the notion. I look at the history of molecular biology to find reasons for this heterogeneity.

What, then, is the methodological chronological ordering of this analysis? Selecting and reading biological works about noise in gene expression could be associated with the first task (i.e. sectioning relevant data). After that, having these works in mind, a further step is asking if this literature could be interesting for philosophical investigations. In the case of noise, these works do give rise to interesting questions. For example, as already mentioned: what do biologists have in mind when they write about noise? Attempting to make explicit the underlying assumptions concerning noise and the attempt to discover its historical roots can be associated with the second task (i.e. explicating underlying assumptions). Finally, trying to create a conceptual map for the different concepts of noise provided by biologists can be linked to our third task, establishing coherence. What is still lacking is the fourth task. Yet I have not identified it. The fact is, this task is tricky. Let us focus on the first steps of the analysis in which I said that we select and read biological papers and then, with this literature in mind, see if interesting philosophical questions might arise. I now propose focusing attention in between the act of selecting and reading and the act of developing questions in order to open up a further epistemological space. It is here that I want to locate the fourth task, i.e. mutual adjustment. Indeed, what happens here is something that it is not yet fully stressed. Here we have what I call a "mental process" pursued by philosophers that consists of the following steps: selecting and mapping the works analyzed, synthetizing them into a general scenario, and seeing the full picture. Using a metaphor, we can say that this mental process is like a bird that, after having explored all the specific places, flies up to survey the landscape it has just explored. Thanks to its altitude, the bird can see, and link, very distant places. With this metaphor, what I am

[64] The other questions that I will address in Chapters 4 and 5 are: to what extent we can say that chance has been conceived mainly as noise? Where does the concept of noise comes from? Can we actually reduce chance at a cellular and molecular level to noise?

suggesting is that philosophers have the advantage of flying between different biological domains (e.g. molecular, cellular and developmental biology) and actively reconstructing a general picture from which new philosophical questions could arise. I call this "flying" the *bird's-eye* process (see Figure 4).

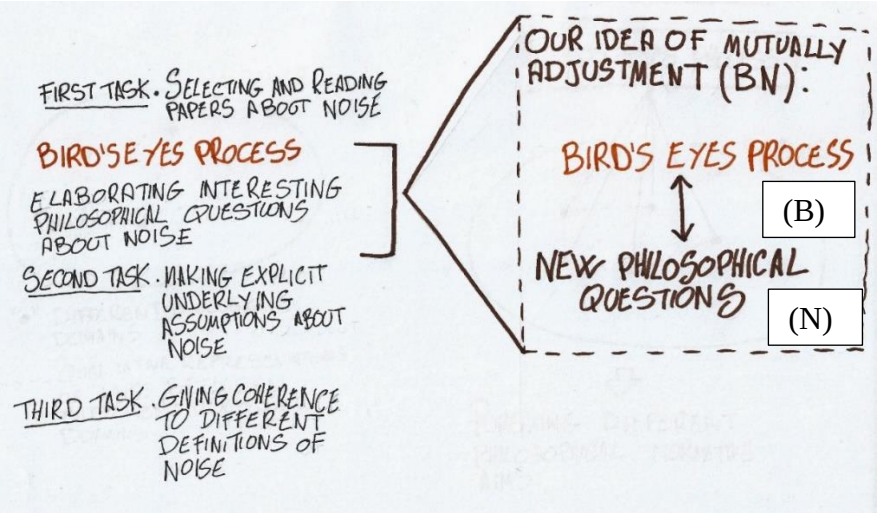


Figure 4. On the left, our methodological ordering of our three tasks (black), and the emerging fourth one (orange). On the right, our idea of mutual adjustment with the bird's eyes process (B) and the new philosophical questions (N). For further details see main text.

This idea of a *bird's-eyes* process could be located in an epistemological space between selecting and reading biological works (first task) and elaborating interesting philosophical questions. But to what extent could this approach help us to make more explicit the notions of mutual adjustment and reflective equilibrium?[65] Indeed, this section tries to elucidate these notions in order to avoid the is-ought and the back-to-theoretical-assumption fallacies. I just wrote that mutual adjustment is intended as a general idea of "going back and forward" between general pre-concepts and empirical evidences. With this idea of a *bird's-eye* process in mind, I can be more precise, especially concerning what the elements involved in this "back and forward" movement can correspond to.

Kaiser (2019) characterizes mutual adjustment as a movement between two quite general elements: pre-concepts and empirical evidence (P↔E). By contrast, we can see, looking at (B↔N), that the *bird's-eyes* process is, itself, the first element of this movement. We can also see that in (B↔N) we do not have any sort of abstract "pre-concept" that philosophers could have *a priori* in mind (such as in P↔E). Nonetheless, it could also be objected that in (B↔N) we do not even consider empirical evidence either! But in order to understand the relevance of (B↔N) we have to relinquish the philosophical *vs* empirical dichotomy. Indeed,

[65] In Figure 4, the reflective equilibrium is not depicted but is meant to be the result of the mutual adjustment between the bird's eyes process and the new philosophical questions. See after.

the *bird's-eyes* process already synthetizes the two aspects: philosophical, because it is the philosopher who mentally initiates the process of mapping biological works; biological, because the "stuff" that is mapped is scientific work. In other words, since in (P↔E) we can read a symmetry (i.e. philosophy is on the right of the double arrow and science on the other side), in (B↔N) this symmetry is lost because the *bird's-eye* process already mixes the two domains. Following (B↔N), the *bird's-eye* process (B) permits the development of new philosophical questions (N). In the same vein, in the new philosophical questions (N), the dichotomy of philosophy vs empirical is also lost. Indeed, philosophical questions could be developed from (and refer to) scientific/biological descriptions, such as the question of whether noise in gene expression can be reduced to the thermal agitation of biomolecules or not. Let's sum up these reflections in the following schema:

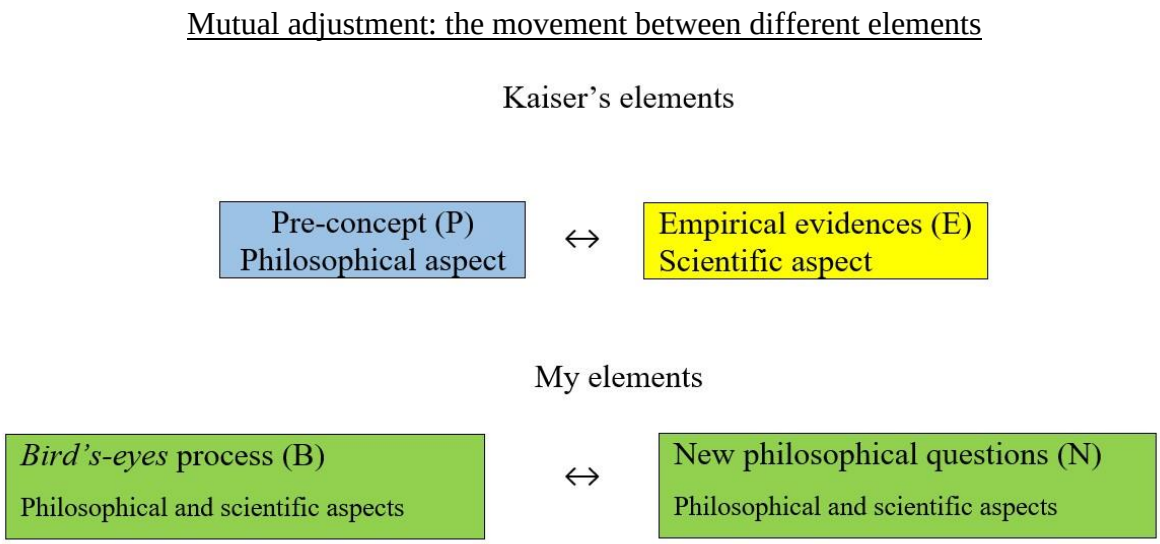


Figure 5 The different elements concerned with the mutual adjustment. At the top, Kaiser's proposal of (P)↔(E) in which philosophical and scientific aspects are separated. At the bottom, my proposal of (B)↔(N) in which the philosophical *vs* empirical dichotomy is lost. For further details, see the main text.

Note also that the movement between (B) and (N) is not unidirectional (i.e. *bird's-eye* process → new philosophical questions, or *vice versa*) but bi-directional (i.e. *bird's-eye* process ↔ new philosophical questions).[66] For example, the *bird's-eye* process allows us to realize that noise is conceived by biologists in very different ways. This awareness raises a new question: could it be the case that all the different notions of noise share some features (i.e. *bird's-eye* process → a new philosophical question)? In light of this general question, one could ask more specifically if, for example, all these definitions (or the majority of them) share a reference to

[66] To be fair, I have to also stress that the original idea by Kaiser of "mutual adjustment" implies a kind of bidirectionality, i.e. the "↔" in (P↔E).

molecular thermal agitation (cf. Chapter 5), as a possible cause of noise. In order to answer this second question, one has to return to looking at biological works (i.e. a new philosophical question → *bird's-eye* process) in order to see if it is the case. To sum up, let us formulate Kaiser's idea of mutual adjustment (P↔E) in terms of my own proposal, (B↔N).

But what about reflective equilibrium? I think this equilibrium is reached when we are able, starting from (B↔N), to pursue the other tasks (i.e. 2) making explicit underlying assumptions. More specifically, this equilibrium is reached when new philosophical questions stemming from (B↔N) enable the pursuit of these other two tasks. For example, in a (B↔N) situation about noise, I mentioned the strategy of asking about thermal agitation as a possible feature shared by the different definitions of noise. For the sake of argument, let us assume that the reference to thermal agitation is a feature shared by virtually all the definitions of noise (the scenario is more complex than that; for an extensive analysis see Chapters 5 and 6). This can then lead us to establish a certain degree of coherence between different definitions, in that we can affirm that most of them do share, at least, the same idea of the origin of noise (i.e. thermal agitation).

The advantage of my proposal (B↔N) compared to (P↔E) is that I give a precise character to the notion of mutual adjustment (i.e. the *bird's-eye* process and new philosophical questions) along with a precise chronological order. I also mentioned one possible meaning of reaching equilibrium as the movement from (B↔N) to the other tasks. Nonetheless, I want to stress again that this clarification does not aim to be universally applicable to all kinds of philosophical projects. Indeed, for this clarification to work, I was forced to be as specific as possible (i.e. exemplifying methodological time ordering with our analysis about noise). This was necessary to provide an argument that did not rest on abstract terms. Therefore, this analysis can work for PSP projects that assume, at least, a bottom-up philosophy of biology (cf. BUPB) attitude. Nonetheless, I cannot exclude *a priori* that other kinds of projects might find benefit from this section – to the extent that they could find certain similarities between their work and my reflections on methodological time ordering and the *bird's-eye* process.[67]

Thanks to the bird's-eye process, I am able to escape from the is-ought and the back-to-theoretical-assumption fallacies (cf. section 4.1) by highlighting the methodological time

[67] For example, in a personal communication, Kaiser suggested to me that it might be interesting to see the reaching of this equilibrium not as the achievement of other tasks but rather as the elaboration of new philosophical questions relevant to scientific practice. Although the idea is interesting, I will leave it aside for future developments of this part of the work that go beyond the ultimate goal of this dissertation.

ordering of my PSP project.[68] This has been the entire aim of this section. Focusing on Figure 5, my project stems neither *entirely* from biology nor *entirely* from philosophy because it is developed through mutual adjustment between the bird's-eye process and the emergence of new philosophical questions. (B↔N) is not entirely biologically-based because it actively reconstructs (i.e. first three tasks) and maps (i.e. the fourth task) biological works. Nor is (B↔N) *a priori* philosophically based, because new philosophical questions stem directly from the bird's-eye process that is not an *a priori* analysis. By contrast, it maps and deals with *corpuses* of biological works. I hope that with my argument regarding (B↔N) and methodological time ordering (cf. Figure 4) might be seen as an exemplification of how normative claims can be "closely […] but not too closely" (Kaiser 2019, p. 44) linked to the descriptive ones, avoiding any fallacies (i.e. the is-ought and back-to-theoretical-assumption fallacies).

One objection to my idea might be the following. At the end of the day, the *bird's-eye* process is not so *useful* and *original* as I suggest. In fact, it is nothing but a revival of the old idea of philosophy as an enterprise with an "eye on the whole" (Sellars 1963, p. 3), that is to say as a discipline that always tries to frame issues from a panoramic point of view to allow a more agile conceptual discussion of the notions under attention. I could agree with this objection if my work were *pure* philosophy. But, as I have stressed in this present chapter, my work is interdisciplinary. As such, the *bird's-eye* process can be, at a minimum, nothing but obvious for a PSP projects, and at best, of great use for these projects.

As a discipline, biology has become a cluster of different and very fine-grained domains of study that seldom speak to each other. This lack of communication could be called the "dark side" of specialization in science. Although specialization has allowed biologists to advance sophisticated questions and provided satisfactory results, it has nevertheless enclosed them in "watertight compartments" – or in "epistemological silos" (Miller *et al* 2008, p. 5). These can be hard to open up. Some scientists understand and underscore the importance of communication and collaboration between different scientists (e.g. Ellis-Davies 2013, p. 64, p. 71; Van de Vijver 2009). There exist interesting philosophical proposals to bridge disciplines (e.g. Eigenbrode *et al* 2007; Miller 2008; O'Rourke and Crowley 2013). For example, developing their idea of "a toolbox",[69] O'Rourke and Crowley (2013) argue that "the challenges

[68] I return again to the fact that this methodological order refers specifically to my PSP project (cf. see the first lines of section 4.3). It should not be intended as *the* order that can be found in all PSP projects. This does not mean, however, that other PSP projects can find similarities with the proposed analysis.

[69] The toolbox approach "aims to address philosophically-based communication issues through a structured dialogue in which participants abstract away from specific disciplinary differences toward the conceptual common

included in [cross-disciplinary research] are, crucially, *communication* challenges" (O'Rourke and Crowley 2013, p. 2; emphasis added). Nonetheless, the value of this collaboration is not universally recognized. In this regard, Miller *et al* (2008) write that "[d]espite decades of attempts to encourage interdisciplinarity, many stakeholders are holding onto a system framed by disciplinary boundaries" (p. 2). In light of this evidence, the *bird's-eye* process could be of great assistance to science and, even while not being philosophically revolutionary, could play an important role in interdisciplinary projects by permitting general frameworks for reflection that are frequently absent in science.

Conclusion

In this chapter I made explicit certain meta-philosophical preliminaries that define my philosophical project as a PSP project (i.e. a project that gives a main role to scientific work). The project deals with the idea of methodological naturalism (i.e. the idea of continuity between science and philosophy) and with the three tasks of the active reconstruction of biological phenomena (focusing on relevant data, explicating underlying assumption, establishing coherence). I also argued that normativity is not a vague term but rather a multifaceted phenomenon. With the help of Kaiser, I linked different parts of my project with different normative dimensions: in practice metanormativity, methodological normativity, and object normativity. Finally, I proposed the original idea of the bird's eye process as a clarification concerning the origin of the normative part in a descriptive-normative PSP project.

What is still lacking is a philosophical argument that justifies the object normativity of my project. Chance is usually associated with disorders, unpredictability and ignorance – which makes my task even harder to pursue. How can chance have any sort of explanatory power? How can it have any epistemic value? The next two chapters will be dedicated to providing a philosophical argument that explains the extent to which I can justify the link "R" provided in Figure 3, that is to say my proposal of objective normativity. In short, it will address the extent to which chance can be said to have explanatory power.

ground they share as research scientists" (O'Rourke and Crowley 2013, p. 8). Even more explicitly, the toolbox provides "structured discussion that facilitates cross-disciplinary understanding by exposing its philosophical dimensions" (Eigenbrode et al 2007, p. 62).

Chapter 2: A critique of the notion of mechanism and mechanistic explanation

Introduction

Why open this second chapter with a critique of mechanism and mechanistic explanation?[70] What is the link between the literature on mechanisms and the object my interest, chance and explanation? I justify this choice for the following three reasons:

1 – The first reason is general but important. When philosophers address explanation, they very often consider mechanistic explanation.[71] In the last two decades, mechanistic explanations have become one of the most widely debated kind of explanation in philosophy of biology (Illari and Williamson 2012). Since my work deals with chance *and* explanation, it is therefore crucial to deal with one of the main accounts that the literature presents concerning explanation.[72] In addition, the mechanistic account of explanation is developed specifically in philosophy of biology and not in philosophy of science more generally.

2 – Philosophers of biology who work, propose and explore the mechanistic account of explanation very often use examples from molecular biology (Baetu 2019), the level of biological organization my work is focused on. It is essential then to consider how philosophers use these molecular examples, which examples they use, and for what purposes.

3 – The last reason is the most important as it is the bridge that allows me to pass from critical analysis to my original account of chance. In the 2000s, the pioneers of the philosophical account of mechanism focused on the notion of regularity, namely the idea that a mechanism can be defined as something that takes place in a regular fashion. Nonetheless, very soon it was clear that what scientists call "mechanism" in actual biology does not behave in regular ways but, on the contrary, often shows stochastic behavior. This was a blow to the very heart of the

[70] Mechanisms, mechanism models, and mechanistic explanations could be linked in the following way: "[it] is useful to distinguish between *mechanisms*, which are actual chains of appropriately causally connected entities in the world, and *mechanism models*, which are descriptions or schemas of such mechanisms used *to explain* the phenomena for which the mechanisms are responsible" (Andersen 2012, p. 416; emphasis added). This distinction appeals me even if it is not central to my main argument. For further detail, see also Glennan (2005) and Illari and Williamson (2011).

[71] These philosophers are also called "new mechanists" and their domain, the "new mechanistic philosophy of science" (Levy and Bechtel 2013, p. 241; see also Skipper and Millstein 2005). Nicholson (2012) and Andersen (2012) note that the terms "mechanists" and "mechanistic" are misleading since they suggest that this contemporary account of mechanism is related in some way to the more classic mechanistic philosophy – which is not correct. If these philosophies share anything, it is only the assonance of the names "mechanism", "mechanistic", mechanists". For a detailed historical overview on this point see also Schiemann (2019). For a further critique see Franklin-Hall (2016).

[72] For the philosophical literature on chance see Chapter 4.

idea of regularity in the definition of mechanisms. In response, some authors (Andres 2012; Bogen 2005) proposed an account of stochastic mechanism in order to integrate stochasticity into mechanistic explanations. I will argue that this integration is not consistent and that it is time to go *beyond* mechanistic explanations to account for stochastic phenomena.

Numerous works already show that mechanistic explanations can be inadequate. To name just one, Moore (2012) underlines that the goal of many structural biochemists who are developing mechanistic "movie[s] of proteins synthesis" is misplaced because "all the functionally significant movements of the ribosome, both internal and external, are biased random walks,[73] and it is most unlikely that any given ribosome will ever do exactly the same thing twice as it elongates some polypeptide" (Moore 2012, p. 8). Skillings (2015) uses Moore's paper to argue that the synthesis of proteins cannot be seen as a mechanism and thus cannot be explained using a mechanistic explanation:

"The mechanism of protein synthesis does not fit neatly into the basic mechanistic account. There is no productive continuity between stages, where earlier stages directly produce later stages. Rather, the mechanism proceeds *stochastically*. Movement through the process of protein synthesis is a *biased random walk*. Watches reliably tell time, and ribosomes reliably produce proteins, but the relations between the temporal stages of the mechanisms are much different" (Skillings 2015, p. 1146; emphasis added).

In recent years, "[b]iological explanations that adhere to a basic mechanistic account […] have often been criticized as inadequately representing and explaining many biological phenomena" (Skillings 2015, p. 1143). Philosophers have suggested thinking about alternative and "non-mechanistic forms of explanation" (Woodward 2013, p. 40), that better explain phenomena that cannot be described in mechanistic fashion. For instance, Green and Jones (2016) propose a "constrained-based explanation in biology" that "[does] not gives as much explanatory weight to mechanistic details about properties of system components" (p. 361) but instead "is driven primarily by the wish to understand the generic features of a class of systems shaped by a shared set of constrains" (p. 361-362). Huneman (2010) suggests going "besides mechanistic explanations, [and relying] explanation upon 'topological' properties of the

[73] The OED (Oxford English Dictionary) defines random walk as "the movement of something in successive steps, the nature of each step being governed by chance independently of preceding steps".

systems in order to derive the *explanandum* as a consequence, and which does not consider mechanistic explanations or causal processes" (p. 213).[74]

In the same vein, I want to propose an "alternative" and "non-mechanistic" philosophical account that I will call "stochastic explanation". The aim of this account is to replace the problematic account of stochastic mechanism and to provide philosophical tools that enable both philosophers and biologists to better understand the explanatory role of chance in biology. From a philosophical point of view, the aim of this project is to provide a convincing argument for the statement that chance can have an essential explanatory role: in certain explanations, stochasticity is fundamental to the *explanans* in order to account for the *explanandum*. From a biological point of view, the aim of this project is to accompany biologists (Kaiser 2015, p. 40) in rethinking and extending the way in which chance is approached and described.

In the present chapter, I proceed as follows. In section 1, I highlight the main problems raised by the notion of mechanism and by definitions of mechanistic explanations. In section 2, I ask whether "more details are better" is a necessary property of mechanistic explanation. In section 3, I provide a critique of stochastic mechanism which challenges the idea of high-failure mechanism. In section 4, I start to introduce my own account by providing the first two criteria of adequacy for an explanation to be labelled as stochastic. In light of the importance of the third criteria, I devote a whole chapter to it (Chapter 3).

1. What is a mechanism and what, then, is a mechanistic explanation?

In this section, I underscore some ambiguities concerning mechanism and mechanistic explanation. I will be not exhaustive, since my aim is only to convince the reader that, even from a general point of view, there are numerous issues difficult to solve. I focus on four points. First of all, I ask why the notion of mechanism is so popular today in philosophy of biology. Second, I show that there is no agreement concerning the definition of mechanism. Third, I highlight that it is also difficult to make a clear distinction between mechanism and causal chains. Finally, I show that linking the notion of mechanism with a mechanistic explanation is no easy task.

[74] Another example of alternative explanation is the design explanation proposed by Wouters (2007). Concerning specifically topological explanation (Huneman 2010) and design explanation (Wouters 2007), Brigandt, Green, and O'Malley (2016) write: "[r]ather than assuming that a model is useful only insofar it explains a biological system in *concrete detail*, current philosophical investigations of design explanations (and of topological explanations) are motivated by the goal of making sense of why some scientists rely on abstract models even in situations *where more detailed models exist*" (p. 8; emphasis added).

1.1 The return of mechanism in philosophy of biology

Since the elaboration of the theory of quantum mechanics (1900s – 1930s), the idea that the micro-world and its behaviours are deterministic and can be explained through classical physics has been seriously questioned (DesAutels 2014, p. 6. on this point, see also Illari and Williamson 2012, and Nicholson 2012). Then "[w]hy, after its spectacular fall from grace at the hands of quantum mechanics, have mechanisms *returned* to favor among philosophers of sciences interested in biology?" (DesAutels 2014, p. 7; emphasis added).

DesAutels suggests three factors that help answer this question. The first tracks back to the Deductive-Nomological (DN) model developed by Hempel (1941). The famous flag-shadow argument against this model shows that there is a deep problem with DN explanation (DesAutels 2014, p. 8): the DN model enables the length of a flag to explain the height of its shadow thanks to reference to the solar altitude. But this model also enables explanation of the length of the flag with the height of the shadow, which is, of course, odd.[75] Philosophers take this as an indicator that something is missing in the DN model. Salmon argues that the reason why the flag pole explains the shadow and not vice versa is that a *causal relation* exists between the former and the latter (see DesAutels 2014, p. 8; see also Salmon 1984, 1989, 1999). This focus on causality has drawn attention to mechanisms that are in fact often thought of as regular causal sequences (see after). The second factor explaining why mechanisms are in vogue in philosophy of biology concerns the well-known problem of there being no laws in biology. Philosophers tried to escape this problem by suggesting generalizations that are not "not-law like" (e.g. Mitchell 1997, 2000, 2003, Woodward 2010). But even among this recent literature, there is, as yet, no consensus about what kind of generality can replace the "law-like" ones.[76] DesAutels suggests, "*it would be nice if we could do explanation in the special sciences without laws*. That is precisely what mechanisms portend to do" (DesAutels 2014 p. 10; emphasis in original).[77] What DesAutels proposes is that the revival of mechanistic explanation is also due to the fact that it permits having general explanations that are "not-law like". A third and final factor is the evidence that science *actually* works and achieves important progress in location and describing mechanisms.[78] The revival of new mechanistic philosophy can thus be explained

[75] For a discussion of this problem, see Gebharter (2013).
[76] The absence of law in biology is also used to critique the DN model. Bechetel and Abrahamsen (2005) write: "[t]he received view of scientific explanation in philosophy (the deductive-nomological or D-N model) holds that to explain a phenomenon is to subsume it under a law [...] Hovewer, most actual explanations in the life science do not appeal to law in the manner specified in the D-N model" (p. 421)".
[77] For a critique on this point, see Illari and Williamson (2011).
[78] DesAutels (2014) writes: "the manifest fact is that life scientists, as a matter of actual practice, have done (and continue to do) incredibly successful science by searching for and describing mechanisms" (p. 11).

by a synergy between the advent of the notion of causation in scientific explanation, the lack of laws in biology, and the practical success of the notion of mechanism and mechanistic explanation. But when did this revival take place?

1.2 The ambiguity of the notion of mechanism

I identify the start of the revival in the 1990s (Bechtel and Richardson 1993) and the early 2000s when work by Machamer *et al* (2000) provided a well-known definition of mechanism:

MCD: "[m]echanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (p. 3).

Even if this is one of the most famous definition of mechanisms, virtually all aspects of MCD have been criticized. For example, a recent work by Kaiser and Krickel (2018) proposes replacing the term "entities" and "activities" with "object" and "occurrence", these last terms being metaphysically less ambiguous.[79] More generally, since in the literature there are numerous definitions of mechanism, achieving any agreement presents a hard task.

"A mechanism for a behavior 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, p. S344)

"A mechanism is a structure performing a function in virtue of its components parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena". (Bechtel and Abrahamsen 2005, p. 47).[80]

These last quotations are only two examples of definitions of mechanism but their differences are sufficient to convey the vastly different approaches concerning the topic. Illari and

[79] More specifically, Kaiser and Krickel (2017) argue that in metaphysical discussions "the term 'entities' is used as an umbrella term for everything that exists" (p. 754), and that the term "activities" is a "problematic notion" for the reason that "activity" plausibly characterizes a specific kind of process, namely a "causal processes that involve changes and an object that is 'active'—whatever this exactly means" (p. 755). "Machamer *et al*. introduced the notion of an activity primarily to account for the fact that 'mechanisms do things', that '[t]hey are active' ([2000], p. 5). Whatever is meant by these quotes, the idea that 'mechanisms do things' seems to exclude the possibility that mechanisms are composed of states that merely consist in the instantiations of dispositions (rather than in their manifestations)" (Kaiser and Krickel 2017, p. 755).

[80] The collection of these definitions comes from DesAutels (2014, p. 12). Another list can be found in Illari and Williamson (2012).

Williamson (2011) note that the only agreement among mechanist philosophers is the fact that "mechanisms explain" (p. 818). In light of that, in another work dated 2012, they aim to develop a general definition of mechanisms which are thought to work across science, emphasizing the fact that "there is a core of agreement" (p. 129) among all different definitions which refers to activity, entities and organization.[81] Nicholson (2012) convincingly argues that the term "mechanisms" is often used by philosophers (including the authors of MCD) and biologists in a confusing manner, and that only an analysis of the history of the concept could prevent this confusion.[82] Moreover, some philosophers admit that they do not actually know what scientists have in mind when they write about mechanisms:

"I certainly do not claim to have access to *what scientists actually mean* when they use the term 'mechanism'" (DesAutels 2014, p. 2933; emphasis added).[83]

Woodward (2013) also underlines the ambiguity of the term:

"An obvious worry raised by arguments over whether there are non-mechanistic forms of explanation, either in biology or elsewhere, or whether all biological systems are 'machine-like', is that such disputes are (or threaten to become) *largely terminological or semantic*—the answer depends on what is meant by 'mechanism', 'mechanical explanation', and so on, and these are notions without clear boundaries" (p. 40; emphasis added).

Woodward pushes the problem in a semantic and terminological dimension: "mechanism" is problematic because each philosopher (and biologist) could rely upon different semantics, meaning that when they write "mechanism", they actually mean different things. Woodward questions if all biological systems have to be seen as "machine-like". Machamer *et al* (2000) themselves admit that the idea of mechanism that they have in mind is not limited to the classic

[81] Their proposal is as follows: "[a] mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon" (Illari and Williamson, 2012, p120). Even if one agrees with this definition (and it is not trivial; see Krickel and Kaiser 2018), Illari and Williamson (2012) still do not focus enough on stochastic phenomena nor how a mechanistic account of explanation can account for these phenomena.

[82] More specifically, Nicholson (2012) recognizes three semantics to the term: 1) "mechanicism" that refers to the "philosophical thesis about the nature of life and biology" (p. 152). The root of this meaning could be traced back to Galilei, Descartes, Gassendi, etc., namely the fathers of what he calls the "mechanistic philosophy"; 2) "machine mechanism" refers to "the internal workings of a machine-like structure" (p. 152); 3) and "causal mechanisms" that are nowadays "of fundamental importance in sciencific practice because they enable the identification of causal relation" (p. 154).

[83] In his PhD thesis, DesAutels, more optimistically, writes: "Recently, much work in the philosophy of science has been devoted to understanding *what exactly it is* that scientists look for when they search for mechanisms" (DesAutels 2014, p. 1; emphasis in original).

Newtonian view of "mechanical (push-pull) systems" (Machamer *et al*. 2000, p. 2), because since Newton first articulated the law of universal gravitation, many other forces (different from push-pull ones) have been discovered. Bearing in mind the example of neurotransmitter release, Machamer *et al*. (2000) give a list of all the other phenomenological forces that can take place: geometric-mechanical, electro-chemical, energetic, and electro-magnetic (p. 14). But philosophically speaking, adding a list of new forces does not tell us if biological systems actually correspond to "machine-like" phenomena (e.g., see Henning and Scarfe 2013) for the very simple reason that a list is not an argument.

1.3 The ambiguity between mechanisms and causation

Another ambiguity that I want to stress here is the problematic relationship between mechanisms and causation. In this regard, Krickel (2018) writes:

"[t]here is an ambiguity in how the term 'mechanism' is used in the new mechanistic literature […] Sometimes the term is used to refer to *machine-like structures* (for example, Glennan 1996); at other times, it is meant to refer to *causal sequences*" (Krickel 2018, p. 1127; emphasis in original).

In this quotation, Krickel seems to be telling us that a mechanism is a machine-like structure *or* a causal sequence (for more detail see also Nicholson 2012) as if these two possibilities were mutually exclusive. But could it be the case that a "machine-like" phenomenon *is* a causal sequence? It seems that Krickel excludes that option. But doubt still lingers, and it is legitimate to wonder, more generally, what the relation between causal sequences and mechanisms actually is. Is a description of a mechanism reducible to the causal sequence that underpins it? If yes, why do we not give up the term "mechanism" or "machine-like structures" in favour of the notion of causal sequences? This could be a reasonable choice since theories of causal explanation are quite convincing and sophisticated in the philosophy of science literature (e.g. Woodward 2005; Strevens 2008) and could be of great help in describing biological dynamics. And if no, what are the "further conditions" (Woodward 2013, note 1, p. 45) that make a mechanism something more than a causal sequence? Andersen (2012) writes that what can ensure the boundaries between mechanisms and causal sequences is the concept of regularity. What makes a mechanism, indeed, is the fact that they are regular in a fashion that causal

sequences are not (on this point see Andersen 2012; DesAutels 2014, p. 38; see also section 1 and 3.2).[84]

1.4 The ambiguity of the relationship between mechanism and mechanistic explanation

The last issue that I want to mention here is that the relationship between mechanisms and mechanistic explanations is not clear. Intuitively, one could assert that when the definition of mechanism is settled (e.g. MCD), it is possible to provide mechanistic explanations using that definition. But the matter is more complex than that since there are different accounts of explanation, as well as different definitions of mechanisms, as already mentioned. Traditionally, there are two different views of scientific explanation. The epistemic view postulates that "explanations are *representations* that meet certain epistemic requirements. Events are explained by subsuming them under appropriate representations" (Boone and Piccinini 2016, p. 2; emphasis in original). The ontic view of scientific explanation (Craver 2007; 2014) postulates that "the adequacy conditions for explanations *are facts in the world*. To explain a phenomenon is to identify the explanatory facts themselves" (Boone and Piccinini 2016, p. 2; emphasis added).[85] We can summarize these accounts as follows:

The epistemic account of explanation → explanations = representations of the world that meet certain epistemic requirements

The ontic account of explanation → explanations = things in the world

Which account explanation of mechanisms should be embraced? In the case of an epistemic account of explanation, mechanistic explanations are not things in the world, but they correspond to our representations of real phenomena.[86] By contrast, in adopting the ontic

[84] Certain authors suggest that mechanisms and causes are distinct by stressing a kind of independence beween them, asserting that we can have causes without mechanisms and mechanisms without causes (see Baedke 2018, p. 155; Weber 2008). Others focus on the discussion's metaphysical dimension. For example, Krickel (2018) proposes a detailed distinction between etiological and constitutive mechanisms; Kaiser and Krickel (2017) give a quite sophisticated elucidation of the relation between mechanisms and phenomena in the case of constitutive explanation.

[85] A quote by Krickel sums up the ontic account very well. "According to the 'ontic view of scientific explanation' (Craver 2007, 2014), explanations are 'objective portion[s] of the causal structure of the world' that 'are not true or false', 'they just are' (Craver 2014, p. 40), and do 'not depend on the existence of intentional agents' as, for example, explanatory models do (Craver 2014, p. 36). Furthermore, ontic explanations 'consist in all and only the relevant features of the mechanism in question' (Craver 2014, p. 40). Hence, according to the ontic view of scientific explanations, explanations and mechanisms are one and the same thing, which exists in the external world, independently of us" (Krickel 2018, p. 1124).

[86] I do recognize that this it is at best only one position among others (and at worst just a simplification). The accounts of mechanistic philosophy are more sophisticated than that. For example, Glennan (2002, 2005), echoing Wimsatt (1972), affirms that, if most of the biological work aims to explain phenomena by discovering mechanisms, then it is reasonable to say that the models that biologists develop are models of mechanism. Then

account of explanation, mechanistic explanations should be identified with the facts themselves (i.e. mechanisms in the world). The ontic account is popular in the literature of new mechanistic philosophy (e.g. Craver 2007, 2014). Nonetheless, "[t]he ontic view has been criticized mainly on the grounds that explanation cannot be successful without obeying some epistemic norms" (Krickel 2018, p. 1124), such as simplicity, cognitive salience and precision. It seems then that both epistemic and ontic accounts of explanation have to meet certain epistemic requirements. For the former this is not a big deal. The epistemic account by *definition* has to meet epistemic requirements. But for the ontic account, meeting these requirements could mean collapsing into the epistemic account, since for the latter, nothing else is required but the explanation itself. Adding epistemic notions such as simplicity or generality means biasing the identification of the explanatory factors themselves (see Boone and Piccinini 2016, p. 2).[87] The other way around, it was noted that since "there [are] real mechanisms in the world" to be described (Illari and Williamson 2012, p. 121), both epistemic and ontic accounts of explanation have to imply real mechanisms. If this assumption creates no problems for the ontic account, it does for the epistemic since it seems weakened by the need to have an ontic account of mechanisms. In general, framing mechanistic explanation in the epistemic-ontic dichotomy is quite problematic – not least for the very simple reason that authors are not always explicit about their position and about the meaning of the ontic and epistemic notions (Illari 2013). It is therefore very difficult to take one position over another without getting stuck on the theoretical problems to which two accounts give rise. From my side, since I do not aim to develop an account of mechanistic explanation, I do not feel obliged to take either side. My account of stochastic explanation is based on very different assumptions (see section 2 and 3), and starts from the simple fact that I do not take a position on mechanism. My intention here was, rather, to call these issues to the attention of the reader.

In summary, the original purpose of the new mechanist philosophy was, and still is (e.g. Glennan 2017), to grasp what scientists and biologists mean by mechanism. My fear is that what is meant by "mechanism" is not monistic, clearly demarcated in science itself. The issues

for him, it is not the case that mechanisms are representations of the world but that mechanical models consist of representations of mechanisms in the world. For the sake of the argument, I do not enter into details that are not relevant to the argument, and which consist only in showing the problematic nature of the relationship between mechanism and explanation.

[87] To try to solve the problem, some authors propose a weak ontic account of explanation (see Kaiser 2015, p. 242) which allows certain epistemic requirements. Krickel, embracing a weak ontic account of explanation, writes: "I will take explanations to be descriptions or models that are not identical with mechanisms but are made true by them" (Krickel 2018, p. 1125). What then is the relation between mechanisms and scientific explanation in this weak account? "Mechanisms are supposed to be the truthmakers of scientific explanations" (p. 1129). In this way, Krickel avoids the strong ontic account, in favour of a weak ontic account that accepts some epistemic requirements such as viewing mechanisms as truthmakers of scientific explanation.

addressed in this first section could be proof of that. It could be the case that in different domains, "mechanisms" have different properties and features,[88] in which case, trying to give a unique definition to mechanism would be not the right way to understand biological processes, although some philosophical efforts have been made in this direction (Illari and Williamson 2012). In the next sections (2 and 3) I focus on further issues. More precisely, I ask if the premise that "more details are better" is necessary requirement for a mechanistic explanation, and if a high-failure mechanism is a good example of a stochastic phenomenon. These reflections aim to convince the reader that it is time to go *beyond* mechanistic explanations, and think in terms of "non-mechanistic forms of explanation[s]" (Woodward 2013, p. 40).

2. "More details are better" is not a necessary property of mechanistic explanations

Intuitively one could assert that mechanistic explanations imply that "more details are better" since the more that can be described and explained about a mechanism, the more that mechanism appears clear and understandable. But even if it is true that in new mechanistic literature we can find the idea that maximizing detail makes an explanation better (i.e. "more details are better", hereafter MDB),[89] nonetheless, it remains for most of the time *just* an idea. In this section, I will show that in the hallmark papers on mechanistic explanation, MDB is nothing more than an *implication* with non-explicit arguments and that MDB is not a necessary property of mechanistic explanation. I focus on this MDB aspect because it will be useful to me when I specify what my account of stochastic explanation is *not*. My account does *not* need the criterion of MDB and I manage this by identifying two clusters of papers. The first refers to the pioneers of mechanistic philosophers (Machamer *et al* 2000; Craven 2007; Kaplan 2011) who tried to develop an idea of mechanisms and mechanistic explanation. The second refers to philosophers who developed the more recent account of abstract mechanism (Boone and Piccinini 2016; Levy and Bechtel 2013; Matthewson 2020; Levy 2014)[90] which points out the

[88] Nicholson (2012) writes: "[i]n a nutshell, what a historically informed perspective reveals is that the term 'mechanism' has come to be used in biology *in a number of different senses*" (p. 153; emphasis added).
[89] We follow the acronym MDB proposed by Craver and Kaplan (2020, p. 288).
[90] Actually, I can trace the root of these projects on abstraction to an earlier time. The Hodgkin–Huxley (HH) model (1952) of the action potential of neurons is the model from which this discussion started (Craver 2006, 2007, 2008; Bogen 2005; and Kaplan and Craver 2011). "Craver thinks that the HH case buttresses the mechanistic approach because it demonstrates how models that do not fully decompose a phenomenon suffer from an explanatory deficiency" (Levy 2014, p. 470). By contrast, other philosophers (e.g. Weber 2008; Levy 2014) underline, in different ways, certain explanatory merits of the HH model not spite of, but thanks to, its abstract dimension.

importance of abstraction in definitions of mechanism.[91] The result of my analysis here is that MDB is not a necessary property for the first cluster and that the second cluster recognizes this fact and makes it more explicit than the first cluster itself.

Let us start with the first cluster of papers. Craver and Kaplan (2020) note that "unfortunately, none of the critics [of the first cluster] defines MDB precisely" (p. 307). For example, Machamer *et al* (2000) write about mechanism as "*a truncated abstract description* of a mechanism that can be filled with descriptions of known component parts and activities" (p. 15; emphasis added). But they do not say much more than that. It is rather the reader who may interpret "truncated abstract description" as an insufficient description that has to "be filled" with other details. But Machamer *et al* (2000) do not elaborate the point, leaving it open to interpretation.

The second cluster describes the tendency of the first cluster:

"The ontic conception of mechanistic explanation [first cluster] has been interpreted as *implying* that the more (relevant) details an explanatory description includes, and hence less abstract it is, the better it explains" (Boone Piccinini 2016, p. 2; emphasis added).

"Contrary to what is often *asserted or implied* [by the first cluster], mechanistic explanation does not require maximal detail" (Boone and Piccinini 2016, p. 10; emphasis added).

"*The suggestion* [by the first cluster], it appears, is that the more detailed, concrete accounts are also more explanatory" (Levy and Bechtel 2013, p. 258; emphasis added).[92]

We can see from these quotations how the second cluster defines the first cluster's commitment to MBD only as an assumption, implication or suggestion – not as a necessary condition. In addition, the second cluster not only recognizes this soft commitment but goes further in specifying that often the first cluster is actually looking for a balance between detail and abstraction. Boone and Piccinini (2016), "members" of the second cluster, write:

[91] More specifically, they start complaining that the notion of abstraction is not sufficiently underlined and explored in the literatue, unlike other aspects of mechanist explanations, such as completeness and specificity (see e.g. Darden and Craver 2002; Baetu 2015).

[92] "Craver makes this tendency even more explicits. He defines 'a *mechanism sketch*' to be 'an incomplete model of mechanism'. It characterizes some parts, activites or features of the mechanism's organization, but it leaves gaps (2007, p. 113)" (Levy and Bechtel 2013, p. 258; emphasis in original).

"Some statements by some mechanist philosophers (e.g., Craver 2007, 114; Kaplan 2011, 347) are sometimes interpreted as implying the requirement of maximal detail, *although the same philosophers also defend the use of abstraction within explanation*" (p. 2, note 2; emphasis added).[93]

More recently, (some) authors of the first cluster, who wrote on mechanism in the 2000s, better clarified their position. Craver and Kaplan (2020) confirm the intuition of the second cluster by saying that they were never completely committed to MDB, but were rather looking for a "delicate balance between the appreciation of the importance of a multitude of biological details and the ability to see beyond those details to general principles" (p. 288).

All in all, for at least two reasons, I cannot assert that the literature on mechanistic explanations necessarily implies MDB:

- The first cluster is not clear on MDB. Some suggest MDB, other not. I share the analysis of the second cluster on this point, namely the fact that in the first cluster MDB is nothing but an assumption/implication/suggestion, and the fact that the first cluster is actually searching for an equilibrium between detail and abstraction;
- The second cluster is itself part of the new mechanistic philosophy, and this explicitly rejects MDB. Even speculating that the first cluster deals with MDB, the second does not. Therefore, saying that *all* new mechanistic philosophy deals with MDB would mean ignoring the complexity of this kind of literature.

I have spent a lot of time on this issue because this point will be crucial in section 4.2, where I will go into the detail of my own account of stochastic explanation. For now, it suffices to say that MDB is not a necessary property of mechanistic explanation, for the reasons I mentioned in this section.

Keeping in mind the general problems and issues of mechanism and mechanistic explanation, in the next section I analyze the specific account of stochastic mechanisms. This account aims to extend the mechanistic framework to also account for stochastic phenomena. My aim will be to show that this extension does not work and that it is time to go beyond the mechanism framework in order to properly account for phenomena described as stochastic.

[93] In a similar fashion, Green and Jones (2016) write: "[t]here appears, accordingly, to be a conflict in the existing philosophical literature, between those who prioritize the explanatory force of mechanistic detail and those who emphasize the irrelevance of this same detail" (p. 344). However, at the same time they admit that these two tendencies can co-exist since they recognize "a certain diversity of explanatory strategies in biology" (Green and Jones 2016, p. 344).

3. Stochastic mechanisms

Stochastic mechanism is a notion which has appeared more recently in literature that tries to fix some important issues to which the notion of mechanism gives rise. Notably, *stochastic* mechanism means leaving behind the notion of regularity – which most mechanisms do not seem to possess – and at the same time retaining the notion of mechanism as valuable. Regularity generally refers to the fact that mechanisms take place and provide results all the time in the same way. If a mechanism is structured in the way that A gives B and B gives C, the property of being "regular" means that, if it takes place, it will always (or almost always) take the form A, B and then C.[94] Note that the literature on mechanisms in philosophy of biology typically assumes that a stochastic mechanism *cannot* be regular.[95] Stochastic mechanisms are often also labeled as high-fallibility mechanisms; most of the time they fail to achieve their results. In light of this semantic, it is clear now why stochasticity cannot be associated with the idea of regularity mentioned. If by "regular" I mean a mechanism that takes place and provides results all the time in the same way, a stochastic mechanism is not regular since most of the time it fails to take place and/or to provide the same results (see section 3.2). In this and the next two sections, I show that the notion of stochasticity is quite problematic if it is read in a mechanistic framework (see section 4). In addition, I stress that the very meaning of stochasticity that these philosophers use is inadequate for a deep understanding of chance biological dynamics. The problems that I address are: the problem of regularity (section 3.1), which is to say that phenomena described as stochastic cannot be labelled as regular and that numerous mechanisms are in fact stochastic; and the problem of high-failure mechanisms (section 3.2), which is to say that the description of stochastic phenomena are richer than the description of high-failure mechanisms themselves.

3.1 Regularity and mechanisms

The main definitions of mechanism presuppose some sort of regularity (cf. section 1) as being a necessary (even if not sufficient) property of a phenomenon for it to be labelled as a

[94] Another kind of regularity is also at stake here: "regular" could also refer to the fact that a mechanism takes place at certain fixed time intervals. For example, mechanism X takes place in a system each Δt_x. But since the literature on mechanisms does not seem to focus on this aspect of regularity, I do not take it into account in the present discussion.

[95] In this section, I limit myself to describing the idea of regularity by new mechanistic philosophers. In Chapter 4, section 2, I argue that regularity and stochasticity can be compatibile. Cell differentiation implies different forms of stochasticity, nonetheless embryogenesis shows a high degree of regularity. What I have in mind here is that a stochastic phenomena at one level, for example molecular, can show regularity at another level, for example in tissue formation.

mechanism. We can easily find this "metaphysical prerequisite" (DesAutels 2014, p. 30) in the original MCD definition below:

MCD: "[m]echanisms are entities and activities organized such that they are productive of *regular changes* from start or set-up to finish or termination conditions" (Machamer *et al* 2000 p. 3; emphasis added).[96]

Andersen (2012) challenges this (imprecise) idea of regularity with what she calls "the argument of science" (p. 418). She writes that in science, biologists do not use conceptions of mechanisms that need (whatever sort of) regularity.[97] This argument echoes the work by Bogen (2005) that convincingly shows that the paradigmatic example of neurotransmitter release[98] – used in hallmark papers by Machamer *et al* (2000) and Craver (2007) – is not regular, and that biologists conceive it as a stochastic phenomenon. Bogen writes that this kind of mechanism fails up to 90% of the time (see DesAutels 2014, p. 36; see also Kandel *et al* 2013). "Mechanists need not include regularities and invariant generalizations in their account" (Bogen 2005, p. 399 quoted in DesAutels 2014, p. 32). This empirical based argument is so compelling that Machamer himself, one of the authors of the MCD definition, reconsiders his position.

"I think 'regular' should be dropped from the definition. Jim Bogen argues forcefully that there might be mechanisms that operate only once in a while or even one that works only once" (Machamer 2004, p. 37 quoted in DesAutels 2014, p. 32).

We can find another example of rethinking the concept of regularity by Glennan, another prominent author of new mechanistic philosophy, who proposes the concept of "ephemeral

[96] A similar appeal to regularity can be found in most of the literature on mechanisms (for an overview see DesAutels, 2014, p. 32).

[97] In order to resist to the argument by science, Andersen proposes broadening the notion of regularity in a taxonomy, "something like a map of the territory across which mechanisms can vary with respect to regularity" (Andersen 2012, p. 419). She proposes three main parameters for regularity: the organizational location, the strength of connection, and the failure pattern (for further detail see Andersen 2012).

[98] This is the main example used by virtually all the main philosophers of the new mechanicism. A very famous analysis is found in Machamer *et al* (2000). Roughly, we can schematize the mechanism that explains neurotransmitter release with a causal chain where "→" has to be intended as "causes". The general scenario that we have to keep in mind when describing this chain is the so-called synaptic cleft, namely a scenario in which a neuron axon meets the dendrite of another neuron. Then, an action potential that reaches the axon terminal → calcium channels to open that → calcium to influx into the axiom terminal that → synaptic vesicles filled with neurotransmitters to fuse with the membrane of the axon terminal that → the release of the neurotransmitters in the synapsis cleft. At this point, the neurotransmitter binds to the receptors located at the postsynaptic membrane of a dendrite of a new neuron. Finally, in the dendrite of the new neuron, the chemical signals will again be transformed into electrical signals. This permits the signals to continue the propagation. I based my reconstruction on Krickel (2018, p. 1125) but detailed descriptions of this mechanism can be found in any neurobiological textbook.

mechanisms" that do not appeal to regularity as a defining feature of mechanisms (e.g. see Glennan 2010, 2017). These authors try to overcome the problem by removing the notion of regularity from the definition of mechanisms; other authors follow a different strategy. Instead of giving up, they defend a sophisticated version of regularity designed to resist the issues mentioned. In the next section, I briefly analyze this second kind of strategy.

3.2 High-failure mechanisms as stochastic mechanisms

Generally, the authors that develop a sophisticated version of regularity have to face the fact that the neurotransmitter release mechanism *actually* fails up to 90% of time (Boges 2005). It is not *in prima face* regular. Because of its high fallibility, neurotransmitter release is also labelled as a "high-failure mechanism" (see e.g. Andersen 2012, p. 419; Krickel 2018, p. 1126). In light of that, Krickel (2018) elaborates an interesting and rigorous analysis (which takes its cue from Andersen 2012) with the aim of answering the following question: despite its high fallibility, to what extent can the mechanism of neurotransmitter release be accounted for by a type-level explanation? Very roughly, when we use type-level explanation we are explaining the behaviour of a mechanism in general – we are providing an explanation that works for *all* mechanisms of the same kind. By contrast, when we explain by appealing to token-level explanation, we are explaining an *individual and actualized* mechanism.[99] The very general question is, how could we provide type-level explanations of mechanisms, even if they very often fail to be regular, and seem to satisfy only token-level explanations? More to the point, what kind of generality, if any, can we attribute to the explanation of high-failure mechanisms in order to provide a type-level explanation? (Krickel 2018, p. 1127).

Krickel (2018) writes that for Andersen (2012, p. 421) "high failure mechanisms are regular if in cases of failure, one can identify interfering factors, or if there is a consistent percentage of times where the mechanism succeeds (which might be rather low)" (Krickel 2018, p. 1126). Krickel challenges this idea saying that 1) not all high failure mechanisms are regular in either of these two senses; 2) Andersen does not specify how such notions of regularity can be implemented in an account of type-level mechanistic explanation (Krickel 2018, pp. 1126-1127). In this vein, she proposes two new versions of regularity that have to be added to

[99] I leave open the question of whether a mechanism is something in the world or something that explains a phenomenon in the world. It is not necessary to take sides here, since my aim is only to highlight that stochastic mechanisms are *conceived* by philosophers as high-failure mechanisms regardless of whether they are in the world or only explanations of it. For further discussion on this point see section 1.4.

Andersen's:[100] the comparative regularity that postulates that "the mechanism brings about the phenomenon more often than any other phenomenon" (p. 1124), and the comparative reverse regularity in which "the phenomenon is more often brought about by [a specific] mechanism than by any other mechanism or causal sequence" (p. 1124). She uses these sophisticated notions of regularity to assert that, if high-failure mechanism is regular in at least one of these two senses, it can be still be considered as "mechanism" and so provide for type-level explanations.

My critique of this point is that in addition to the main critique that neurotransmitter release is not a genuine example of mechanism because of its lack of regularity, *nor is it a paradigmatic example of a description of a stochastic phenomenon*. In the neurotransmitter example, the attribution of "being not regular" is reduced to the fact that most of the time the mechanism *does not produce the expected outcomes.* It fails to transmit the signal. As already stated, in the new mechanistic literature the property of "fallibility" is strictly linked to the property of being "stochastic":

"[t]he idea that there are stochastic mechanisms—mechanism types that have instances that *do not bring about the phenomenon*—is commonly accepted among the new mechanists" (Krickel 2018, p. 1134; emphasis added).

I argue that there are at least three reasons why using the label "high-failure" to describe a stochastic phenomenon is misleading:

1. *"High-failure" does not help in the explanation of a phenomenon in general and in the explanation of phenomena (described as stochastic) in particular*. Whether we refer to a mechanism as something in the world, or as a representation or model that explains a phenomenon in the world, if we use the label "high-failure" we are describing a phenomenon that fails to take place and/or that does not give the expected results. The first problem of this approach is that this description is not explicative. Describing a phenomenon as a "failure" prohibits advancing knowledge of it, since this label defines the object of attention through a deficiency. Nonetheless, I am not saying that *it is always the case* that knowledge advancement cannot take place through negative definitions, for example with definitions that specify the object of interest through

[100] We recall that Andersen (2012) provides a taxonomy of regularity, in which she shows different ways in which mechanisms can be (ir)regular.

something that it misses. For instance, Sorensen (2011) proposes a quite convincing definition of parasitic natural kinds in order to characterize something that is not a natural kind but that nonetheless, precisely thanks to this negative characterization, finds its conceptual clarification.[101] What I am saying here is that "failure" does not help progress in explaining and understanding biological phenomena in this specific scenario.

2. *The "high-failure" label implies unjustified evaluative claims*. To say that something (e.g. a process) fails to achieve its result is to say that it does not achieve its goal, its *function*. To some extent, by using the label "high-failure", we are linking the concept of failure with the notion of (non)function. But the concept of (non)function is problematic. On what bases can we outline the functional/dysfunctional difference in biology? In the literature of philosophy of biology, there is a long-standing debate questioning the meaning of function in biology (Cummins and Roth 2009; Kitcher 2003; Lewens 2004; Perlman 2010). In recent years, debate on this topic has started in the biological fields themselves (Doolittle 2018; Keeling *et al* 2009). A careful preliminary analysis is needed in order to adequately frame the meaning of functionality, nonfunctionally and failure. But in its absence, the use of these notions carries issues and evaluative claims which are difficult to resolve and to justify.
3. *"High-failure" is only one way, among many, to describe stochastic phenomena*. The notion of "stochasticity" is richer than the notion of high-failure, the latter being only one instance, among others, of the different ways a phenomenon can be labelled as stochastic. Thus, as such, stochasticity-as-high-failure fails to be representative. In order to illustrate this point, recall that the irregularity most new mechanistic philosophers try to challenge is the fact that the neurotransmitter is not regular because it fails to give its output up to 90% of the time. It is, then, stochastic. Thus, we can say that stochasticity-as-high-failure refers to the fact that a mechanism or a phenomenon could or could not happen. Let us say that M is the mechanism of the neurotransmitter and that (x) represents the successful outcome of M, where M succeeds in giving its output. The unsuccessful outcome of M is represented by not-(x), where M fails to give its output.

[101] Sabbarton-Leary *et al* (2015), sum up very well the idea proposed by Sorensen (2011), writing that "parasitic kinds are *not* natural kinds. They are *absences of actual natural kinds"*. If heat is (roughly) defined as particles in motion, cold is the absence of such motion. The kind "cold" therefore parasitizes the natural kind "heat" and it is therefore defined as parasitic kind. Although parasitic kinds are not a natural kind they "inherit the lawfulness and projectibility of the natural kinds that shape them" (quoted by Sabbarton-Leary *et al* 2015, p. 87 that refers to Sorensen 2011, p. 117).

In this scenario, stochastic behaviour resides on the chancy switch between (x) and not-(x). The literature builds the features of stochastic mechanisms using this as its paradigmatic example. But the methodological problem I stress here is that "M could give (x) or not-(x)" is *only one way* through which a phenomenon can be described as stochastic. Indeed, another phenomenon, P, could be described as stochastic in other ways too, for instance: P could give (x) or not-(x) or (y) or (z) and so on; P could give (x) and/or not-(x) and/or (y) and/or (z) and so on. Thus, by limiting stochasticity to high-failure mechanisms, we risk losing the richness of the different ways in which phenomena can be described and explained as stochastic in biology. In the next chapter, I will argue that the explanatory role of chance consists exactly of these richer ways in which a phenomenon can be described as stochastic.

Let us now leave behind notions of mechanism and high-failure in order to propose an account of chance that emphasizes the role of chance in biological explanations. In the next section, I introduce the first steps of my account of what I call "stochastic explanation".

4. Stochastic explanations

In this section, I provide a description of what I mean by "stochastic explanation" (hereafter SE). Very generally, SE is a philosophical category that refers to any biological explanation in which chance has an explanatory role. I will show that while some SEs are already present in biology, others have to be discovered/formulated. The main aim of SE is to add further philosophical meaning and a value to chance in biological research. SE is the result of a theoretical reflection that I built up from empirical evidence. But science talks with many different voices and what matters here is to make explicit the origins of my examples, since this choice influences the results of my work to the extent that different sources may imply different dimensions of normativity (cf. Chapter 1). I call the biological scenario from which I pick these examples the "example landscape", and I make it explicit in the next section. After that, I give a definition of stochastic explanation which underlies two of the three criteria of adequacy that I want to provide for this account. In the following chapter, I describe the third.

4.1 The biological "example landscape"

In the present thesis, I use examples that come from molecular, cellular and developmental biology. I have two main sources: textbooks and scientific articles. By textbooks I mean all the scientific books provided for any program teaching biology in any university. By scientific

articles, I mean all kinds of academic publication (e.g. journal article, original papers, review, books, monographs, edited books, etc.) that aim to contribute to research.

1. *Textbook examples.* I use several case studies from biological textbooks. For example, in the following chapter, I will use the example of cell wall expansion and its explanation; in Chapter 5, I will describe the explanation of the antibody synthesis, and so on. I think that these explanations have something to say about the explanatory role of chance, namely that they already give an explanatory value to chance that philosophers have to spell out. In normative terms, my philosophical contribution in these cases consists mainly of evaluative claims about the fact that these explanations are good explanations *because* chance provides some kind of explanatory power (cf. see Chapter 1, section 3.1.1). These examples correspond to type-level explanations since they do not refer to a specific instance of a phenomenon, but generalized phenomena of the same kind. For instance, the explanation of DNA replication (see Chapter 3) does not refer to a specific instance of DNA replication but to *all* DNA replication phenomena. The advantage of using type-level examples is that they make my argument more general and more compelling. First, it permits me to assert that the value of chance that I am defending is general, as well as the explanations I use to test my idea. Second, underlying well-known and general explanations – such as DNA replication in which chance plays a central explanatory role (see Chapter 3, section 3.3) – encourage me to ask whether other kinds of explanations, maybe less general and "famous", can or should imply a similar value too.
2. *Scientific article examples.* I use several examples from biological articles. In Chapter 5, I will rebuild numerous explanations that stem from the biological literature on noise in gene expression. In Chapter 6, I will pick up explanations from papers that deal with alternative splicing and in Chapter 7 that deal with alternative start-codon selection. In normative terms, my philosophical work in these cases consists mainly of two types of prescriptive claims (cf. see Chapter 1): 1) I argue that biologists should recognize that chance could have a role in certain biological explanations; 2) I argue that biologists should sometimes favour stochastic explanations (for the definition of this kind of explanation, see the end of the present section) over other explanations. Are these kinds of explanations of type or of token? Explanations from biological papers are difficult to separate into the philosophical categories of type or token level explanations – not least because these categories are relative, not absolute. At first glance, it may be tempting to

answer that papers also provide type-explanations. Scientists work with populations of cells, collect data about certain behaviors, and provide models, explanations and predictions[102] to understand these behaviours and fit to these data. On the one hand, this could mean that they provide models and/or explanations that apply to *all* situations of the same kind. For example, given a population of cells T that present a certain behaviour B, in a certain fixed environmental condition C, scientists typically provide an explanation E and/or a model M that can be applied for all Ts that show B. But the problem is that very often, the authors of papers give/provide E and/or M for *very specific biological scenarios*. Given a *specific* population of cells T that presents a *quite specific* behaviour B, in a certain fixed environmental condition C, authors of scientific papers typically provide an explanation E and/or a model M that can be applied for all *the specific Ts* that show *the specific Bs* in Cs. Do these explanations *still* count as type-level explanations? I think that if compared to the explanation provided by textbooks that are often more general, they are token level explanations, but they can be type level when compared to a more fine-grained explanation. Let's take the example of DNA replication. Its explanation in prokaryotes is type level with respect to DNA replication in E. Coli but token level with respect to DNA replication in general.[103] In light of this relative state, I suggest simply underlining this problem and keeping the answer open, since for my purposes it is not necessary to grapple with it here.

[102] However, it must be kept in mind that explanation and prediction do not always have to go together (see, e.g. Douglas 2009; Hanson 1959, Lee 1979; Rescher 1958; Scheffler 1957; Thom 1992).
[103] Thanks to Philippe Huneman for making this point clearer.

My "examples landscape" can be summarized in the Table 1 below:

	Meta-normativity	*Cases of study*	*Kind of explanations*
Textbooks	Evaluative claims "These explanations are good explanations *because* chance provides some kind of explanatory power"	- Antibody synthesis; - Cell wall expansion	Type-level explanations
Scientific articles	Evaluative claims: "These explanations are good explanations *because* chance provides some kind of explanatory power" Prescriptive claims: "Biologists should recognize that chance could have a role; biologists should sometimes favour stochastic explanation over other kinds of explanation"	- Noise in gene expression; - Alternative splicing; - Cell identity; - Alternative start codon selection in translation	Open question

Table 1 Our example landscape. For further details, see main text.

It is important to specify the biological sources of my analysis for the following reasons. 1) In a bottom-up project in philosophy of science in practice (PSP), biological data plays a principal role in the development of the philosophical argument. Specifying where these data come from renders the argument subject to fallibility – readers can check for themselves if my reconstruction is good, if I use relevant examples, if I understand the purpose of a certain explanation, and so on. 2) As already mentioned, knowing the source of examples opens the door to understanding how general an explanation is and, in turn, to what extent my account can be generalized. If I use an example from a textbook, the level of generality in that case is very high and my account will be more generalizable too, so more applicable. For instance, if I argue that chance has a main role in DNA replication, and if this argument is convincing, then my argument inherits the generality of the explanation considered. 3) I linked the sources of my examples with two kinds of normativity. This link helps as a further clarification of my normativity analysis, which I started in Chapter 1. In this section, I specified the kinds of normativity which are associated with specific kinds of examples. With textbook examples I mainly make evaluative claims; with examples from articles, I make both evaluative and prescriptive claims.

What I am looking for in this examples landscape are explanations that use chance in an explanatory way or/and explanations that might use or make chance explicit to improve their

goodness. But what is the value of chance? First, let's introduce my account of stochastic explanation (SE):

S.E. A stochastic explanation is a philosophical category that refers to any explanation with a chancy element that provides (explicitly or implicitly) (part of) its explanatory power.

This definition immediately brings up many questions. What does "explanatory power" mean in the context of SE? What does "chancy element" mean in the context of that explanation? And more to the point, how can we prove that in an explanation, a hypothetical chancy element C (whatever that means) is responsible for (most of) its explanatory power (whatever that means)? In order to answer these questions, I propose three criteria of adequacy that can help to delineate in more detail my philosophical idea of SE.

The first two criteria of adequacy that I cover in the last part of this chapter answer to *what SE is not*. Indeed, I want to provide an account of explanation for phenomena described as stochastic that: 1) does not rely on the problematic notion of mechanism or mechanistic explanation (section 4.2); 2) does not require MDB in order to be seen as a good explanation (section 4.3).These first two could be called "negative criteria" because they do not provide a precise account of SE but rather designate what must not be associated with SE. The third criteria of adequacy specifies what "chancy element" and "explanatory power" mean in the context of SE, and in general, what my philosophical idea of SE corresponds to. I call this last a "positive" criterium because it provides precise features of SE. In light of its importance, I dedicate the entire following chapter to this third criterium.

4.2 The first negative criterium. "SE is not a mechanistic explanation"

The general idea behind this negative criterium is that the notion of mechanism is problematic and its relation to explanation is not clear, as I showed in section 1 of this chapter. I do not want to ground SE in a philosophical notion that is problematic *right from the start*. More specifically, when I say that SE is not a mechanistic explanation I mean that:

- *The intelligibility of SE does not depend on some sort of productive continuity.* Often a mechanism is conceived as intelligible only if it is ensured by some sort of productive continuity such as "I have A, then B then C, etc...".[104] I refuse this claim. Leaving aside

[104] "The regularity is exhibited in the typical way that the mechanism runs from beginning to end; what makes it regular is the *productive continuity* between stages. Complete descriptions of mechanisms exhibit productive

the term "productive" as it could be very ambiguous,[105] my main objection is that what the authors call "continuity" is actually a very strictly level-loaded concept. What might appear a discontinuity at one level could be seen as a continuity to another. Cell differentiation (see e.g. Golubev 2012; Paldi 2018, Kupiec 2019; see Chapter 4, section 6) could be seen as a discontinuous process at the singular cellular level, but continuous at the level of cell populations or the tissues level. Numerous cells die during differentiation but the development of tissues appears continuous. When Machamer, Darden and Craver (2000) introduced the notion of "productive continuity", they had in mind the example of neurotransmitter release. They imagined the steps of that process as the phases A → B → C (and so on) and postulated that, for example, if B is unknown, then the phenomenon is hardly intelligible. But biological phenomena are well understood even if they are not associated with any idea of continuity (see, e.g. Matthiessen 2017) since other kinds of networks are intelligible too. Gene expression follows stochastic patterns (e.g. RNAs or/and polypeptides burst; see Chapters 5 and 6) and we cannot represent and explain this phenomenon through the notion of productive continuity because it is, rather, a complex matrix of processes that interact and influence each other. But the fact that the processes in gene expression do not show productive continuities does not preclude their intelligibility. There is deep understanding of these processes. Therefore, I want to leave the notion of productive continuity behind and provide an idea of intelligibility (and of understating, cf. section 4) that is more open.

- *SE does not explain biological phenomena that appeal to a notion of modularity.* Often the notion of mechanism is related to scientists trying to decompose and localize parts of a biological system in order to understand how it works (Bechtel and Richardson 1993; 2010; Kaiser 2018).[106] Decomposition and localization often imply the assumption that biological systems are modular, namely that each part of a system could

continuity without gaps from the set up to termination conditions. Productive continuities are what make the connections between stages intelligible" (Machamer *et al* 2000, p. 3, emphasis in original). For the authors, productive continuity makes a mechanism regular. But we can have a mechanism that shows productive continuity in the form of A→ B → C but that but that takes places only once. It seems here that productive continuity and regularity are decoupled, or at least, their relation has a degree of ambiguity difficult to solve.

[105] Kaiser and Krickel (2017) elaborate a critique of the notion of activity proposed by MDC in two main points. 1) "[A]ctivities are conceived of as irreducible and fundamental things that involve an unanalyzable kind of 'activeness' and 'productiveness'" (p. 754) and they point out that this raises philosophical issues difficult to handle; 2) "the category of activities is unnecessarily narrow. It excludes, for example, states of objects such as the voltage-gated sodium channel being open. Furthermore, it excludes passive behaviours, such as the release of calcium ions from the endoplasmic reticulum into the cytosol, from being components of mechanisms" (*ibid.*).

[106] "Localization is a spatial notion— where one identifies parts of mechanisms via their locations—while decomposition" (Glennan 2005, p. 447).

work without the other parts (Woodward 2013). This idea was strongly contested by philosophers. Rice (2015) criticizes Woodward's interventionist account of causal explanation, writing that: "[i]nterventionist accounts of causal explanation require that causes be *modular* in the sense that they can be manipulated independently of other causes within the system" (p. 20; emphasis in original). He continues arguing that this kind of modularity is very hard to find in evolving biological systems that are, by contrast, very "complex systems whose dynamics are chaotic, nonlinear, and involve feedback loops" (p. 21).[107] Cartwright writes that in biological systems it is often the case that "the causal laws are harnessed together and cannot be changed singly" (Rice 2015, p. 21, refers to Cartwright 2004, p. 811; see also Illari and Williamson 2012, p. 122). In light of these convincing critiques, I want to avoid using the notion of modularity in my account of SE (cf. section 1).

- *SE is not something in the world.* Unlike the majority of the new mechanist philosophers, I do not embrace an ontic account of explanation (even in a weak sense) because committing SE with an ontic account of explanation (whether weak or strong) could mean inheriting all the (metaphysical) problems of both the ontic account (see section 1) and the ontological status of chance in biological systems (see Chapter 5 and Chapter 6 concerning reductionism). As a philosophical category, stochastic explanation is thought to enable appreciation of the role of chance in certain biological explanations. To say that explanation X is a stochastic explanation is to make explicit the presence of a chancy explanatory element within. The level at which this analysis is provided is genuinely epistemological: I reflect on the ways the research is pursued, focusing and analyzing explanations. SE does not refer to the phenomena themselves since it actually refers to the *explanations* that biologists provide to understand the world.

[107] Illari and Williamson (2012) go further on this point arguing that Woodward (2002) is talking about *representations* of mechanism and not about mechanisms themselves. If it is hard to argue that mechanisms are modular, it is less problematic to say that models or representation are, or should be, modular. The authors say that modularity in representations and models make intervention and prediction easier (Illari and Williamson 2012, p. 122).

4.3 The second negative criterium. "SE is not committed to MDB"

Following the conclusion developed in section 2, I can say that "more details are better" (MDB) could be associated with mechanistic explanations but that the former is not a necessary condition of the latter.[108] I want to use this conclusion as a basis for this second negative criterium of adequacy. Why am I paying so much attention to this aspect (cf. section 2)? Because the relation between MDB and mechanistic explanation affects the way in which I can develop the present negative criterium. It would have been easier, were we to have found a strong commitment to MDB in the new mechanistic literature. That argumentation could have looked something like this: "I am developing an account *against* the mechanistic one that has a strong commitment with MDB". In a certain way, this would have been an easier strategy, to the extent that I would have clearly had in mind the target of my critique – mechanistic explanation committed to MDB. But the new mechanistic philosophy does not actually have this strong commitment, and I have to refine my argument in light of this evidence.

My aim is to argue in favor of an "alternative" and "not mechanistic" form of explanation (i.e. stochastic explanation (SE)) in the specific case of biological explanations of chancy phenomena. What I need is not definitive proof about the fact that *all* mechanist explanations are framed within MDB, but rather a proof that mechanistic explanations *might be framed* in terms of MDB. In other words, I can argue against MDB even if not all mechanistic explanations are framed in terms of MDB. Therefore, when I say that SE is not committed to MDB I mean that:

- *MDB is not a good epistemic norm for the explanatory power of SE.* What makes SE a good explanation is not the fact that SE provides a good degree of detail but the fact that SE contains a chancy explanatory element;
- *MDB is incompatible with my idea of chance as an explanatory element.* In Chapter 3, I provide a specific idea of what I mean when I write that chance can be explanatory. For now, I can say that chance has explanatory power to the extent that it provides abstract explanations. MDB and abstract explanation are clearly incompatible: the former adds detail while the latter removes them.

[108] All in all, Krickel notes that despite a certain tendency for MDB, in actual science, biologists often *do not even settle the problem*: "[s]cientists take the neurotransmitter release mechanism to explain neurotransmitter release even though they do not know what explains its failures. They seem to provide mechanistic explanations of phenomena *regardless* of whether they think that there could be a more detailed description of the mechanism that would render the relation between the mechanism and the phenomenon deterministic (although it might be an ideal that drives scientific research)" (Krickel 2018, p. 1138; emphasis added).

I have specified the two negative criteria of adequacy for SEs, but to provide the positive criterium, I have to take a longer path. The idea behind this criterium is to argue that the chancy elements present in SEs provide an important explanatory power in term of abstraction. Thus, I have first of all to check the main account of abstract explanations already developed in the literature of the philosophy of science. Some of these elements will come in handy when I develop and detail the norm of abstraction for my philosophical account of SE.

Conclusion

In this chapter, I provided a detailed analysis of the problems and limits of the notion of mechanism and mechanistic explanation. My aim is to convince the reader that it is time to go beyond mechanistic explanation and think about alternative forms of explanation when we reflect on stochastic phenomena. My first point was to make explicit that "more details are better" – if not outright rejected – is only an assumption/implication/suggestion in mechanistic literature. I then criticized the account of "high-failure mechanisms", asserting that it does not properly represent the different ways in which a phenomenon can be labelled stochastic in biology, and that high-failure mechanisms are a subset of the possible ways in which stochasticity can be found in biological explanation. I then introduced my original account of stochastic explanation (SE) with the aim of broadening the meaning of chance and thinking about a more comprehensive theoretical framework from which all kinds of stochastic phenomena can be properly represented, and in which the role of chance is made explicit and clear. I then described the first two negative criteria of SE, which specify what SE is not: 1) SE is not a mechanistic explanation; 2) SE is not committed to MDB. In the next chapter, I will introduce the third positive criterium that makes explicit what SE does refer to. This criterium corresponds to the epistemic notion of abstraction.

Chapter 3: Chance as an abstractor in life science explanations

Introduction

In this chapter, I develop in detail the third criterium of adequacy for my account of stochastic explanation.

Chance has an explanatory power as an abstractor in the context of explanation. It allows for the synthesis of the *explanans* as well as for unnecessary details to be abstracted in order to account for the *explanandum*. Moreover, an explanation in terms of chance as an abstractor (which I call "stochastic explanation" or SE)[109] should, in some explanatory contexts, be preferred to others which are non-chancy and more detailed. Note that my argument for the explanatory power of chance (as an abstractor) is a genuinely philosophical argument that stems from deep analysis and study of biological explanations.[110] I do not rule out the possibility that the argument could be used to highlight the explanatory role of chance in a plurality of different scientific domains – indeed, I hope in future works to be able to generalize this argument across different scientific disciplines.

The chapter is structured as follows. The first section introduces various accounts[111] of abstract explanation already present in existing literature.[112] In the second section, I introduce the idea of chance as an abstractor, my account of stochastic explanation (SE), and make explicit the extent to which chance provides to stochastic explanation (at least part of) its explanatory power. In the third section, I test my account with three biological examples, suggesting that the epistemic value of chance is already operational in certain biological explanations and the contribution made by my philosophical work simply makes this explicit.

[109] In Chapter 2 (section 4.1) I defined a stochastic explanation as follows: a stochastic explanation is a philosophical category that refers to any explanation with a chancy element that provides (explicitly or implicitly) (part of) its explanatory power.

[110] A reader might ask why chance as an abstractor finds application in *biological* phenomena? Here we must be careful because we could slip into metaphysical analyses such as wondering whether these phenomena, when biological, have a special status. What I propose is an inverse question: can we find non-biological explanations, (e.g. physical) in which chance as an abstractor has an essential role? My impression is that it is difficult to find them, but I sow the seed for ontological reflections that I will discuss later on in this chapter.

[111] The list of authors I will provide is not intended to be exhaustive. I focus on the accounts that are most useful for developing my argument.

[112] As I already stressed in the previous chapter (cf. the "second cluster"; section 2), there is also a literature on mechanisms that underlines the epistemic role of abstraction. Nonetheless, to the extent that I argued that mechanistic explanation is problematic, I do not want to ground my discussion of abstraction on this kind of explanation.

1. Abstract explanations

1.1 How can we unpack the notion of explanatory power?

There is a rich literature on explanatory power (or "explanatory depth") in scientific explanations (e.g. Jackson and Pettit 1992; Strevens 2008; Woodward 2005; Ylikoski and Kuorikoski 2010; Ylikoski 2001, 2007). What gives an explanation value? To what extent can we claim that one explanation is better than another? Numerous authors focus on the notion of explanatory power in trying to answer these questions. Often, authors provide different conditions – traditionally called "epistemic norms" or "dimensions of goodness" – to specify the value of an explanation, which is to say its explanatory power.[113] For example, Ylikoski and Kuorikoski (2010) provide a list of parameters that highlight what explanatory power can refer to. These parameters are: sensitivity, cognitive salience, precision, detail, factual accuracy and degree of integration. They note that these conditions are not necessarily consistent with each other. For example, if an agent wants an explanation with high precision, it may as a result be less cognitively salient (cf. Levins 1966). Depending on her epistemic interests and goals, each agent can have a different list of these conditions in order to establish whether an explanation is a good one. For example, Baedke (2018; Chapter 5) proposes that precision, sensitivity and cognitive salience are good parameters by which to establish the explanatory power of scientific explanations and that likeness, causal power, attractiveness and riskiness are not (Baedke 2018, p. 171).

The choice of these conditions, then, is a pragmatic task insofar as it depends on the attitude of the agent.[114] Some authors determine that any kind of choice regarding epistemic parameters made at this point is a product of subjective decisions. Sober (1999), following Railton (1980, 1981), writes that there is a complete explanation towards which we must aim. If the complete explanation is not (yet) available, the choice of partial explanation that we will use is simply a "matter of taste" (Sober 1999, p. 551) and depends on subjective factors, such

[113] Even if I focus on explanatory power, other criteria have been proposed to evaluate explanations. Hempel and Oppenheim (1948) focus on the law of nature, Salmon (1971) on statistical significance. More recently, Woodward (2003) argues that we can evaluate an explanation in relation to the number of what-if-it-had-been-different questions the agent can answer. Lewis (1986), and more recently Strevens (2008), focus of the importance of causal factors and high probability (Strevens 2000).

[114] In this scenario there are two possible interpretations of the fact that the conditions for a good explanation depend on "the attitude of the agent". The first possible attitude is to consider that the choice for certain conditions instead of others depends on what I want to explain. For example, Baedke chooses precision, sensitivity, and cognitive salience because his goal is to provide epigenetic explanations. In this case, it would seem that the conditions for a good explanation are pragmatically established depending on the object of the explanation. I would like to resist this possibility by opting instead for a second attitude. The conditions that would make for a good explanation depend on an agent's beliefs as to whether certain conditions are better than others with respect to the epistemic advantages they may have. From this perspective, Baedke would have opted for precision, sensitivity and cognitive salience, thinking of them as good conditions for obtaining, in general, good explanations.

as the agent's interest, inclinations, etc.[115] Then, Sober does not believe that the context objectively matters in developing a theory of explanation. By contrast, van Fraassen (1977,1980) develops an account of explanation showing that interests, inclinations, etc., are not unconstrained and subjective but "'pragmatic virtues', having to do with the 'application' of science" (*Scientific Explanation*, Stanford Encyclopedia, section 6.2). Van Fraasen emphasizes the relevance of the context in which the why-question is developed. Following him, whether I develop one explanation over another does not depend on my own internal subjective inclinations, but rather on my research context.[116] Aspects – such as choosing the subject of the why-question, choosing only certain contrastive classes, and establishing the types of question that can be accepted by the questioner – are all pragmatic choices since they are developed to maximize the quality of the explanation to be developed for/in a certain research context[117].

In the same vein, I am pragmatic (in van Fraassen's sense) since I choose the epistemic norms (in this case only one) that better allow me to maximize the quality of my account of stochastic explanation. I can pose this question in the following way: What condition should I choose (pragmatic choice) in order to best underscore the explanatory power of chance? I answer, abstraction. But before entering the merits of my proposal, it is necessary to draw the reader's attention to philosophical studies that have already worked with this notion of abstraction.

1.2 Abstraction in explanation

Reflecting on abstraction as an epistemic norm is nothing new. Numerous papers have already discussed the explanatory power of abstraction: 1) the literature on alternative forms of explanations in which abstraction is often compared with the notion of idealization (Godfrey-Smith 2009; Jones 2005, 2013; Levy 2018), generalization (Green and Jones 2016; Putnam 1973, 1975; Rice 2015; Sober 1983; Weslake 2010)[118]; 2) and the literature on the statistical vs causal explanation debate (Ariew *et al* 2017; Lange 2013, Matthen 2009). Let's use this section to briefly analyze the first group. In the section that follows, I will deal with the second group.

[115] Within this debate, Rosenberg – a defender of ontological reductionism in biology – claims that biological explanations are "erotetic" or "protagorean" (Rosenberg 2006, p. 366).
[116] In a similar vein, the choice of the level of explanation is actually context dependent. In particular, it depends on the "investigator's purpose" (Woodward 2010, p. 297), i.e., what "the investigator wished to explain or understand" (see also Godfrey-Smith's contextualism 2003).
[117] For a critique to van Fraassen see Kitcher and Salmon (1987).
[118] Furthermore, Levy underlines that other authors write about abstraction but in an indirect way: they write about it to address other issues (Levy 2018, p. 2).

1.2.1 The confusion between abstraction and generality

With respect to the first group of studies proposed above, I will focus on authors working on abstraction and generality rather than on abstraction and idealization, since this second relation is generally less problematic than the first. Indeed, they all more or less agree that *abstracting* consists of removing a certain degree of detail while *idealizing* – meaning distorting or introducing some (often) clearly false elements – to make the theory/explanation more effective. In the present section, I focus solely on the works of Sober and Weslake. They are the most problematic, and therefore most philosophically interesting, with respect to the relationship between abstraction and generality.

Sober's account of equilibrium explanation does not refer to abstraction but generality. However, as we will see below, it can also be understood as an account of abstract explanation. Sober (1983) proposes an account of equilibrium explanation referring to generality which challenges the utility of causal explanation (see also Weslake 2010). "[While] causal explanation shows [that] the event to be explained was in fact produced, equilibrium explanation shows [by contrast] how the event would have occurred *regardless of which of a variety of causal scenarios actually transpired*" (Sober 1983, p. 202; emphasis added). Equilibrium explanation is an account that permits the explanation of an *explanandum*, even when different *explanantia* are possible, the full details of which we do not consider. Take the example of a ball rolling down a bowl. Regardless of its specific trajectory (i.e. all possible *explanantia*), the ball will always stop its journey at the center of the bowl, at a situation of equilibrium (the unique *explanandum*; see Sober 1983; Strevens 2019). Even without considering all the possible ways in which this equilibrium is reached, it can be explained by recalling the physical laws the ball follows.[119] The generality of the equilibrium explanation resides in the *explanans* being general enough to cover very different possibilities. In other words, Sober makes explicit that generality refers to a specific element within a single explanation, in this case the *explanans*. But Sober does not write about abstraction. By contrast, this specification could be useful to describe his account more precisely. In Sober's equilibrium explanation, do we have a process of abstraction at the level of the *explanans*? If so, of what kind? I speculate that in the *explanans* of Sober's equilibrium explanation we have a process of abstraction in the sense of abstracting away from all the possible ways in which the *explanandum* can be realized. No matter which pattern the ball follows, it will always arrive at

[119] I do not discuss whether this example falls within the DN account of explanation or not. I use it only to give an example of how we can think about generality in an equilibrium explanation (see Sober 1983).

the same equilibrium position, which is the center of the bowl. This position is explainable regardless of (that is to say, abstracted from) the specific path followed by the ball.

Generality and abstraction are not equal notions. Levy (2018) writes that while generality concerns scope, "the number of things (objects, processes, phenomena) to which a representation applies" (p. 15), abstraction concerns details, "not which things a representation covers, but how much it says about them" (*ibid.*). In Sober's account, the notion of generality refers to the number of situations in which the equilibrium explanation can apply. This explanation has a high degree of generality because it can be applied to all situations in which a ball is dropped into a bowl. By contrast, abstraction refers to the degree of detail in the *explanans*, that in this case is very low, in order to cover all the possible ways in which the *explanandum* could be realized. However, I stop here, since it is not my job to extrapolate from Sober's account a satisfactory notion of abstraction (nor is it my desire).

Another author who links abstraction to explanation is Weslake (2010). He argues in favor of the independence of nonfundamental science (e.g. biology) from fundamental science (i.e. physics). Weslake's position could be framed in the general debate on (anti)reductionism in biology initiated by Kitcher (1984) and Fodor (1976) (see Chapter 6). Very briefly, Weslake (2010) writes that "[…] the nonfundamental explanation *supervenes on* and *is multiply realizable by* the fundamental explanation" (p. 288; emphasis added) if the relation between fundamental and nonfundamental explanations in different situations is one of the following: 1) both explanations apply; 2) nonfundamental explanations apply and fundamental explanations do not apply (*ibid.*). Weslake calls "the degree to which a whole explanation applies to a range of possible situations *abstraction*" (p. 284; emphasis in original).[120] More explicitly, the epistemic notion of abstraction consists of the supervenience and multiple realizability of nonfundamental explanation as compared to fundamental explanation. I appreciate this strategy for giving an *explicit* meaning to abstraction, because it is also what I want to do in the context of stochastic explanation (SE) (see section 2). Nonetheless, since Weslake does not give a precise definition of supervenience and multiple realizability, and since these notions are hotly debated in philosophy (Polger and Shapiro 2016), the notion of abstraction remains vague.

Furthermore, even if Weslake made strides in defining abstraction, there still exists a confusion between generality and abstraction, or at least a degree of ambiguity. When he gives the definition of abstraction, it seems that this definition can work for generality too:

[120] Weslake explicitly echoes Sober and Putnam (Weslake 2010, pp. 290-291).

"[c]all the degree to which a whole explanation applies to a range of possible situations *abstraction*" (Weslake 2010, p. 284; emphasis in original).

This could equally be rephrased as, "call the degree to which a whole explanation applies to a range of possible situations *generality*." But, in line with Levy, I write that it is generality that concerns the extensiveness of an explanation, not abstraction. Abstraction concerns the detail of the explanation. It seems that Weslake calls abstraction what is actually generality.

1.2.2 Abstraction in statistical explanations

This section aims to focus on a corpus of philosophical studies on statistical explanation, within the context of the debate on statistical vs. causal explanation in the history and philosophy of biology. The reason for this focus is that these kinds of studies gesture explicitly or implicitly to a central role in the notion of abstraction. What I want to show is that even when dealing with statistical explanation, none of these accounts treat abstraction in the way I propose in the following section. This reinforces the idea that the account of abstraction related to chance that I propose is original.

Lange (2013) proposes that in the study of population genetics, we can find a hitherto neglected non-causal explanation that he calls "really statistical (RS) explanation". The explanatory power of RS resides in some "particular signature of statistical processes that the explanandum exemplifies" (Lange 2013, p. 173). If we ask how to explain and predict the distribution of weight in a given population, Lange would answer that we can explain it using the statistical tools of "regression toward a mean" (or "normal distribution"). Lange calls this account "real" statistical explanation since 1) the explanatory power stems entirely from statistical signatures that the *explananda* exemplifies, and 2) no causal details are required for this kind of explanation.[121]

Lange does not explicitly talk of abstraction. Nonetheless, the strength of his account resides in the abstract properties of statistical signatures.[122] Population phenomena (i.e. natural phenomena) are explained only by using statistical tools, abstracting away from causal details.

[121] He also adds that if we were to add causal details to the explanation, we will not have a less abstract explanation but instead a completely different one. Looking for causal details slides the *explanandum* from asking about the behavior of a population to asking about the behavior of a single individual by looking to its single causal history (e.g. reproductive activities, geographical movements, proximity to prey and or predators, etc.) (cf. Chapter 5, section 3.1.1).

[122] Lange can be considered ~~as~~ a "statisticalist". Typically, the aim of statisticalists is to show "that evolutionary theory can be understood as purely statistical phenomenon, thereby casting the aim of population genetics as mostly about relationships among abstract statistical properties" (Holmes 2021, p. 1).

Here, abstraction is intended as the property of statistical models that allows for the explanation of natural phenomena lacking detail, such as the causal history of individuals. In RS explanation, abstraction is found within the statistical model that provides explanation.

An account similar to Lange's but somewhat more radical is that of Ariew *et al* (2015, 2017), who propose talking of statistical autonomous explanation (SAE). In this account, not only are the causes abstracted, all references to natural entities are too. Ariew *et al* postulate that if a certain system (e.g. natural populations) satisfies statistical requirements, nothing else is needed to explain its dynamics. Very generally, if a real scenario (such as a natural population) meets certain statistical assumptions (which they call the "minimal material condition" – see after), then the explanation that can be given of this natural population is genuinely statistical; no causal detail is necessary. The development of an SAE proceeds in two stages. 1) The first step is to assume that the population conforms to the proprieties of an idealized statistical distribution (Ariew *et al* 2015, p. 13). It does this by conforming to three minimal material conditions: the population must feature characters that can be sampled randomly; the trials by which they are sampled must be independent; the trials must be sufficiently numerous (Holmes 2021, p. 3; see Ariew *et al* 2017). 2) The second step consists of deducing[123] the *explanandum* from statistical regularities that "govern" statistical distribution (Ariew *et al* 2015, p. 13). Ariew *et al* give a historical example: Galton's question of why inheritance preserves a normal distribution of certain characteristics (e.g. height) over generations. Ariew *et al* answer that it is because the population can be idealized with minimal material conditions and these conditions enable a real phenomenon, such as the inheritance of a certain characteristic, to be explained by statistical law.[124]

As is also the case with Lange's account, SAE does not make explicit a precise concept of abstraction. Nonetheless, we can extrapolate its importance. One of the main points of this "autonomous" statistical explanation concerns autonomy from the real details out there (whenever causal or not). If the minimal material conditions are satisfied, we are able to explain a complex natural situation, such as natural population dynamics, by using a simple statistical tool, such as a diagram of the normal distribution. Here, abstraction can be seen as playing a role in specifying why this account can be seen as "autonomous". SAE is autonomous since it

[123] Note that in response to the objection that deduction from statistical law is not enough for an explanation, the authors add: "we argue that autonomous statistical explanations are not merely deductions, but are sufficient explanations when they are also able to provide counterfactual information that reveals the salient relationships of dependence" (Ariew *et al* 2015, p. 23).

[124] In a personal communication, André Ariew told me that there is a deeper reason why real phenomena respect the minimum material conditions, but for the purpose of the present section is not necessary go into detail concerning the development of their account.

abstracts away real detail in order to explain natural phenomena. Focusing only on normal distribution, the explanation is abstracted away from all the real situations that a natural population can present. These include geographic location, non-random reproduction, prey-predator proximity, contingent dynamics, etc. In a word, both RS explanation and SAE abstract away all causal histories of individuals since they are unnecessary for explaining population macro-proprieties (cf. Garfinkel 1981). In both Lange's RS explanation and Ariew *et al*'s SAE, we find an explanation whose explanatory power stems from the fact that relatively simple statistical models can explain complex biological phenomena.[125]

What is the difference between the kind of abstraction that Lange and Ariew *et al* (implicitly) propose and my idea of chance as an abstractor? Very generally, I can say that there are various ways in which an explanation can abstract away detail. Their respective accounts of statistical explanation exemplify two possible ways of doing that. In statistical explanations, both RS and SAE, statistical signatures and/or models allow the causal detail of a phenomenon to be abstracted away. Here, the explanatory power lies in the statistical signatures/models. In my case, abstraction resides in a property of chance itself *which, as such,* allows for the different ways in which the *explanandum* can be realized to be abstracted. In a word, in their accounts, the explanatory power resides in statistical models and signatures; in my account, it resides in the notion of chance itself.

1.2.2.1 Causal vs non-casual explanation, the agnostic solution

The big difference between RS explanation and SAE is that Lange aligns with the need to be clear about the causal/non-causal status of explanation under question. RS is a non-causal explanation since it abstracts away causal details and explains only by using statistical tools. By contrast, were we to ask Ariew *et al* whether SAE is a causal explanation or not, they would answer that this question is actually misplaced and "has produced an intellectually unsatisfactory literature" (Ariew e*t al* 2017, p. 63). In their view, the relevant question is not whether statistical explanation is causal or non-causal. Rather, in order to shift attention from complicated metaphysical questions and towards reflection on history of science, the question

[125] Actually, there is another account of statistical explanation that does explicitly discuss abstaction. Matthen's (2009) "Statistically Abstractive Explanation" aims to highilight a hitherto neglected form of explanation where factors are abstracted away even if they are probabilistically relevant. To explain a certain characteristic of the market, a model in economics abstracts away, for example, the fact that for certain women a color is more eye-catching than another. This is a factor that is probabilistically relevant but theoretically inadmissible. I do not analyse this account in the main text since Matthen uses abstraction in a quite different fashion compared to my idea, so I prefer to discuss other cases (Lange, Ariew *et al*) that, even if implicilty, imply a way to abstract away details in ways more similar to mine.

should instead be, "what is the historical genesis of an explanation?" In describing the case study of Galton's explanation, they show that SAE explains macro-regularities (e.g. through normal distribution) and that causal details are useless for this purpose (even though these same details can prove useful for a different kind of statistical explanation interested in causes). They conclude: "causal […] explanation and statistically autonomous explanation are distinct and epistemically irreducible forms of explanation" (Ariew *et al* 2015, p. 13). It is, then, clear that Ariew *et al* are deflationary (cf. Holmes 2021, p. 8) concerning cause and do not ask if either the things they are abstracting away or their SAE accounts are causal or not. The ambiguity of their notion of "material" in their minimal material conditions is voluntary: the entities that account for these minimal conditions (being random and independently sampled, and being present in a large number) correspond to individuals of a biological population – that is to say natural entities. They do not deny this. Nevertheless, they leave unanswered the question of whether the entities satisfying these minimal material conditions are in a causal relationship with each other or not. They simply do not answer. It is not their aim.

The two accounts sketched in these last two sections prompt me to formulate the following question: besides any technicalities, am I developing a causal or non-causal account of explanation? This question gains particular relevance to the extent that this choice makes an important difference in defining my concept of chance. If we talk about causal explanation and chance is also to be conceived as causal, is that to say that chance is a cause as, for example, Aristotle suggests (cf. Dudley 2012; Ross 1936)? By contrast, if I develop a non-causal account of explanation, what is chance then? An absence of cause as Hume proposed? From my side, I am sympathetic with Ariew *et al* in diverting attention from the question of causal/non-causal explanation.[126] Nonetheless, I do not align myself with Ariew *et al* who say that the whole debate around non-causal causation is completely "intellectually unsatisfactory". I do, however, unquestionably recognize that this is a topic that has already been covered and abundantly discussed in the literature. For this reason, it is not my intention either to go into too much detail or not to propose yet another account. Nonetheless, I want to be clear with respect to the kind of position I am taking concerning my project. My attitude, that I will to call "causal agnosticism,"[127] does not stem (only) from the fear of not being able to say something original

[126] Holmes (2021) provides a critique of SAE arguing that its status "is at best unclear and at worst doubtful" (Holmes 2021, p. 8) exactly because it is a deflationary account. For the purposes of the present section, which is only to mention some accounts relating to abstraction, I do not need to enter in these detail.

[127] Thanks to Charles Pence to coming up with this expression. The "causal agnosticism" argument goes against the idea proposed by Humphreys (1989) that chance, not being a cause, cannot be explanatory. I say that chance can be explanatory regardless of whether the explanation under consideration is causal or not.

about this causal vs. non-causal debate. In fact, I do not even need to take a position regarding the issue of causality in order to develop my account of stochastic explanation in biology. Or at least, taking one position or another would not jeopardize my general account. Indeed, the argument I will provide in the next section is entirely focused on the structure of explanation. Its strength does not stem from determining causal relations between processes (processes which I am nonetheless interested in explaining), but rather from an abstract relationship that chance enables between the possible *explanantia* and the *explanandum*, within a single explanation. Whether the examples that I provide to substantiate my account are explicitly or implicitly causal or not *does not trouble* my argument. The kind of abstraction that chance provides "takes place" between the *explanans* and the *explanandum* within a single explanation. I will argue that, at least for some explanations, when the *explanans* contains notions such as "stochasticity", "randomness" and "chance", these notions enable the *explanans* to abstract unnecessary details in order to account for the *explanandum* (cf. Sober 1980).

All the accounts that I sketched in this first section will contribute to the development of my idea of stochastic explanation. From Sober and Weslake I learned the necessity of carrying the notion of abstraction to its fullest extent while at the same time doing important conceptual work to separate it from the concept of generalization. From Sober I also learned that a good way to ground discussions on explanation is by using the *explanans-explanandum* schema. Lange's RS and Ariew *et al*'s SAE accounts have shown me some ways in which abstraction comes into play in statistical explanation. Even if statistical explanation deals with chance in terms of probability, the purpose of these accounts is not to discuss the notion of chance or abstraction, and at any rate they do not make abstraction a key point of the argument (although implications with respect to the latter are present, as shown in the previous sections). This is good news for me since it means that the account I am going to provide is original at least in the sense of linking chance with abstraction. Of course, originality does not mean relevance. This is why, in the rest of the chapter, I aim to convince the reader of the relevance of my account of stochastic explanation from both a philosophical and biological point of view. Moreover, Lange's and Ariew *et al*'s accounts have allowed me to make my position clearer regarding the causal vs. non-causal explanation debate and my stance on causal agnosticism. More generally, this survey of the literature has given me many insights into the sides present in different debates. In the next section, I bring order and consistency to these ideas by proposing the third criterion of adequacy that an explanation must have in order to be counted as a stochastic explanation (the other two are described in Chapter 2, sections 4.2 and 4.3).

2. The epistemic norm of abstraction. The third positive criterium

I argue for chance's explanatory power in terms of the epistemic norm of abstraction. First of all, I must make explicit what I mean by abstraction. As showed by Levy (2018), abstraction is a comparative notion. An explanation is abstract with respect to something else – another explanation, for example. More specifically, abstractness "is whether, given two representations, one includes more detail than another *vis-à-vis* the same subject matter" (p. 4). Levy also adds that these representations do not have to be *pre-specified* (p. 6):

"[…] I think that in making judgments of abstractness with regards to a given representation we typically only presuppose that there *could be* a representation with a different degree of detail. In other words, we often rightly speak of a representation as abstract even though we are not comparing it to some specific, more concrete alternative. In so doing we simply express the judgment that a more detailed description is possible, i.e. that more *could* be said" (p. 6; emphasis in original).

I agree with Levy that we do not need a clear idea of both representations to provide a comparison. We can have a representation (a) and an idea (even confused) of an alternative representation (b), and still be able to compare the two. I am sympathetic to this idea since we can look at a tree and assert that it is small by comparing it with our *idea* of "tree". My criticism of Levy is that, in the quotation mentioned above, he does not specify the context in which he talks about these representations. When Levy writes about representation, is he referring to an explanation? If this is not the case, are these representations just descriptions? What I am saying is that Levy does not clarify the theoretical framework from which he writes about "representations". By contrast, and following Sober, in what follows I want to talk about abstraction in the context of explanation, highlighting the relationship between the *explanans* and the *explanandum*. I intend to frame my analysis in the context of a single explanation.

What does it mean that the notion of abstraction is referred to as the *explanans* of an explanation, in particular stochastic explanation (SE)? What gives this notion its epistemic value? Levy (2018) distinguishes between the process of abstraction and the product this process gives rise to, namely poorly detailed representations. He calls "abstraction" the process that moves from a detail-rich to a detail-poor representation and "abstractness" the representation itself and its level of detail. Avoiding confusion, when I talk about Levy's concepts I add "LY" to his terms: "LY abstraction" and "LY abstractness". He writes that often LY abstraction is not so important, because sometimes it is not essential to specify how we arrive at abstracting away details (p. 4). He shows that the statement "the speed of light in a

vacuum is constant" is less detailed and more abstract than "the speed of light in a vacuum is c= 299 792 458 m/s." He adds though, "[t]his is irrespective of whether the former was arrived at by leading out information contained in the latter" (p. 4). Therefore, he focuses attention on "LY abstractness", the product of abstraction or, in other words, the less detailed representation itself.

Abstraction Box
LY abstraction = the process of abstraction that moves from a rich to a detail-poor representation
LY abstractness = the abstracted representation itself and its level of details

In my context of study, the way in which chance abstracts away details, that is to say LY abstraction, is not negligible since it is the exact element that could make explicit the sense by which chance provides abstraction in SE. I state that chance shows its LY abstractness thanks to the process of LY abstraction I propose to define more precisely below:

THE PROCESS OF LY ABSTRACTION (PLA). *To any given stochastic explanation (SE), the* explanans *can hold and synthesize disjunctive and/or cumulative representations of a variety of possible ways in which the* explanandum *(e.g. the phenomenon to be explained) can be realized, without the need to provide the details of all those representations.*

Here, the comparison becomes more specific than in Levy's account. I am comparing the *explanans* that contains a chance element with other possible and more detailed *explanantia.* We find the process of abstraction precisely in this comparison. *Explanans* with such chancy elements (that are in fact used in the explanation) are more abstract than any other more detailed *explanantia* that might explain the same *explanandum*. In addition, this kind of *explanans* does not only abstract away the more detailed *explanantia*, they even synthesize them.

But this definition of the process of LY abstraction (PLA) easily leads to a new question: how can the *explanans* "hold" and "synthesize" the disjunctive and/or cumulative representations of the different ways in which the *explanandum* could be realized without providing the detail of all those representations? In other words, how can the actual *explanans* abstract, and at the same time synthetize, all the other possible, more detailed *explanantia*? Sober might answer that for an equilibrium explanation it suffices that the *explanans* is general enough to cover all the different ways in which the *explanandum* can be realized. My analysis is inspired by Sober's account, and even were I do provide an account of a more specific kind of process, I want to focus mainly on abstraction, not on generalization. I argue that, in stochastic explanation (SE), the power of the *explanans* to abstract and to synthesize more

detailed *explanantia* derives from the notion of chance that is present within it.[128] But first of all, I have to elaborate my definition of chance that will be used to specify this purpose.

2.1 My notion of chance as present contingency (CPC)

At the beginning of this thesis, I introduced a first general definition of the notion of chance developed by reworking some elements from the debate on contingency (cf. Introduction, section 1.5).

Given some fixed initial conditions, chance is a property of a description of a process. This property specifies that the described process could be otherwise (or at least that some events in that process can be otherwise).

In the present context, it is necessary to make this definition more specific. The first part remains the same because it defines precisely what I mean by chance in my epistemological framework, i.e. *a property of a description of a process*. By contrast, the second part is ambiguous: in what sense can a process (or an event) *be otherwise*? It is fundamental to understand this point because my proposal is precisely that chance holds and synthesizes in a single *explanans* all the different ways in which a process (or event) can be otherwise (providing nonetheless the same *explanandum*). It is therefore necessary to a unpack the meaning of this *can be otherwise*. I propose then that a process (or event) can be otherwise insofar as it could take place, not take place or take place in different ways.

CHANCE AS PRESENT CONTINGENCY (CPC). *Given some fixed initial conditions, chance is a property of a description of a process. This property specifies that the described process (or at least some events in that process) could take place, not take place or take place in different ways.*

In this definition, when I say that a process, described as chancy, could take place, could not take place or could take place in different ways, I refer to the fact that a process P could (x) ("P takes place") or could (¬x) ("P does not take place") or could take place in a way (x) and/or in a way (y) and/or in a way (z), etc.

[128] Having said that, I maintain that once the notion of abstraction is established, the notion of generality follows. Indeed, if we have an *explanans* X that has abstracted and synthesized all the more detailed *explanantia* Y, X, Z etc. then the explanation is general insofar as it can be valid for both Y, X and Z.

In the present discussion, I develop CPC in terms of processes rather than results since my aim is to argue that representations of stochastic *processes*[129] can have explanatory power. But still, why do I focus on processes instead of results?[130] I have two reasons for saying that focusing on results is not a good research direction in this current context:

- It does not make sense to propose a reflection on results since biological explanations are typically in terms of *processes*. Biological textbooks provide copious evidence of the fact that biologists are interested in the *processes* that produce certain results. Since my project is a bottom-up and PSP project, I want to stay as close as possible to what biologists do and work with. In addition, the processes that I am interested in are the ones that biologists *already* label as stochastic. My question is not whether these processes are really stochastic, but rather how the involvement of chance in their descriptions/representations can have explanatory power;
- Glennan (1997) writes that a stochastic result can be perfectly obtained by both stochastic and deterministic processes (see also Earman 1986; Merlin 2009; Millstein 2000c). The result of a process does not tell us anything about the nature (deterministic or indeterministic) of that process. In this situation, why use results to show the epistemic value of chance if we are able to work directly with processes that biologists themselves already count as stochastic? This would be an explanatory regression since we would no longer be able to establish the status of processes knowing only their results.

These two arguments are enough to enable me to continue my analysis, focusing mainly on processes. Once the definition of CPC is specified, what remains is to tie together the process of LY abstraction with this notion of chance (CPC) in order to be clear about the sense in which chance as an abstractor can provide an explanatory power.

[129] A reader may have noticed that in working with the notion of stochastic process, I have not yet given a definition for a process. I have no intention of getting into difficult metaphysical discussion with respect to what a process is. I propose then a deflationary definition of a process. I define it as a series of events over time that can produce certain results.

[130] *A priori*, for a given explanation, we could imagine an argument that shows that results can have an explanatory power too. Think broadly about work in biological practice. Scientists do experiments in the laboratory collecting data, often in the form of results, that allow them to explain, or at first perhaps only interpret, a certain phenomenon.

2.2 Providing the synthesis between chance as present contingency (CPC) and the process of LY abstraction (PLA)

In this section, I elaborate a framework to specify in what sense chance as an abstractor in the context of (at least certain) biological explanations underlies the special relation of abstraction between *explanans* and *explanandum*. More technically, I elaborate a synthesis between chance as present contingency (CPC) and the process of LY abstraction (PLA). I consider three ways in which CPC can provide abstract *explanatia* in term of disjunctive and/or cumulative representations:

- In the first case, chance as present contingency (CPC) provides an *explanans* of a stochastic explanation (SE) though the process of LY abstraction (PLA). It does this by synthesizing the **disjunctive representations** of a variety of possible ways in which the *explanandum* can be realized. Given a fixed initial condition X, the behaviour of a biological process P could be (x) or not (x); or could be (x) or (y) or (z) or (t) and so on.

  Short example
  Explanandum: The fact that a single neuron expresses just one of 1,500 genes in order to develop an odorant receptor.
  Explanans: A transcription factor "TF" can stochastically (CPC) initiate the expression of a single, or a few, receptors per cell – that is to say, a transcription factor "TF" can stochastically activate gene (A) or (B) or (C) or… (X). These ways (A), (B), and (C) are not listed in the explanans but abstracted and synthetized in the verbalization "can stochastically activate" (PLA).

- In the second case, chance as present contingency (CPC) provides an *explanans* of a stochastic explanation (SE) though the process of LY abstraction (PLA) by synthesizing the **cumulative representations** of a variety of possible ways in which the *explanandum* can be realized. Given a fixed initial condition X, the behaviour of a biological process P could be (x) and (y) and (z) and (t), etc.

  Short example
  Explanandum: Plant cell walls grow spherical in shape.

Explanans: "If the cellulose microfibrils are randomly oriented [CPC] in all walls, the cell will expand equally in all directions, tending to become spherical in shape" (Evert and Eichhorn 2013, p. 61). Then microfibrils are randomly oriented through directions (A) and (B) and (C), etc., that is to say through all directions. These directions are not listed in the explanans but abstracted and synthetized in the verbalization "random rearrangement" (PLA).

- In the third case, chance as present contingency (CPC) provides an *explanans* of a stochastic explanation (SE) though the process of LY abstraction (PLA) by synthesizing the **disjunctive and cumulative representations** of a variety of possible ways in which the *explanandum* can be realized. Given a fixed initial condition X, the behaviour of a biological process P could be (x) and/or (y) and/or (z) and/or (t), etc.

  Short example
  Explanandum: The addition of nucleotides to the new strand during DNA replication.
  Explanans: A random sampling (CPC) at the beginning of the process of elongation of the new DNA strand. The sampling process can add nucleotide (A) and/or nucleotide (B) and/or nucleotide (C), etc. These ways (A), (B), and/or (C) are not listed in the explanans but abstracted and synthetized in the verbalization "random sampling" (PLA).

In summary, chance as present contingency (CPC) permits the process of LY abstraction (PLA) that, in turn, permits the abstract explanation that I call stochastic explanation (SE). In what follows, I provide the structure of a stochastic explanation. I depict only the third case since the other two share the same structure and differ only in "and/or" formulas:

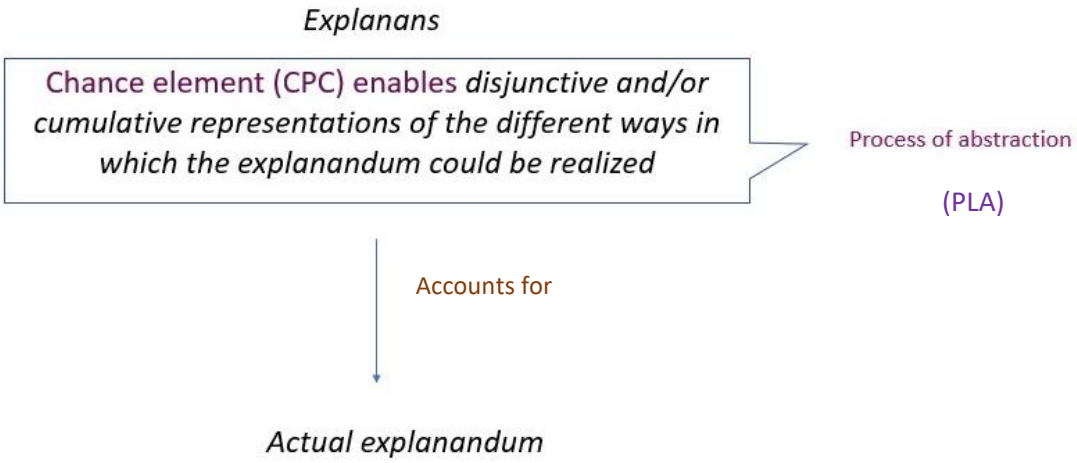


Figure 1 Structure of SE in terms of *explanans* and *explanandum* in which I depict CPC and PLA. For further details see main text.

In providing the definition of SE (see Chapter 2, section 4), I mentioned that its explanatory power could be due (in part) to a chancy part present in it. Relying on the process of LY abstraction (PLA) and chance as present contingency (CPC), I can now be more precise. Biological explanation does not, of course, refer to CPC (nor to PLA). CPC is a philosophical tool that allows me to identify chance concepts scattered in explanations. Often chance is identified by nouns such as "stochasticity", "randomness", "chance", or adjectives such as "stochastic", "randomly" or "chancy" etc. I use CPC to make explicit that chance has a central explanatory role, no matter the words. Once CPC is identified in a given explanation, I show that it is precisely its presence in the *explanans* (see Figure 1) that enables the explanation to be abstract and good. "Abstract" because CPC in the *explanans* holds and synthetizes the different ways (cumulative and/or disjunctive) in which the *explanandum* could be realized. "Good" since insofar as this kind of explanation (even if not the proposed philosophical extrapolation) can be found in textbooks and scientific articles (see Chapter 2, section 4.1), then it is reasonable to think that they are valid and shared by the scientific community. I do assume that biologists agree that these explanations are already "good", since explanations that are neither validated nor shared by scientists should have no place in biological textbooks, or more generally in scientific publications. I am not saying anything new here. I want only to underline in a deeper sense why these explanations are as good as they are. Because chance is present in the *explanans*, this allows all the possible ways in which the *explanandum* can be realized to be synthesized and held.

2.3 An orthogonal question and a transversal one for my account of stochastic explanation

The account so far sketched could raise two orders of questions with respect to the epistemological perspective on chance adopted in the present work – one orthogonal and one transversal. Both deal with a basic interrogation of where the explanatory power of stochastic explanation stems from. But concerning this origin, these two questions suggest different possibilities. Let's start with the transversal one. Is the explanatory power of SE due to the general fact that this explanation is abstract (that is to say that unnecessary details have been abstracted)[131] or is it rather due to a peculiar property of the concept of chance as an abstractor? I call this question transversal, since its possible answers still concern the epistemological level. The second question reflects an orthogonal movement. Does the fact that chance as an abstractor has an explanatory power depends only on the epistemological factors mentioned above, or does it also reflect specific properties of the phenomena that chance can account for? This last question I call orthogonal since we ask whether, in addition to the epistemological virtues, chance as an abstractor also reflects the real properties of the explained phenomena. Let's summarize these two questions in the following Figure 2:

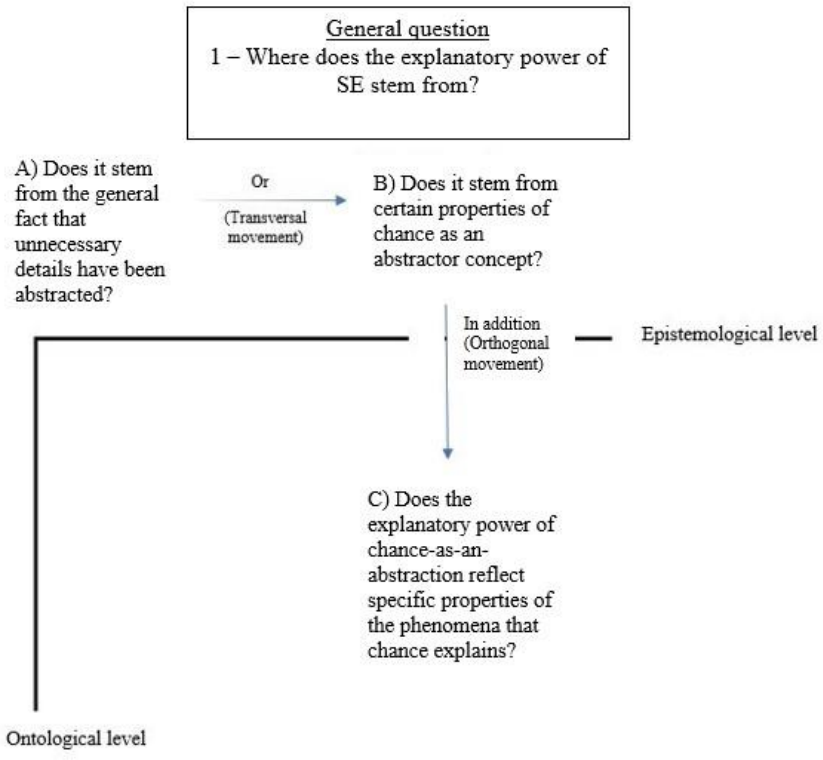


Figure 2 of the transversal and orthogonal question. For further details, see main text

[131] In literature we can find the argument that abstract explanation could be a virtue, at least for the explanation of certain phenomena (see e.g. Levy 2014; Strevens 2019).

In the Figure 2 we can see that from the general question, different sub-questions can follow. The first two sub-questions A and B are transversal since they move but remain on an epistemological level. Sub-question C shifts to the ontological level, because it concerns the ontological properties of the explained phenomena. Let's try to answer questions A and B first.

Where does the explanatory power of SE stem from? Does it stem from the general fact that details are abstracted away, or from certain properties of the concept of chance as an abstractor? (Figure 2 questions 1, A, B.) I answer positively to question B, the explanatory power of SE stemming from the concept of chance-as-an-abstraction. To justify this answer, it might be helpful to compare SE with the two accounts of statistical explanation sketched above. In Lange's and Ariew *et al*'s accounts described in section 1.2.2, it is implied that statistical signatures or/and statistical models enable causal details to be abstracted away. "To abstract away" in this context means there is no room for causal details in the explanation. The details are ruled out of the explanation because they are not needed (for instance, an individual's behavior won't explain certain dynamics in population genetics). By contrast, in my account, if it is true that the *explanans* (complete with the chance element) abstracts away the different ways in which the *explanandum* could be realized, it is also true that this *explanans* does not rule these ways out of the explanation. Instead, it *synthesizes them*. How? Thanks to the concept of chance. When we say that the microfibrils are "random arranged" we are abstracting away all their possible orientations – but we are also synthetizing them by using the notion of "random". In defining the process of LY abstraction, I wrote that the explanans can *hold* and *synthesize* the different ways in which the *explanandum* could be realized. In this context, the verbs to "hold" and "synthetize" can become more consistent. In Lange's and Ariew *et al*'s accounts it is a matter of abstracting in the sense of preventing causal details from entering the explanation. However, in my case it is a matter of abstracting by not explicitly noting in the explanation the variety of possible ways in which the *explanandum* can be realized. However, it is also a matter of "holding" these ways together (i.e. including them in the explanation). This inclusion occurs through a synthesis which is possible thanks to the use of notions such as "random", "stochastic", chancy" etc. In summary, in order to answer the Figure 2's question B in a few short words, I can also say that chance as an abstractor holds and synthetizes *into* the explanation the multiple realizability of a process. So, the explanatory power does not come from the simple fact that there is an abstraction of unnecessary details (answer to A) but rather from the fact that the properties of abstraction and of synthesis stem directly from the notion of chance (answer to B).

What about sub-question C? Does the explanatory power of chance-as-an-abstraction reflect specific properties of the phenomena that chance explains? In the introduction, I formulated questions such as: if chance is as an abstractor in the *explanation* of a given phenomenon, what can we say about the explained *phenomenon*? Does it mean that chance as an abstractor can be used to account for a given phenomenon by virtue of some particular features? In other words, does chance as an abstractor imply that the explained phenomenon itself is chancy? And if so, what is the relationship between chance as an abstractor and chance out there? These questions suggest that since chance as an abstractor works only to explain *certain* phenomena and not others, it is legitimate to question *what kind* they are.

It is not one of the goals of this thesis to answer this question. In fact, I would like to return to my epistemological stance, which is the reframing of the whole present work (cf. Introduction, section 1.4). The fact that SE can only be applied to certain phenomena and not others obviously depends on certain properties of those phenomena. There is some interesting ontological work that could be done taking question (C) as a starting point. However, I do not include it here. Future work can address this kind of analysis. In this context, regarding the link between chance and phenomena, it is sufficient to refer to biological practice. Indeed, one of the starting points of my work is that biological explanations already use chance to explain biological phenomena. And this instance, which regards the addition of nucleotides to new strands of DNA during replication, stems from biological practice and provides a more than sufficient basis for the epistemological requirements of this thesis (see section 1.4).

In the last sections, I defined what I mean by chance as present contingency (CPC) and the process of LY abstraction (PLA) (section 2.2), asking relevant questions that this account can give rise to (section 2.3). I provided some "short" examples of the three ways in which CPC and PLA can be coupled and the purpose of the next section is to provide a more detailed discussion concerning these examples. The main reason I will go into so much detail with respect to these examples is to show that my philosophical account of SE is already implicit in many, or at least in certain, biological explanations. Being able to argue this position is crucial to the value of my work, which takes a bottom-up approach to Philosophy of Science in Practice (PSP) and thereby aims to propose philosophical projects that are "potentially relevant" (Kaiser 2015, p. 40) for scientific practice (cf. Chapter 1).

3. Three examples of stochastic explanations

3.1. Stochastic explanation (SE) in terms of disjunctive representations

The first example of stochastic explanation (SE) that I propose is the one in which we have an abstraction in terms of the synthesis of disjunctive representations of the *explanans*. I discovered this explanation in research on the olfactory receptor gene (hereafter OR genes). The odorant receptors are encoded by a large family of OR genes (Buck and Axel 1991). The most accredited – but still questioned[132] – hypothesis is that a single olfactory sensory neuron (hereafter OSN) expresses a single OR gene. This is called the "one neuron – one receptor" hypothesis (for a review, see Mombaerts 2004). In each olfactory sensory neuron (OSN), the olfactory receptor genes (ORs) are expressed in a specific zone called the olfactory epithelium (hereafter OE).

Odorant receptor Box
OR genes = olfactory receptor genes
OSN = olfactory sensory neuron
OE = Olfactory epithelium

I focus on a paper review by Fuss and Ray (2009) in which the authors assume that "each neuron […] expresses just one odorant receptor gene [i.e. the "one neuron – one receptor" hypothesis]" and ask: "how does a single neuron choose to express just one of many possible odorant receptors and exclude expression of all the others?" (p. 101). This question is particularly relevant because individual neurons "choose to express only one (and sometimes two) odorant receptor from a genomic repertoire that can *be as large as 1500 genes*" (p. 102; emphasis added). They stress that even though several years have passed since the identification of OR genes, "yet little is known about the mechanisms[133] regulating their expression" (*ibid.*). Nonetheless, Fuss and Ray summarize the main models of OR genes expression developed up to 2009. The authors emphasize that these models are all compatible with available data and that they are not mutually exclusive (p. 102).

[132] "[T]he general validity of the 'one neuron – one receptor' hypothesis is hard to prove" (Fuss and Ray 2009, p. 107).

[133] In this context, using the notion of mechanism could give rise to an objection. How could it be the case that in giving an example of SE, namely of a non-mechanistic explanation, I am providing an explanation in which the biologists themselves use the notion of "mechanism"? This could be seen as a short-circuit since I am actually giving an example of mechanistic explanation. But this is not the case. In this specific example, when the authors ask about mechanisms, they simply ask how a neuron can choose one specific OR gene among 1,500 possibilities. There is no metaphysical commitment to entities, activities or regularity. In light of this, we can simply call and think of this dynamic in the less loaded concept of "process". For a critique of mechanisms and mechanistic explanation see Chapter 2. For a discussion on mechanisms, stochasticity and metaphysics, see the conclusion of the dissertation, section 3.

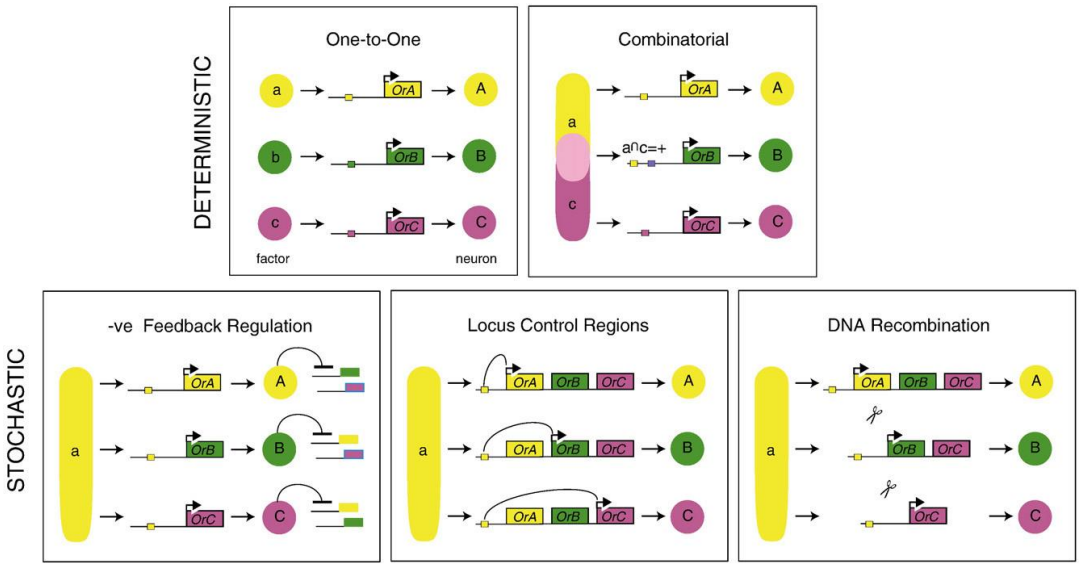


Figure 3 "Various models for patterning of one-receptor-per-neuron in the olfactory system: deterministic (top) and stochastic (bottom)" (for further details see main text; image and quote by Fuss and Ray 2009, p. 103).

The specificity of an olfactory sensory neuron (OSN) depends on which olfactory receptor (OR) gene is expressed in its olfactory epithelium (OE). The models in figure 3 explain how this specificity can take place. Two are deterministic and three are stochastic. Let's begin with the first deterministic model on the top left. This model depicts how specific transcription factors[134] ("a", "b", "c" on the left) activate specific OR genes. The OR genes, in turn, encode for specific OR receptors. For example, the yellow transcription factor "a" (top, on the left) activates the yellow gene OrA which, in turn, encodes for the yellow olfactory receptor A. The green transcription factor "b" (on top, on the left, second line) activates the green gene OrB which, in turn, encodes for the green olfactory receptor B; and so on. This model is deterministic since there is a linear relationship between the transcription factors and the consequent olfactory receptors encoded (e.g., if factor "a", *then* olfactory receptor A through the OrA gene).[135]

The stochastic models in figure 3 are what is interesting here. I focus only on the first model on the left because the others share the same idea. In it, the yellow transcription factor "a" can give rise to different olfactory receptors (i.e. yellow, green, purple olfactory receptors "A", "B", "C") through the activation of different OR genes (i.e. yellow, green, purple OR gene

[134] Transcription factors are complex systems of proteins that bind DNA and permit RNA polymerases to translate sequences of DNA. RNA production is usually described as *bursts*, namely a random explosion of RNA copies (see also Eldar and Elowitz 2010, p. 168; the term *bursts* is also associated with the explosion of protein copies during the process of translation mediated by ribosomes). Note that in some works scientists talk about "intrinsic disorder" in transcription factors, i.e. it is observed that the binding sites are not specific but can bind contingently different molecules (Liu *et al* 2006).

[135] The second deterministic model (Figure 3, on top, on the right) differs from the first to the extent that, in addition to the deterministic link just described (e.g. if factor a, then olfactory receptor A), it could be also the case that a *combination* of different transcription factors activates a specific OR gene which, in turn, encodes for an OR receptor (e.g. if a∩c factors, *then* olfactory receptor B through OrB).

"OrA", "OrB", "OrC"). Then, in the right part of the model, we can see a representation of a feedback mechanism through which the expressed olfactory receptor blocks the expression of the other two.[136]

Let's frame this model in terms of *explanans* and *explanandum*. The *explanandum* is the expression by a single neuron of just one OR gene out of 1,500 possibilities. How can that be explained? By chance! More precisely, the *explanans* is the fact that the yellow transcription factor "a" "*may stochastically* initiate expression of a single of a few receptors per cells" (Fuss and Ray 2009, p. 102; emphasis added). Then, a feedback mechanism provided by the expressed olfactory receptor avoids the expression of other OR genes. The paper by Ray (2009) does not specify if their conception of stochasticity refers to the fact that each gene has the same probability of being expressed or follows a specific probability distribution. But this specification is not necessary since the explanatory power of chance as present contingency (CPC) resides in its abstract virtues (cf. section 2.3) and not in the specification of its probability distribution (if any). In this first case, the *explanans* of this explanation provides an abstraction in terms of the synthesis of *disjunctive* representations of the possible ways in which the *explanandum* could be realized. In Figure 3's first stochastic model, the represented stochastic behavior of the yellow transcription factor "a" could be intended as a specific case of abstraction in this example of stochastic explanation. Let's see why. I asserted that disjunctive representations could be framed as follows: the behaviour of a biological process P could be (x) or not (x); or could be (x) or (y) or (z) or (t), etc. Let's say that process P, which has to be explained (i.e. the *explanandum*), is the expression by a single neuron of just one OR gene out of 1,500 possibilities. And let's say that the *explanans* is the stochastic behaviour of the yellow transcription factor "a". Finally, let's say that (x), (y), (z)[137] are the possible disjunctive representations of the expression of OR genes (i.e. OR genes "OrA", "OrB", "OrC").

Then, in this case,

It is the explanans *of the stochastic model depicted in Figure 3 (first stochastic model on the right) that contains the notion of chance as present contingency (CPC). This CPC provides an* explanans *of a stochastic explanation (SE) though the process of LY abstraction (PLA) by*

[136] For example, the yellow olfactory receptor A inhibits the expression of green and purple OR genes. With the representation of this feedback process the authors answer the second part of the question: "how does a single neuron choose to express just one of many possible odorant receptors *and exclude expression of all the others?*" (Fuss and Ray 2009, p 101; emphasis added).

[137] In this case study, we have only three disjunctive elements (x), (z) and (y) that correspond to three expressions of OR genes (i.e. "a", "b", "c"). But in general, there are no limitations on the numbers of disjunctive elements. We will see in the example of the synthesis of antibodies that these elements can also tend to infinite (see Chapter 5).

*synthesizing the **disjunctive representations** of a variety of possible ways in which the explanandum could be realized. Stating that transcription factors "may stochastically initiate" (CPC) means that a transcription factor "a" may activate yellow OR gene or green OR gene or purple OR gene. The* explanans *holds and synthesizes representations of different disjunctive OR gene expression, without spelling out their details (PLA).*

3.2. Stochastic explanation (SE) in terms of cumulative representations

Let's now give an example that indicates the second case of chance as an abstractor in terms of cumulative representations of the *explanans*. This example comes from a textbook of plant physiology (i.e. Evert and Eichhorn 2013) and accounts for cell wall expansion.[138] The cell wall is one of the most famous differences between animal and plant cells. The wall has numerous cellular functions: it prevents the rupture of the protoplast (i.e. cytoplasm and nucleus) due to water pressure; it determines the size and shape of the cell and the texture of the tissue; it plays an active role in defense against external pathogens, "processing information from the surface of the pathogen and transmitting this information to the plasma membrane of the plant cell" (Evert and Eichhorn 2013, p. 57); and it contains a variety of enzymes and plays important roles in the absorption, transport and secretion of substances in plants (p. 57). The cell wall is principally composed of cellulose, i.e. repeated monomers of glucose linked end to end. Different chains are bound together through hydrogen bonds, forming microfibrils of 10 to 25 nanometers in diameter. The other components of the wall are hemicelluloses, pectins and glycoproteins (e.g. extensins).

Evert and Eichhorn write that the "[e]xtension of the wall is a complex process under the close biochemical control of the protoplast" (p. 60-61) in which are involved, among others, wall proteins like expansins and hormones. The authors propose two models to explain cell wall expansion that reflects two possible ways in which "the orientation of cellulose microfibrils […] influences the direction of expansion" (p. 61):

"(a) If the cellulose microfibrils are *randomly* oriented in all walls, the cell will expand equally in all directions[139], tending to become spherical in shape. (b) If the microfibrils are oriented at right angles to

[138] I take Evert and Eichhorn (2013) as a reference to describe that case study but the reader can easily find similar examples in any other plant physiology textbooks.

[139] But a cell wall is of a certain thickness, therefore they cannot be oriented any which way in 3D – only in 2D, or on a plane of a certain thickness. If they were 3D, they would grow into a sphere, leaving no space for the interior of the cell. On a plane, they grow in a sphere but with a hollow at the center where the rest of the cell is. Therefore, more correctly we should say that microfibrils can orient themselves in all directions only in a certain

the ultimate long axis of the cell, the cell will expand longitudinally along that axis" (*ibid.*, emphasis added; see Figure 4).

What is of interest to me here is the explanation provided with model (a), namely that the spherical growth of cells is *explainable* by fact that there is a "random arrangement" (p. 60)[140] in the orientation of the microfibrils during cell growth. The chance element of the explanation, in this case the random rearrangement, is essential to explaining the phenomenon under attention.

In order to underline chance's essential role, let's use an explanation that does not refer to random rearrangement. In this hypothetical scenario, we could have two situations: 1) We don't know anything about the orientation of microfibrils and we cannot say anything about their shape; therefore, in this case we have no explanation at all. 2) We provide a model in which the microfibrils are all oriented at a specific angle. In this case, we turn to explanation (b), which could not replace (a) because it refers to another biological scenario and is thus *another explanation*. So, random rearrangement is essential to explaining the spherical growth of plant cell walls (cf. section 4.2).

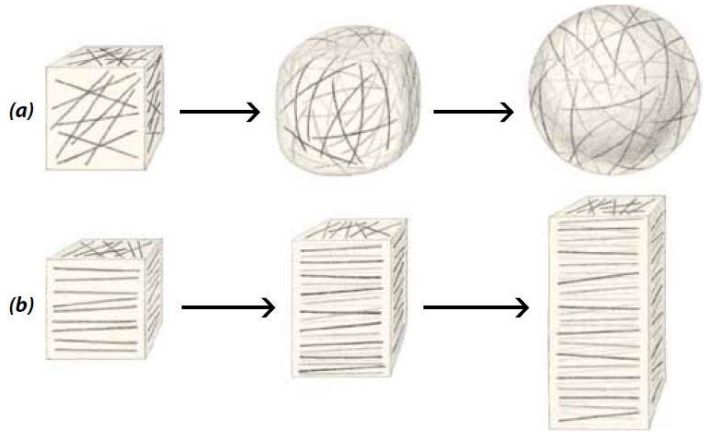


Figure 4 Two models to explain two different shapes in the wall cell expansion (see main text; image by Evert and Eichhorn 2013, p. 61).

From where does chance's explanatory power originate in this explanation? The answer has to be found by looking to the notion of abstraction that this explanation (implicitly) possesses. Let's say that process P, the spherical shape of a plant cell, has to be explained (i.e. the

range of space, that is, within the thickness that will constitute the wall of the cell. Thanks to Katherine Waters for making this point clearer.

[140] Curiously, we find the same expression, "random rearrangement", for a different case study in which chance has another crucial explanatory power: the synthesis of antibodies in B-Lymphocytes (see, e.g., Mak and Saunders, 2008). I will provide an analysis of this process in Chapter 5.

explanandum). And let's say that the *explanans* is the random rearrangement of the orientation of microfibrils in all walls. Finally, let's say that (x), (y), (z), (t), etc. are all the possible rearrangements of the orientation of microfibrils in the wall.

Then,

The explanans *in the explanation of cell wall expansion (Figure 4, a) contains a notion of chance as present contingency (CPC). This CPC provides an* explanans *of a stochastic explanation (SE) though the process of LY abstraction (PLA) by making the synthesis of* ***cumulative representation*** *of a variety of possible ways in which the* explanandum *can be realized.*[141] *In this case, the random arrangement of the orientation of microfibrils in the cell wall (CPC) holds and synthesizes cumulative representations of the microfibrils' directions, without actually spelling out their details (PLA).*

3.3. Stochastic explanation (SE) in terms of cumulative and disjunctive representations

Now, I will provide a third relevant example to help highlight the third case of CPC in which the *explanans* holds and synthesizes cumulative *and* disjunctive representations of the different ways in which the *explanandum* can be realized. This example is the explanation of DNA replication that can be found in any molecular biology textbook. DNA polymerase is the enzyme that synthesizes DNA molecules from nucleotides, the building blocks of DNA. It synthesizes 50-500 nucleotides per second, positioning the new nucleotides on the template strand in the correct way. By "correct way", I mean the situation in which the new nucleotide fits the nucleotide of the old DNA strand, respecting Chargaff's rules (Chargaff Lipshitz, and Green 1952; Elson and Chargaff 1952) – so that a guanine is always bound with a cytosine and a thymine is always bound with an adenine. The new nucleotide connects to the newly synthesized DNA strand in two stages: (1) the new nucleotide creates a hydrogen bond with the nucleotide of the template strand (see Figure 5, stage 1, new hydrogen bond in orange); (2) polymerase creates covalent bonds between the new nucleotide and the newly synthesized DNA strand (see Figure 5, stage 2, new covalent bond in red).

I propose that this explanation contains an important and hitherto neglected additional stage. In order to make sense of this process, an additional stage (0), which precedes stages (1)

[141] In this context, we can also talk in terms of the principle of indifference, that is to say that since we have no reason to say that the orientation of microfibrils is in one direction rather than another, we give to all directions an equal probability. But my argument seeks to go beyond this principle. In fact, I am proposing that it is exactly because CPC abstracts and synthetizes all these directions that we have an explanation of why the cell wall is sperical in shape.

and (2), should be considered. In the nucleus, nucleotides move by diffusion; there is no active transport toward DNA polymerases. In light of that, I hypothesize that at stage (0), there is a random sampling of nucleotides (i.e., an indiscriminate sampling process) in which the physico-chemical differences between nucleotides play no causal role in the choice of the nucleotides of the newly synthetized DNA strand. In fact, there is a long-standing debate around what is more important for fidelity in DNA replication: hydrogen bonds between nucleotides or geometry. Geometry here designates the "size and shape of the molecule" (Stamos 2001, p. 174). Whatever the answer is, this question does not affect stage (0), since the hydrogen bond and the geometry are established/checked after stage (0), that is to say at stage (1). Stage (0) then could be seen as an indiscriminate sampling process to the extent that it has a role in randomly sampling new nucleotides for the new template (see Figure 5, stage 0, indiscriminate sampling process in green). More explicitly, the nucleotides are indiscriminately sampled. DNA polymerase then checks if a certain nucleotide is the good one for a certain position,[142] and after that, stages (1) and (2) take place. Going even further, it could be argued that this chance element at stage (0) is necessary for the *explanation* of hydrogen bond formation between parallel nucleotides as well as for that of covalent bond formation among the new nucleotides. This would mean that stage (0) allows stages (1) and (2).

It is often written in biological textbooks that the affinity between polymerase and the correct nucleotide is higher than the affinity between the polymerase and the incorrect nucleotide "because the correct pairing is more energetically favorable" (see Alberts *et al* 2008, p. 269). This could be seen as an objection to my hypothesis since this evidence jeopardizes my assumption that stage (0) can be described as an *indiscriminate* sampling process. It seems that the sampling process is in fact *discriminated by* energy affinity. Nonetheless, the movement of free nucleotides in the nucleus microenvironment is described by diffusion which is "the ubiquitous *random motion* of small particles due to thermal agitation" (Brangwynne *et al* 2009, p. 423; emphasis added). Even if the affinity between a correct nucleotide and the moving polymerase is higher than the affinity between an incorrect one, the possibility of their interaction is still chancy since it still depends in *prima facie* on nucleotide diffusion. Therefore, even if at the beginning of stage (1) the sampling process is *discriminate*, I can maintain that stage (0) is *indiscriminate*.

[142] This process is often called "double-check" in which the DNA polymerase's "fingers" tighten around the active site and check if the new nucleotide fits the complementary one (see Alberts *et al* 2008, p. 269).

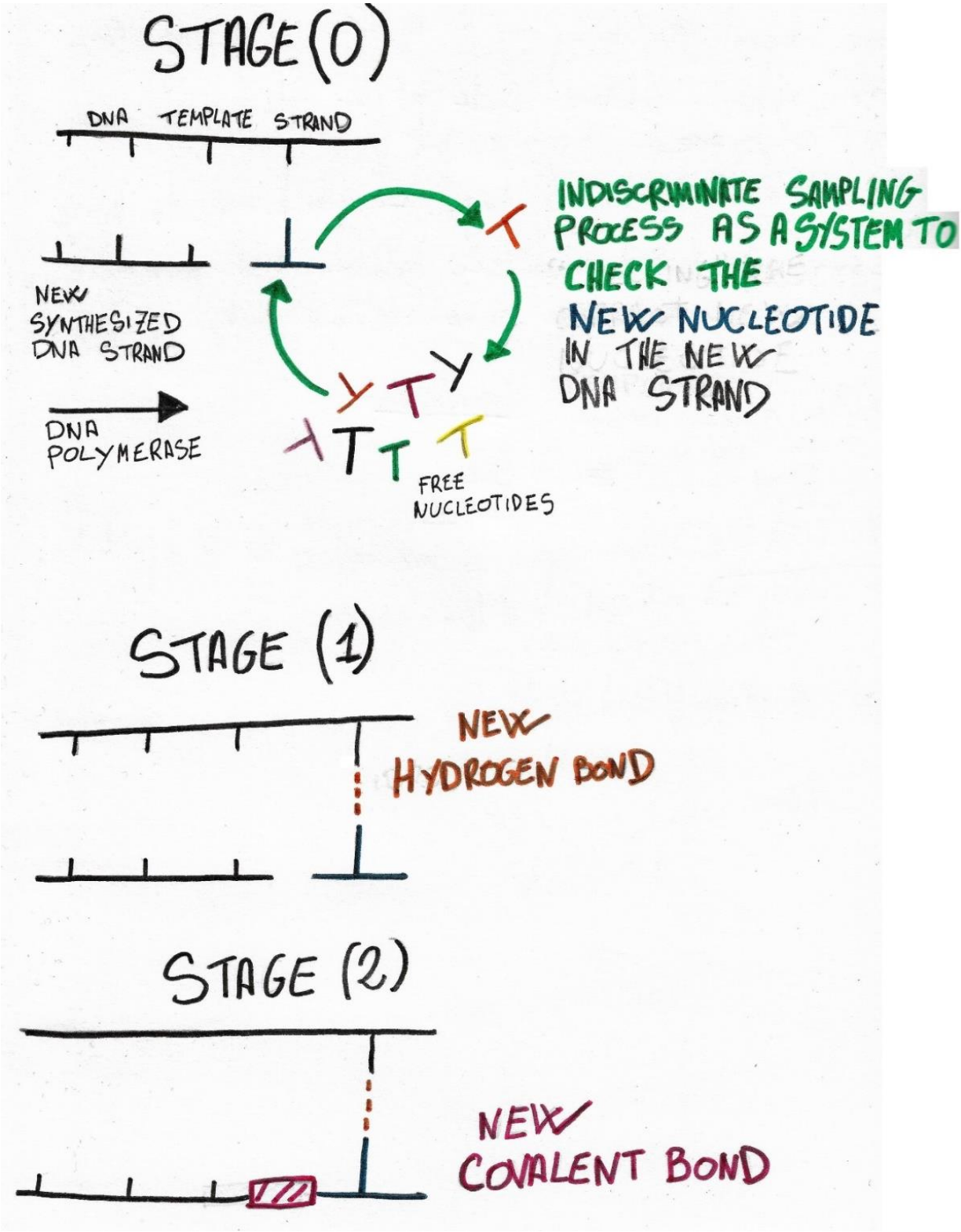


Figure 5 DNA replication model. Qualitative representation of the stages (0), (1) and (2). For further detail, see main text.

Let's frame this model in terms of the *explanans* and *explanandum* of an explanation. The *explanandum* is the addition of nucleotides to the new strand in a way that follows Chargaff's rules. What is its *explanans*? My answer is that the *explanans* consists of the three stages that I have just described (stage, 0, 1, 2).[143] I mainly make explicit the relevance of stage (0) since it is the one related to chance. Note that the description of stage (0) is my original contribution to the explanation of DNA replication and that it is not highlighted in molecular biology textbooks. Yet it is essential to explaining and understanding stages (1) and (2), and provides an excellent example of the third case of chance as an abstractor in this stochastic explanation. The disjunctive and cumulative representations could be framed as follows: the behaviour of a biological process P could be (x) and/or (y) and/or (z) and/or (t), and so on. Let's say that the fact about P that we have to explain (i.e. the *explanandum*) is the addition of nucleotides following Chargaff's rules. Let's say also that the *explanans* is stage (0) because it is the key process that enables stage (1) and stage (2). Finally, let's say that event (x) corresponds to the correct positioning of an adenine on the DNA strand, and event (z) to the correct positioning of a guanine. In the following Figure 6, I depict chance as an abstractor in term of cumulative representations:

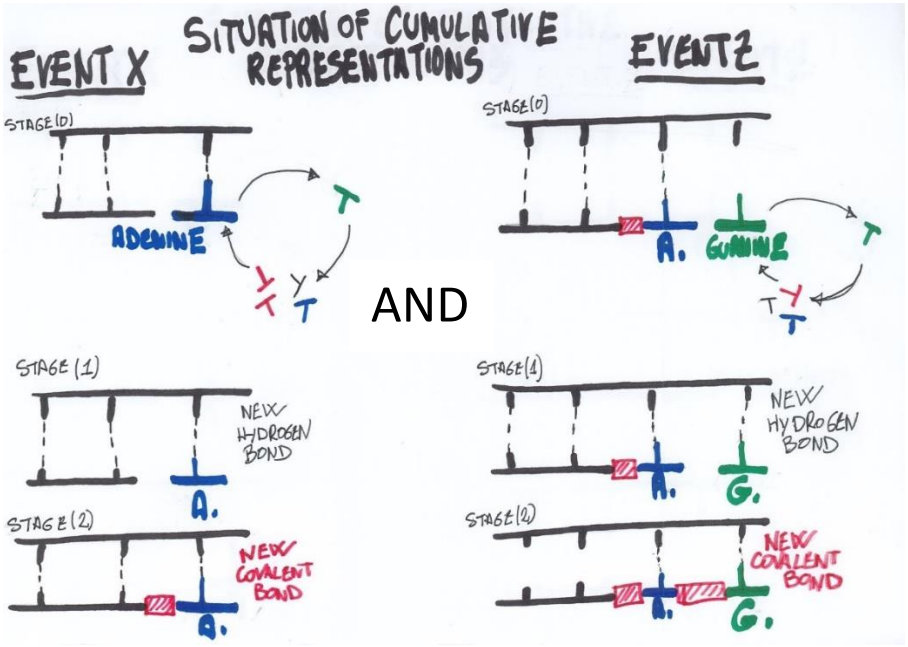


Figure 6 A situation of cumulative representations, in which both event X and event Z take place. For further detail, see main text.

[143] I am aware that this is only a part of the explanation of DNA replication. Here it lacks the representation of DNA polymerases, Okazaki fragments, primase, DNA ligase, replication forks, and so on. Nonetheless, giving the complete explanation of DNA replication (as it is presented in any textbook of molecular biology) is not of use here. In this context, I want only to underline how this kind of explanation is actually an example of SE in which the CPC provides PLA in terms of cumulative and disjunctive representation of the different ways in which the *explanandum* could be realized.

In Figure 6, we can see a possible situation in which both events X and Z take place. In this case, chance as present contingency (CPC) provides an *explanans* of a stochastic explanation (SE) though the process of LY abstraction (PLA) by synthesizing the **cumulative representations** of a variety of possible ways in which the *explanandum* could be realized. More specifically, as an indiscriminate sampling process, stage (0) permits explanation of the case in which we have event X *and* event Z: random sampling of different nucleotides at stage (0) permits explanation of the addition of an adenine at T_1 and of a guanine at T_2. In Figure 6, we can see how through stages (0), (1) and (2), an adenine (left) *and* a guanine (right) are bound to the new strand.

Stage (0) can also explain hypothetical cases in which one nucleotide *or* another can be added at T_x. In this case, the focus is on a single T_x. In Figure 7, I depict the case in which only event (x) *or* event (z) takes place, namely cases in which an adenine (left) *or* a guanine (right) are bound to the new strand. Chance as present contingency (CPC) provides an *explanans* of a stochastic explanation (SE) through the process of LY abstraction (PLA) by synthesizing the **disjunctive representations** of a variety of possible ways in which the explanandum could be realized. More specifically, stage (0) in terms of an indiscriminate sampling process permits explanation of both the cases in which we have event (x) *or* event (z). Random sampling of different nucleotides at stage (0) enables the explanation that we may have an addition of an adenine *or* a guanine a single T_x.

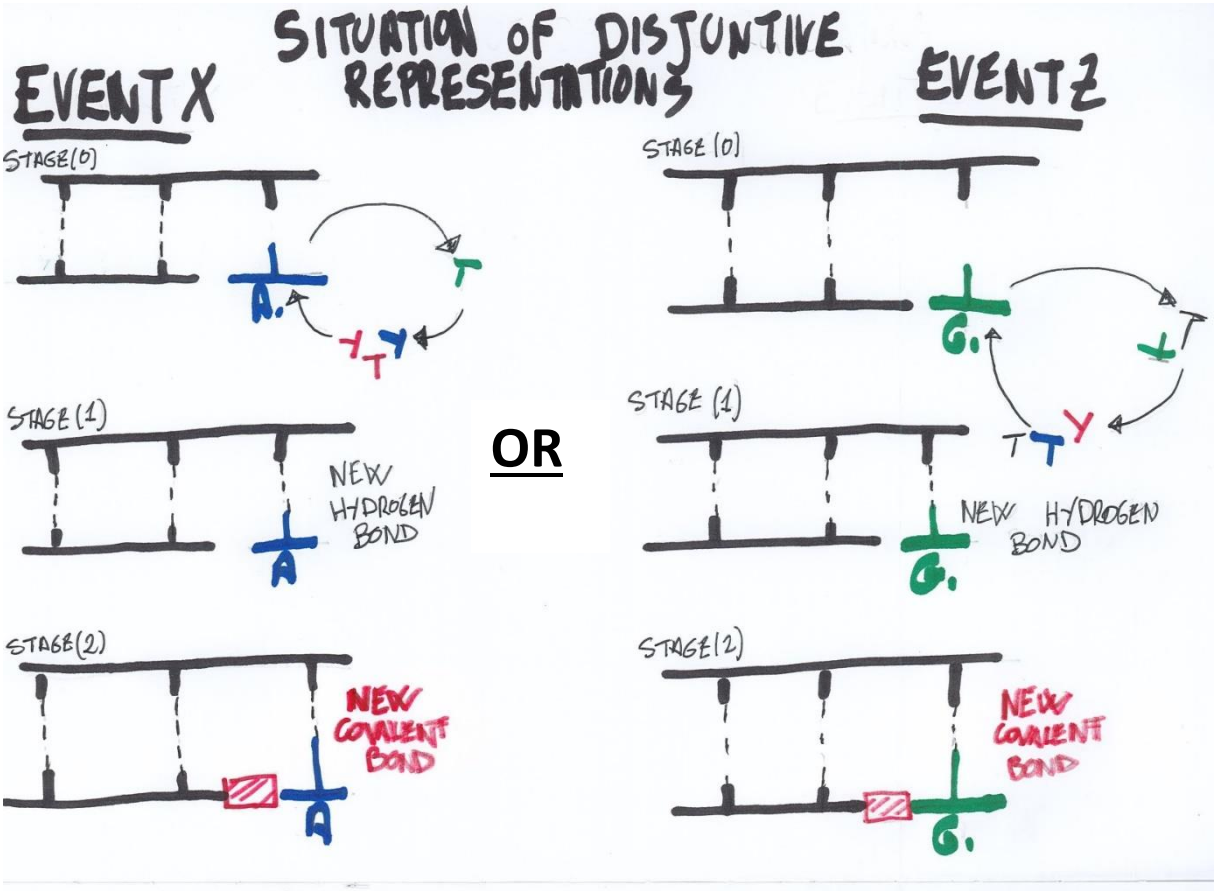


Figure 7 A situation of disjunctive representations, in which event X or event Z can take place. For further detail, see main text.

What I am proposing is that the explanation of DNA replication is an SE in which chance has a main explanatory role. Moreover, this explanation is an interesting case in which we can have both cumulative and disjunctive representations.

Then,

The explanans *of the explanation of DNA replication (Figures 5, 6, 7) contains notions of chance as present contingency (CPC). This CPC provides an* explanans *of a stochastic explanation (SE) through the process of LY abstraction (PLA) by synthesizing the* ***cumulative and disjunctive representations*** *of a variety of possible ways in which the explanandum could be realized. In this case, CPC is the random sampling of the new nucleotide – that is to say stage (0) – which holds and synthesizes cumulative and disjunctive representations of events X and/or Z without actually spelling out their details (PLA).*

In summary, through these three detailed examples, I have shown that certain biological explanations are stochastic explanations (SEs) insofar as they already implicitly possess an idea of chance as present contingency, and consequently, of chance as an abstractor. This idea is

hard to grasp if we do not approach these explanations with a philosophical analysis, and I have provided one in these last sections. The philosophical work here has been to make explicit the explanatory value of chance, which I have achieved thanks to the conceptual tools I developed in this chapter – namely chance as present contingency (CPC), the process of LY abstraction (PLA), and, more generally, stochastic explanation (SE). In making explicit this epistemic value I also made evaluative claims, namely, I claimed that certain biological explanations *are good* explanations thanks to their stochastic character (i.e., the explanatory chance element in them). However, I still have not justified why SEs are, in fact, good explanations. That is precisely what I attempt to do in the following section.

3.4. SEs are "good" and "better" explanations

In this section, I want to propose that the SEs just provided are better than hypothetical explanations of the same phenomena which use "more details are better" (MDB) as a criterion for developing a non-stochastic explanans. This is because SEs have a higher cognitive salience and lower sensitivity.

Cognitive salience and sensitivity as a measure of the goodness of an explanation were highlighted for the first time by Ylikoski and Kuorikoski (2010)[144] who called them "dimensions of the goodness" (p. 201). By this, they mean that these notions are developed to unpack the (very often) vague concept of explanatory power in the context of explanations (cf. section 1). In their 2010 paper they open a reflection, asking what makes one explanation (x) *better* than another explanation (y)? They develop five dimensions of goodness, including cognitive salience and sensibility, to provide the framework from which to answer that question:

"[c]ognitive salience refers to the ease with which the reasoning behind the explanation can be followed, how easily the implications of the explanation can be seen and how easy it is to evaluate the scope of the explanation and identify possible defeaters or caveats" (p. 214; emphasis added).

"[t]he more *sensitive* an explanation is to changes in background factors, the less powerful it is" (*ibid.* p. 208; emphasis added).

With respect to the first criterion, the description is quite clear and I would like to use the same definition of cognitive salience for specifying why SEs provide a better explanation. The clearer

[144] The work by Ylikoski and Kuorikoski (2010) is one of the first to make explicit in which sense an explanation can be seen as a better explanation, using epistemic norms. The notions that they proposed have inspired other philosophers. See for example the work by Baedke (2018).

the reasoning, implications and scope, the more we gain in term of cognitive salience of an explanation. I differ slightly regarding the meaning of sensitivity. Ylikoski and Kuorikoski develop this dimension (and all the others) by embracing a specific account of explanation, namely the contrastive-counterfactual account of explanation (Woodward 2003). In this framework, the more sensitive an explanation, the less powerful it is. If I alter a small detail in the background conditions and see that the relation between *explanans* and *explanandum* changes, then I am in front of an explanation with a high degree of sensitivity. An explanation with high sensitivity is fragile. Note that the expression "background conditions" is a technical term which refers directly to Woodward's account of explanation. But Ylikoski and Kuorikoski (2010) underline that their dimensions are generalizable also to different accounts of explanation and to explanation more broadly. They write: "a supporter of a competing theory of explanation might accept our taxonomy of explanatory virtues but prefer a different account of their source" (p. 203). I take this advice seriously since I want to use these two dimensions to make clear this section's point without, at the same time, embracing Woodward's account of explanation.

More generally, I would like to note that these authors, in order to be able to measure the goodness of an explanation in terms of cognitive salience and sensitivity, actually refer to *better* explanations since they provide comparative scenarios. If an explanation (x) has a higher cognitive salience and lower sensitivity of explanation (y) – and assuming that we use these two dimensions as a good parameter to measure the goodness of explanation – then (x) is a *better explanation* than (y).[145] Although the two notions are often used interchangeably in the article by Ylikoski and Kuorikoski (2010), I find instead that emphasizing their difference is useful in the present context. In fact, I find use of the notion of "good" problematic when referring to an explanation because it implies the existence of some absolute measure for the goodness of explanation. But finding this absolute measure is not a good strategy because whether an explanation is "good" or not depends strictly on the research context and the goals set by the researchers. In order to avoid this "goodness" impasse, I plan to argue that the SEs just prosed in the previous section are not good in absolute terms but *better* than other explanations. The discussion then, will be *comparative*. The problem now arises, however, of specifying the "other" explanations to which we are comparing SE. I propose using the notion of chance as an abstractor to allow us to show that SEs are better explanations compared to a

[145] Note that Ylikoski and Kuorikoski (2010) themselves stress that it is not always so easy to compare explanations through the dimensions they provide.

hypothetical explanation accounting for the same phenomena (i.e., the same *explanandum*) that uses a "more details are better", non-stochastic *explanans.*

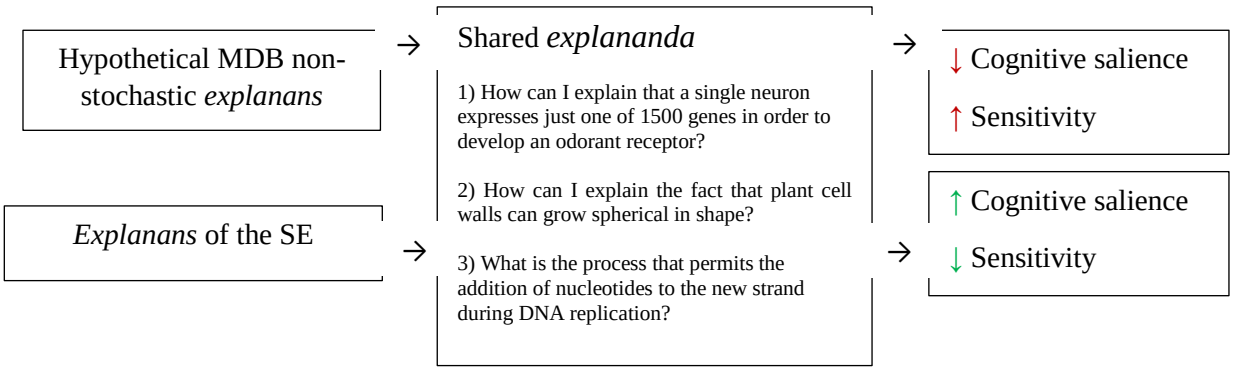


Figure 8 Representation of the differences in cognitive salience and sensitivity between a hypothetical mechanistic non-stochastic *explanans* and the SE *explanans.*

Following Figure 8, we can see on the left the hypothetical, non-stochastic MDB *explanans* and the SE *explanans*. At the center, I depicted the three SE *explananda* from the previous section framed as questions: 1) odorant olfactory gene expression, 2) cell wall expansion, and 2) DNA replication. On the right, I show increases or decreases in cognitive salience and sensitivity. For example, if my *explanandum* is the spherical shape of a cell wall, using the SE *explanans* of the "random rearrangement" of microfibrils is better than giving a non-stochastic MDB *explanans* that describes all the orientations the microfibrils can take. The former is better since it increases cognitive salience and decreases sensitivity. By contrast, describing all the possible directions that microfibrils could take it would make the explanation too complex and sensitive to changes in background factors.

The following quote by Douglas allows me to explain what I am arguing for in a different way:

"[w]e are finite beings, with finite mental capacities. We need explanations to grapple with […] this [complex world]. Explanations help us to organize the complex world we encounter, making it cognitively manageable." (Douglas 2009, p. 13)

MDB decreases cognitive salience since the more details we have, the harder it is for a finite human being "to grapple" (Douglass 2009, p. 13) with them all in a single explanation. By contrast, chance as present contingency (CPC) with its proprieties of abstraction and synthesis enables increases in cognitive salience. Furthermore, MDB increases sensitivity since the more

details included in the explanation, the more it risks failing if these details change. The abstraction in CPC permits robust explanations, which in this case means explanations that work even if there are changes in the background factors.

Conclusion

In this chapter, I provided a brief overview concerning abstraction and explanation in philosophy of science. From these sketches of the literature, I claimed that linking chance with abstraction is an original and unexplored idea. Furthermore, in analyzing statistical accounts of explanation, I specified my position regarding the causal vs. non-causal debate of explanation by embracing what I termed "causal agnosticism". Whenever causes are specified or not included in explanations that I label "stochastic", my argument remains untroubled as it has at its core not causes but rather the abstract relation between the *explanans* and the *explanandum* within a single explanation.

I then provided a proposal concerning the third positive criteria of adequacy of my account of SE. I dedicated many sections to this criterium because it is the "positive" one which prescribes what stochastic explanation (SE) corresponds to. In presenting this criterion, I developed the notion of the process of LY abstraction (PLA) and the notion of chance as present contingency (CPC). I propose the following Table 1 in order to sum up the ideas I have up to now developed:

Stochastic explanation (SE)	A stochastic explanation is a philosophical category that refers to any explanation with a chancy element that provides (explicitly or implicitly) (part of) its explanatory power.
The process of LY abstraction (PLA).	To any given SE, the *explanans* can hold and synthesize disjunctive and/or cumulative representations of a variety of possible ways in which the *explanandum* (e.g. the phenomenon to be explained) could be realized without the need to provide the details of all those representations.
Chance as present contingency (CPC)	Given some fixed initial conditions, chance is a property of a description of a process. This property specifies that the described process (or at least some events in that process) could take place, not take place or take place in different ways.

Table 1 of the different definitions developed in the present chapter

Interlude

Chapter 4 (Interlude): Chance in philosophy of biology. What has already been done and what happens next

Introduction

This chapter has been designed as an interlude, a middle space between the two moments of the previous three chapters and three that follow. What I have not yet proposed in this dissertation is a state of the art which gives the reader an idea of what has already been done in philosophy with respect to chance in molecular and cellular biology. However, it is not my intention to develop an exhaustive state of the art because there are not many philosophical reflections on chance in cellular and molecular biology and because in the present chapter I want instead to focus on studies that can be functional for my argument. These studies can 1) substantiate the relevance of my account of stochastic explanation (SE) that I developed in the previous two chapters and 2) serve as inspiration for the development of the next three chapters.

In the first section, I emphasize that developing an account of chance without including the notion of probability (cf. Introduction, section 1.4) has already been proposed at least in other study contexts. Montévil *et al* (2016) developed the idea of "biological randomness" that arises from the dialogue between physics and biology. The relevance of their work encourages me to state that my account of stochastic explanation (SE) is not the only attempt to emphasize that chance and probability are not synonymous, and that it is possible to develop an idea of the former without at the same time being (necessarily) involved with the latter[146]. In the second section, I discuss the work of the molecular biologist Jean-Jacques Kupiec with respect to chance in cells. This section emphasizes how biological reflection on chance is heavily focused on ontological investigation, as opposed to my work which is purely epistemological. In the third section, I show that a particular definition of chance as noise is used in two very different domains of study and that in both cases it prevents epistemic progress. The first domain is the study of drift in evolutionary biology, the second is noise in gene expression. If in the first domain this problematic aspect is well analyzed in the context of philosophy of biology, there are not many philosophical reflections with respect to noise in the second domain of gene expression. This section intends to convince the reader that a philosophical study of the notion of chance in gene expression, I will provide in the next three chapters, is therefore necessary. Through Merlin (2010, 2016)'s study, the fourth section aims to draw attention to the importance of the study of time in the context of chance in molecular and cellular biology. In

[146] An example of how chance and probability are used as synonyms see Handfield (2012).

biological practice, time is rarely made explicit. In this section, I will show that the link Merlin proposes with respect to chance and time could be an excellent starting point for the exploration of chance in cellular and molecular biology. Indeed, in Chapter 7 I show that highlighting different models of time is crucial to understanding the different ways in which stochasticity can operate at the molecular and cellular level.

1. The principle of variation and chance without probability

In this section, I focus on Montévil *et al's* (2016) article to show that there has been at least one other attempt in contemporary literature (in addition to mine) to develop a conception of chance. Montévil *et al's* (2016) "biological randomness" is not grounded in consideration of probabilities.[147] Even if they do not deny probability can be present and used *a posteriori*, they do not use probability in the development of their definition. In other words, if I understand the authors' work correctly, they propose a conceptual hierarchy. They first propose developing a notion of chance in terms of "biological randomness", and then specify that randomness could be measured with probability. I am sympathetic to this idea because I do not want to develop my own definition in terms of probability either. I argued in Chapter 3 that the notion of chance is explanatory because it is an abstractor, holding and synthesizing all the possible ways in which the *explanandum* can be realized. Of course, these "possible ways" could be then measured with probability. Let's take as an example the development of the plant cell wall already proposed in Chapter 3. In this case chance (i.e. CPC) holds and synthetizes all the possible directions in which the microfibrils can be oriented. An idea of equiprobability is clearly implied here: the microfibrils are equally likely to be oriented in any direction in space. Nonetheless, not all explanations already imply a notion of probability. Other examples of biological explanations may instead require further investigation in order to quantify chance in terms of probabilities.

Let us come back to Montévil *et al's* (2016) original concept of chance – which is, incidentally, not explicitly linked to any other of the classic debates of chance in philosophy of biology (e.g. debate of the status of evolutionary theory, difference between drift and natural selection, evolutionary chance, etc.). They start from the broad notion of "randomness" defined as "unpredictability with respect to an intended theory and measurement" (Calude and Longo, 2016, p. 1) and specify that this definition can work across sciences – it is applicable to domains such as physics, mathematics, computer sciences and biology. Their account is epistemological.

[147] Actually, Darwin can be described as another thinker who developed a conception of chance without referring to probabilities (see Morizot 2012).

This is separately confirmed by Longo, one of the authors of the article, who argues that the difference between ontology and epistemology becomes meaningless since the concept of randomness can be developed only by starting from an intended theory.[148]

"[w]e stress that randomness, for us, means *unpredictability in the intended theory* (Calude and Longo 2016). This relativizes randomness to the theory and its symmetries […]. Without an at least tentative theoretical frame, one cannot talk of unpredictability: in order to '(un-)pre-dict' one needs to try first to 'say' something ('dicere', in Latin)" (Longo 2018, p. 446; emphasis in original).

Here I think there is an analogy. Just as my notion of chance as present contingency (CPC) finds its role in explanation, Montévil *et el* – and more generally, Longo – repeatedly specify that randomness finds its definition, and role, only in the framework of an intended theory and measurement. I think that we are both interested in developing an epistemological account of chance. But the definition I mentioned is general and only a starting point from which the authors develop their idea of randomness *specifically* for biology. They characterize their account of randomness through the comparison between physics and biology:

"[i]t is worth emphasizing that, although we will elaborate on the concept of variation by analogy with and in contrast to the physico-mathematical perspective, we by no means advocate a physico-mathematical treatment of biological phenomena" (Montévil *et al* 2016, p. 38; see also Longo 2018, p. 445).

This paper and others[149] provide a theory of organism based on three theoretical principles, one of which is the principle of variation.[150] The authors first deal with physical objects. They claim that they are *generic* objects because their behavior in a phase space –the "mathematical description of the object" (Montévil *et al* 2016; p. 3, note 5) – do not change over time:

[148] I would like to point out that this definition of chance as "biological randomness" is reminiscent of the definition of "contingency relative to a given theoretical system" proposed by Gayon in his 2005 article (p. 397). Curiously, the authors do not mention this work.

[149] In addition to those already mentioned, we have: Longo and Montévil (2012a) for a discussion on symmetry; Soto *et al* (2016) for an overview of three theoretical principles of organism; Longo and Montévil (2012b) for the relation between randomness and order in biology; and Buiatti and Longo (2013) for the different levels of randomness present in biology.

[150] The other two principles are *the principle of organization* (see Mossio *et al*, 2016) and *the default state of proliferation with variation and motility* (see Soto *et al*, 2016).

"[G]eneric objects are objects which *are all of the same kind from the point of view of the theory* (they typically obey the same equations[151])" (Montévil *et al* 2016, p. 38; emphasis added).

"Generic objects are, for the most part, defined by the transformations that *preserve them*, and that enable us to define stable mathematical structures […] *symmetries* (in general) are transformations which leave the relevant aspects of an object invariant" (*ibid.*, p. 39; emphasis added).

For example, physical objects understood as physical bodies always have mass and weight and invariably follow the Newtonian principle F = ma. For this reason, Montévil *et al* also define physical objects as *invariant under transformation* – their phase space does not change over time:

"An 'invariant' is characterized by the transformations that preserve it—typically, transformations in space and time as reference systems, like in Galileo's or Einstein's relativity" (Longo 2018, p. 444).

With this concept of a general object in mind, Montévil *et al* then argue that a biological object is not general at all. Biological organisms are *specific* objects since they qualitatively change their organization over time. In particular, the symmetry of biological objects such as organisms or ecosystems[152] *changes* over time. The principle of variation concerns randomness in the precise sense of *unpredictable symmetry changes* (see also Longo 2018, p. 447). Physical objects maintain their symmetry through time and therefore they are *invariant* under transformation. By contrast, biological objects, such as organisms' ecosystems, vary their symmetry through time and therefore are *variant* under transformation. The difference between variant and invariant is at the heart of Montévil *et al's* account. From a physical-mathematical perspective, continuous symmetry change means that the phase space also changes continuously. It is necessary to keep in mind that in physics it is not possible to describe a phase space that changes continuously. Descriptions of biological objects are therefore not possible – or at least, it is not possible to describe them with the tools provided by physics.[153]

[151] Indeed, the authors stress that a phase space is not sufficient for understanding the behavior of a generic object (Montévil *et al* 2016, p. 38). An equation is necessary to understand the trajectories of generic objects through a phase space.
[152] "Ecosystems might be conceived of as biological systems [biological objects], although not necessarily as organisms" (Mossio *et al* 2016, p. 2).
[153] What makes a biological object a specific object is the fact that it is "constituted by a particular history of relevant and unpredictable symmetry changes over time" (Montévil *et al* 2016, p. 40).

To better explain this idea of chance, the authors use Galton's quincunx device as an analogy (Galton, 1894; see Figure 1).[154] In it, several balls fall one after the other, encountering numerous obstacles. Due to these obstacles, each ball changes its trajectory continuously. In the lower part of the device there are different spots into which the balls can finally fall. The device has pre-defined possibilities because, even if we do not know exactly which spot each ball will fall into, they will form a normal or Gaussian curve. The device is an analogy of the idea that in physics, the phase space is an *a priori* fixed phase space. Following the analogy, the device is the phase space and the balls, the physical objects. Even if we do not know the specific trajectory of a generic physical object, nonetheless, we have its possible results in a fixed phase space.

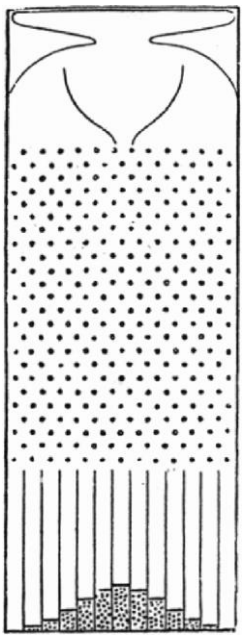


Figure 1 Galton's quincunx (by Galton, 1894, p. 63, figure 7). See main text for more detail.

Let's summarize by turning to the words of the authors:

"A ball falls but obstacles lead it to move randomly to the right or to the left. The outcome is variability in the position of the balls at the bottom of the device. This device illustrates variation in a *pre-defined set of possibilities*. Biological variation, by contrast, sometimes involves the constitution of new possibilities, which would amount for the ball to *jump outside of the quincunx*" (Montévil *et al* 2016, p. 45; emphasis added).

[154] For a discussion on Galton and statistical explanation, see Ariew, Rohwer, and Rice (2017). For a historical overview on Galton see Pence (2021). From the relation between Galton and Mendel see Radick (2011). For an investigation on probabilistic causation in developmental behavior through the example of Galton's Quincunx, see Kaplan and Turkheimer (2021).

With this, Montévil *et al* (2016) elaborate an original idea of *biological randomness*, defining it as "the very fact that biological objects undergo *unpredictable* symmetry changes" (p. 40; emphasis added).

What I want to underline is that these authors stress that biological objects' relevant symmetries and changes of symmetry are not pre-ordained (p. 40) and "can only be listed *a posteriori*, that is, after their realization" (p. 40; emphasis in original). This framework influences the way in which Montévil *et al* (2016) conceive probability:

"Biological objects are – by hypothesis – specific, but when we describe a particular change of symmetry, it is studied *a posteriori* as a generic aspect of the object, and can be added to the past possibilities of a system. Randomness is then not correctly framed by *a priori* probabilities. Probabilities, if any, are defined *a posteriori*. A specific possibility is accommodated by the space of possibilities, but this space is obtained *a posteriori* and obviously does not include all future possibilities" (p. 42; emphasis in original).

What the authors show in this passage, if I understood correctly, is that the notion of chance has a conceptual priority over probability and, more importantly, that this notion is not constructed on the basis of probability. The fact that in biology the phase space is always changing is at the heart of Montevil *et al's* notion of randomness. Probabilities can only be used to measure this change retrospectively. This is an important and original philosophical move since philosophers mostly use probability and interpretation of probability to define chance, or at least use chance and probability as interchangeable notions (e.g., see Figure 2).

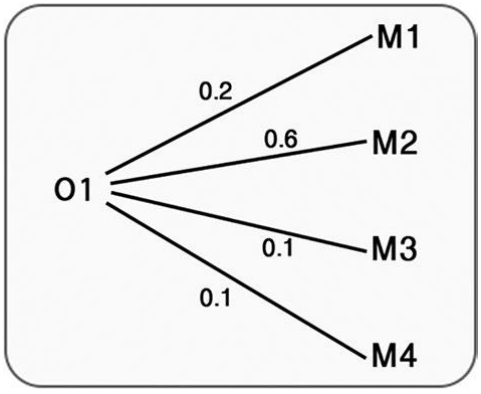


Figure 2 Wong's representation of mutational outcome array. For further detail, see the main text.

In Figure 2, we see a representation of chance concerning mutational outcomes. It depicts an initial condition (O1) that represents a given genome in which the mutation has not yet occurred. (M_x) are the possible results, that is to say the possible mutations. The numbers are the

probability of each of these (M_x). Wong (2020, p. 7) discusses chance through this representation of fixed probability distribution.[155] Millstein (2011) too, although she defines her Unified Concept of Chance (UCC) in terms of causes and not in terms of probability. Soon after, she tries to integrate that definition with probability, specifically with conditional probability:

"The translation between UCC and a (conditional) probability is as follows. The UCC corresponds to the *probability of a particular outcome given the specified subset of causes*. Or, more formally, a given instantiation of the UCC can be translated to *Pr (outcome/subset of causes)*, where 'subset of causes' are the considered causes and 'outcome' is one of the possible outcomes, for that instantiation of the UCC" (Millstein 2011, pp. 439, 440, *emphasis in original*).

I think these authors use probability to define chance but their purpose is quite different from mine. They do not aim to propose a reflection on chance with respect to biological explanation. They just want to clarify the uses of this concept in evolutionary biology, and chance in this area of study is mostly developed in terms of probability (e.g. Denny and Gaines 2002). So, in a sense these philosophers are just tracing the way chance is defined in biological practice. That's a good strategy, indeed. But it is not the only strategy. Chance can also be studied as a notion with respect to the role it plays (or should play) in some biological explanations. This is a completely different project, and the focus of the present dissertation.

But different areas of research aside, the difference also lies in a theoretical choice. My assumption is that chance and probability are not synonymous and my philosophical stance is that, in line with Montévil *et al*, chance takes conceptual priority over probability. Why? Because in biological practice we find many explanations that contain chance (cf. the three examples in Chapter 3) but that nevertheless do not contain (at least explicitly) the notion of probability. I think then it is an interesting philosophical project to explore these case studies. In fact, few philosophers have ventured to study the notion of chance without bringing in probability at some point in their reflections (for example, Montévil *et al*).

[155] In his paper, Wong also proposes Shannon's information entropy (p. 24) to quantitatively measure chance in random mutations. It is noteworthy that Shannon's information entropy measures a degree of disorder and it is not so obvious that measuring disorder means measuring chance (e.g. stochasticity). At the very least, an explicit and clear philosophical argument is needed to defend this idea.

2. Kupiec's ontology of chance

In this section, I describe in detail Kupiec's ontological account of chance in biology. The term "ontological" here refers to what is out there, that is to say chance as an inherent property of biological phenomena. Even though my own analysis is epistemological in nature, in this section, I discuss some ontological issues that stochastic processes give rise to. Is stochasticity fundamental and/or inherent to biological phenomena? Is a stochastic process reproducible? In the Introduction and Chapter 3, I suggested that providing an ontological analysis of chance might constitute only the second step in a philosophical analysis, following an initial epistemological exploration. In this section, I want to make the argument more compelling, by confronting Kupiec's ontological work on chance. My primary aim is to show an example of ontological work on chance in the exploration of biological processes. My secondary aim is to argue that ontological questions often call on epistemological issues, which testifies once again to the priority of epistemological over ontological work. Analyzing Kupiec's work is also a first step in getting to the heart of the case studies analyzed in the next three chapters, all of which will be about gene expression.

Since the 1980s, Jean-Jacques Kupiec has fervently defended of the idea that chance is at the heart of the cell (Kupiec [1983] 2020). The French biologist argues that chance variations (and selection) act at the molecular level too. In recent works, he has argued that 1) chance has an inherent and fundamental role in biological process; 2) stochastic processes are variant but reproductible processes. The first stochastic models of cell behavior date back to 1960s[156] and Kupiec claims he was one of the first biologists to propose[157] a stochastic model of gene expression (Kupiec [1983]2020, 1997). In various works, Kupiec complains that even if his idea that the expression of genes is fundamentally stochastic has found followers, contemporary biology still struggles against the tendency to privilege a deterministic idea of cellular processes – especially of those in gene expression. Indeed, Kupiec shows us that determinism has not disappeared but acquired a new form (see Duchesneau 2012). He calls this new view "determinism with noise" (Kupiec 2014b, p. 2; 2019, p. 21-23), and that it consists of adding (and admitting) "a nuisance" to an otherwise strictly deterministic process.[158] Kupiec argues

[156] Kupiec (2014b) notes that one of the first stochastic models of cell behavior is by Till *et al* (1964).
[157] "I believe I proposed in 1981 one of the first stochastic gene expression models (if not the first)" (Kupiec 2019, p. 236; my translation from the French).
[158] We can find an example of this in Corson and Siggia (2012) in which the authors propose a geometrical and mathematical model to explain and make predictions about the differentiation of the vulva of *Caenorhabditis elegans*. One of the main assumptions of this model is that the relevant factors for vulva development are the two signaling pathways that induce cell differentiation: inductive EGF signaling and Notch signaling. In the paper they do include noise in term of Brownian motion and in terms of interference between the two signals (see Corson and

against this view and for stochasticity as an *inherent* and *fundamental* element of cell processes and cells variations.

2.1 Noise vs stochastic trains

To explain the difference between "processes with noise" and his account of "intrinsically stochastic processes" (Kupiec, 2019, p. 21-23), Kupiec uses the metaphor of a train. Imagine a train has a fixed schedule going from A to B. Imagine also that the train cannot deviate from the main route and that the conductor knows all the speed needed for the different parts of journey (i.e., slowing down at critical curves, speeding up on straight parts, and so on).[159] Even if the behavior of the train (and its conductor) is well programmed and all necessary maneuvers are planned and streamlined, some degree of noise could still be present. Why? Because the behavior of the conductor will be not *exactly* the same from A to B. We are talking about a human agent that is fallible and does not follow – each time she travels from A to B – the exact same moves and maneuvers. In addition, adverse weather conditions, unexpected events, accidents, damage to the train and so on might require changes to the train's speed. Therefore, "although [the train] follows a strict program, it never carries it out in a rigorously identical way. The functioning of the train can thus be described as 'deterministic with noise'" (Kupiec 2019, p. 21; emphasis added).[160] This metaphor is Kupiec's way of making the point that a determinist account with noise is nothing more than a disguised version of an approach that is determinist anyway.

Now imagine a train, continues Kupiec, in which the changes of speed "would be chosen at random thanks to a random number generator" (p. 23)[161] embedded in the engine of the locomotive. In this case, stochasticity in the speed of the train is not due to external fluctuations (e.g. weather conditions, conductor behavior, and so on) but stems from the train's very engine. The difference between the two trains resides in the source of stochasticity. In the deterministic train with noise, the speed variation can change and this change is caused exclusively by external disturbances. If there were no such perturbations, the train would behave in a deterministic and reproducible way (going from A to B it will invariably have the same speed

Siggia 2012, p. 5570, equation 2, the parameter "$\eta^{(t)}$"). But it is just an additional a parameter as an error term to an otherwise quite deterministic model.

[159] In reporting this example from Kupiec (2019), I took the liberty of adding some details to help properly understand the metaphor under attention.

[160] The English translation is mine. The French original reads: "[b]ien qu'il suive un programme strict, il ne le réalise jamais de manière rigoureusement identique. Le fonctionnement du train peut ainsi être qualifié de '*déterministe avec bruit*'" (Kupiec 2019, p. 21; emphasis added).

[161] "[La vitesse] serait choisie au hasard grâce à un générateur de nombres aléatoires" (p. 23).

and perform the same maneuvers). This is why the source of this kind of stochasticity is called *noise*. By contrast, in the stochastic train the source of stochasticity is integral to the train itself. In this case, it is impossible to say that variation is due to external noise because the random number generator is part of the engine. By analogy, Kupiec argues that stochasticity is *intrinsic* to biological systems. Like with the stochastic train, their source of stochasticity does not come from outside, but it is inherent to the very biological dynamics.

At this point, I would like to address a criticism to Kupiec. The metaphor he uses is indeed a good heuristic tool for understanding his position regarding the fact that in biological systems stochasticity is intrinsic. But at the same time, this metaphor does not help with understanding the other half of the argument. It does not show how stochasticity can be fundamental to the train and therefore for biological systems. In the case of the train, the stochastic engine tends to change the speed randomly, but this does not change the fact that the train is predestined to go from A to B. In this case, stochasticity is not fundamental to getting from A to B: the train will go regardless of the random speeds it will assume during the journey. A piece is missing here. To try to complete the picture I would like to add another element to the metaphor. I propose visualizing the stochastic train with a magical addition: the tracks are not present, as Kupiec – following the real-life scenario – suggested, but they appear *during* the journey. Pushing the metaphor further, this then means that the direction of the train is stochastic, and that the final destination X of the train (which is not necessarily B) *depends* on these stochastic directions. This allows us to assert that stochasticity is an inherent *and* fundamental property for the realization of the journey. Stochastic directionality brings the train to destination X.

This criticism, however, may not be entirely appropriate since Kupiec and I may have different biological processes in mind when we think about the stochastic train. Kupiec agrees that at the genome level, the activation of different genes can be stochastic. However, when he talks about the activation of a single gene (i.e. when referring to the processes that led to the synthesis of proteins through transcription and translation) he maintains the idea that a gene always gives the same proteins through the same processes. This might account for the fact that what he is interested in is only the stochastic train speed/stochastic expression of different genes and not in the magic tracks which lead to different destinations/the expression of a single gene through stochastic processes which can give rise different proteins. This second scenario will be the focus of my interest in the following chapters. Kupiec would be right to talk about the train that, although stochastic, always goes from A to B. Translated into biological terms: although in the process of cell differentiation there are many stochastic dynamics (leading to

the death of many cells, see after) the formation of tissues follows similar *patterns*. In other words, in development, we will always have the formation of the same tissues or, more precisely, embryonic leaflets (e.g. mesoderm, ectoderm, endoderm). This would justify his choice of using a metaphor with real trains (i.e. with fixed rails that will always start at A and arrive at B).

So, perhaps instead of criticizing Kupiec, what I have just suggested is rather a way to broaden his frame of reference. The stochastic train with the rails appearing as the train progresses might be a good metaphor for the processes I will be dealing with in the chapters that follow. There, in fact, I will analyze the processes of gene expression of single cells starting from a single sequence of DNA and/or a single RNA. Specifically, I will focus on transcription in Chapter 5, alternative slicing in Chapter 6 and translation in Chapter 7. In contrast to Kupiec, what I therefore want to emphasize is that, even when talking about the expression of a single gene in a single cell, we can explain processes with the notion of stochasticity. The train without rails here is appropriate. Even though it starts from point A/gene A, it is impossible to tell where the train will end up/know what or which protein or structural RNA will be produced.

2.2 Kupiec's argument: probabilistic phenomena are both variable and reproducible

After having highlighted the difference between my project and Kupiec's, let's continue our exploration of his work. My ultimate goal here is to emphasize how ontological questions often lead to epistemological issues. I agree with Kupiec when he writes that "while a metaphor can be useful in illustrating a theory, it can also be a source of confusion" (2019, p. 230).[162] So, to avoid overstretching the metaphor of the train, let us leave it aside and focus on a specific biological process: cell differentiation. Can we say that this is at the same time an inherently, fundamental and reproducible stochastic process? In the words of Kupiec: "how could a phenomenon as precise and reproducible as embryonic development be the result of random events?" (Kupiec 2019, p. 236; see also Kupiec 2014a, p. 165).[163] Faulty assumptions are present here, which prevent the thought that embryonic development is *in fact* due primarily to stochastic phenomena. To switch perspectives, following Kupiec, we have to start by breaking these incorrect assumptions:

[162]"[S]i une métaphore peut être utile pour illustrer une théorie, elle peut aussi être une source de confusion" (2019, p. 230).
[163]"Comment un phénomène aussi précis et reproductible que le développement embryonnaire pourrait-il résulter d'événements aléatoires ?" (Kupiec 2019, p. 236).

1. *Embryonic development is invariant*. This assumption is wrong since is biased by "an a priori will to find invariant stages, which leads to ignoring its variations" (Kupiec 2019, p. 237).[164] In fact, in developmental processes, variability is the norm. Kupiec justifies this statement, saying that numerous cells *die* during development. This fact is rarely underlined in literature: "when we think of, or look at, cell differentiation, we usually only consider cases that succeed, so that it seems an identical process. *But if we consider the cases when embryonic mortality occurs*, we must consider that this process is not always the same" (Kupiec [1983]2020, p. 85; emphasis added; see also p. 82). Development, then, is usually conceived as invariant since we do not usually consider the variations that actually occurred. Kupiec says that the death of cells proves this variation: cell death shows that development tries all (the) possible pathways and many fail while others are carried on. Development then is not invariant, but a variable process that progresses by "learning" from its wrong paths.
2. *A probabilistic phenomenon is not reproducible*. This is another wrong assumption that Kupiec refutes. "When the number of random events of a phenomenon is large, its variance is negligible" (Kupiec 2019, p. 237).[165] And since its variance is negligible, he continues, such phenomena can be said to be reproducible. As it is clear from the last passage, his argument is based on the law of large numbers. "The law of large numbers […] states that if you repeat an experiment independently a large number of times and average the result, what you obtain should be close to the expected value" (Wikipedia). This law is typically used in physics to the extent that the number of atoms (for example) is so large that at the macro level we do see their collisions, the empty space between them, etc., but only the macroscopic properties. For example, even if a table is made of billions of atoms in thermal agitation in a vacuum, what we see at the macroscopic level is solid and compact wood. Historically, Schrödinger's *What is life?* (2013[1947]) had an enormous impact, clearly establishing that the law of large numbers only applies in physics and *not* in biology. The book sought to justify the order we find in biology from perspective of physics (without, however, reducing the former to the latter). At the end of the second world war, in light of what the study of atomic energy has caused – namely, the atomic bomb – physicists were looking for new ideas and stimuli for study

[164] "[U]ne volonté *a priori* de trouver des stades invariants, ce qui conduit à ignorer ses variations" (Kupiec 2019, p. 237; emphasis in original).
[165] "Lorsque le nombre d'événements aléatoires d'un phénomène est grand, sa variance est négligeable" (Kupiec 2019, p. 237).

to move away from that psychological trauma (see Morange 2000). Schrödinger was among these physicists who began to take an interest in biology. How is order possible in biological systems, he asked? His argument goes that in physics, order is usually justified by the fact that the number of molecules at a molecular level is so large that the law of large numbers can be applied and the macroscopic level averages out (e.g. a table of wood). On this basis he develops the principle that, in physics, "order comes from disorder". But in biology, he continues, we do not have a sufficiently large number of biomolecules inside the cell and so the law of large numbers cannot be applied. In light of this, he attempts to find other explanations, and it is in this context that he theorizes an aperiodic crystal (notably proposed in advance of discovery of DNA's double helix in 1953 by Franklin, Watson, and Crick) that contains all the information necessary for an organic system to organize and live[166]. In biology therefore, the guiding principle is that "order comes from order", or in other words, order comes from the information present in the geometry of the aperiodic crystal. This principle influenced several generations of biologists in the 50s, 60s and 70s. Kupiec adds that the idea of order by Schrödinger indirectly reinforces the deterministic idea of biological process. If order comes from order, everything that is written in the aperiodic crystal is executed with no margin of error and therefore in a deterministic way. Schrödinger's idea was of course revolutionary at the time, but Kupiec and Nicholson (forthcoming) write that it has caused a major delay in the study of stochasticity in the cell. Specifically, ideas in Schrödinger's book have prevented biologists from studying stochastic cellular dynamics to which, in fact, the law of large numbers can indeed be applied. It is not the aperiodic crystal (today, DNA) that dictates the instructions to the cell, but rather stochastic processes that create "order from disorder". For example, the differentiation of a single cell is a stochastic process (more precisely it is a stochastic process followed by stabilization, see next section). But if we analyze the cellular differentiation of many cells, we can observe a reproducible and regular process. Cellular differentiation is a probabilistic phenomenon but in light of the fact that "[c]ellular phenomena rely on a very large number of stochastic molecular interactions" (p. 238)[167] at the macrolevel (e.g. tissue formation), we see a reproducible phenomenon. In each embryo (of the same

[166] Delbruck (1971) provocatively argues that the true discoverer of DNA was Aristotle instead. He writes: Aristotle's "'unmoved mover' perfectly describes DNA: it acts, creates form and development, and is not changed in the process" (p. 55).

[167] "Les phénomènes cellulaires reposent sur un très grand nombre d'interactions moléculaires stochastiques" (p. 238).

species), the cell differentiation occurs in the same way. Even in biology, then, the principle "order comes from disorders" can be applied.

The value of the arguments developed in these last two points resides in the fact that Kupiec makes variation and reproducibility compatible with stochasticity. Cell differentiation is variant but, since it follows the law of large numbers, it can be conceived as a stochastic *and* reproducible process. I agree that the law of large numbers can be applied to biology too, but I have one last concern. Kupiec writes about the law of large numbers in an abstract manner, without addressing the problem of how it can be applied. The law of large numbers comes up several times in Kupiec's arguments.[168] But still, in which scenarios this law can be applied? It is *actually* reasonable to call this law into account in the context of cell differentiation? And if so, is it possible to be more specific about the elements that are associated with it? To which biological items do large numbers refer? Molecules? Processes? Events? All of them? How large does their number need to be in order to be understood as "large numbers"?

2.2 Models for cell differentiation: Inductive Instructive Model (IIM) vs Selection Stabilization Model (SSM)

The final thing that needs to be specified in this brief journey through Kupiec's work is to understand more deeply in what sense, according to him, stochasticity is fundamental for biological systems. In order to do that, we need to dive briefly into the history of modeling cellular differentiation. Spemann (1938) was the first to propose a model of cell differentiation based on induction (Kupiec 2014a). The general idea beyond this model is that the differentiation of one cell depends on – and is caused by – the signal sent by another cell. In other words, this information *induces* cell differentiation. The notion of induction helps greatly in developing a determinist conception of cell differentiation: if cell (a) sends a signal *d* to a non-differentiated cell (b), this latter differentiates into cell (d) (see Figure 3).[169] The reasoning behind this conceptualization is that if we have A, we then have B in a deterministic way (i.e. that is, all the times that I have A I *will necessarily* have also B). Nonetheless, Kupiec (2014a) notes that this idea of "induction = determinism" was developed only after Spemann. It was the rise of molecular biology (1940s – 1960s; see Chapter 5) that consolidated this identity relation. Indeed, the notion of induction is soon reabsorbed into the theoretical framework of this

[168] See also Kupiec (2019, p. 237 - 238) and Kupiec (2014a, p. 165).
[169] The other case depicted in Figure 3 is that cell (b) gives the instruction in the form of a signal *c* to cell (a). Cell (a), receiving this signal, differentiates from cell (c).

molecular domain, which speaks mainly in terms of information. Induction is therefore an *instruction* that a cell conveys to another in the form of *information*. This idea of induction-as-information is the base of the (classical) Inductive Instructive Model of cell differentiation (often called IIM; see, e.g., Kupiec 2019, p. 198; Kupiec 2014a, p. 162).

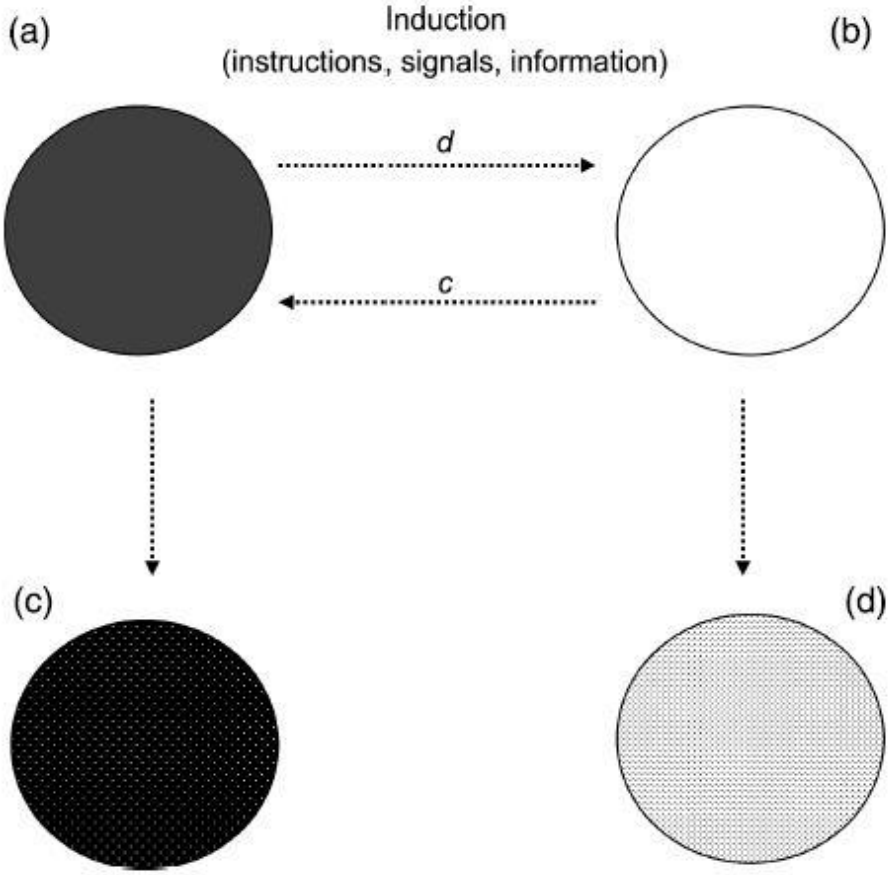


Figure 3 Inductive instructive model (IIM; image by Kupiec 2014a, p. 161). For further detail, see main text.

Kupiec is critical of IIM and underlines two main problems:

1. On the one hand, this model does not fit the empirical data[170] provided by the scientific community concerning cell differentiation. For this reason, the author conceives of it as inappropriate. Kupiec insists that this inadequacy is due to the fact that IIM does not take seriously the inherent and fundamental stochasticity of biological processes (see Kupiec 2014a).

[170] Kupiec (2014a) gives an overview of these dates on page 157 in the section, "demonstration of stochastic gene expression".

2. On the other hand, in Figure 3 we can see that cells (a) and (b) are of different colors (black and white), meaning that they are already different (in some fashion). This implies that the model attempts to explain cell differentiation starting from an initial situation where the cells already show difference. A paradox exists here which Kupiec does not miss the opportunity to underscore: "the IIM does not explain how different cell lines are produced from a single cell […]. *The model therefore presumes a diversity of cells, the appearance of which it is supposed to explain*" (Kupiec 2014a, p. 161; emphasis added)[171].

Advocates of IIM attempt to justify the above two points by saying that a morphogenetic gradient pre-exists in the egg and that this can explain (and causes) the initial asymmetry of cells. But this hypothesis has never been confirmed by data. In fact, the biggest stumbling block for IIM is the fact that it is problematic to keep believing it in light of the fact that it does not fit with data (for further detail see also Kupiec [1983]2020, p. 81 and Kupiec 1997, p; 201). This insurmountable problem leads Kupiec to conceive an alternative model.

Kupiec's reflection culminates in the development of an alternative model he calls the Selection Stabilization Model (SSM) ([1983]2020). This model aims to fill the exact gap left by the IIM: by discarding the determinist idea of cell differentiation, it takes seriously the fact that biological processes are fundamentally and inherently stochastic.

[171] We can find a similar critique in Kupiec (2019): "[t]he [IIM] model presupposes cellular heterogeneity which must be explained" (p. 194).

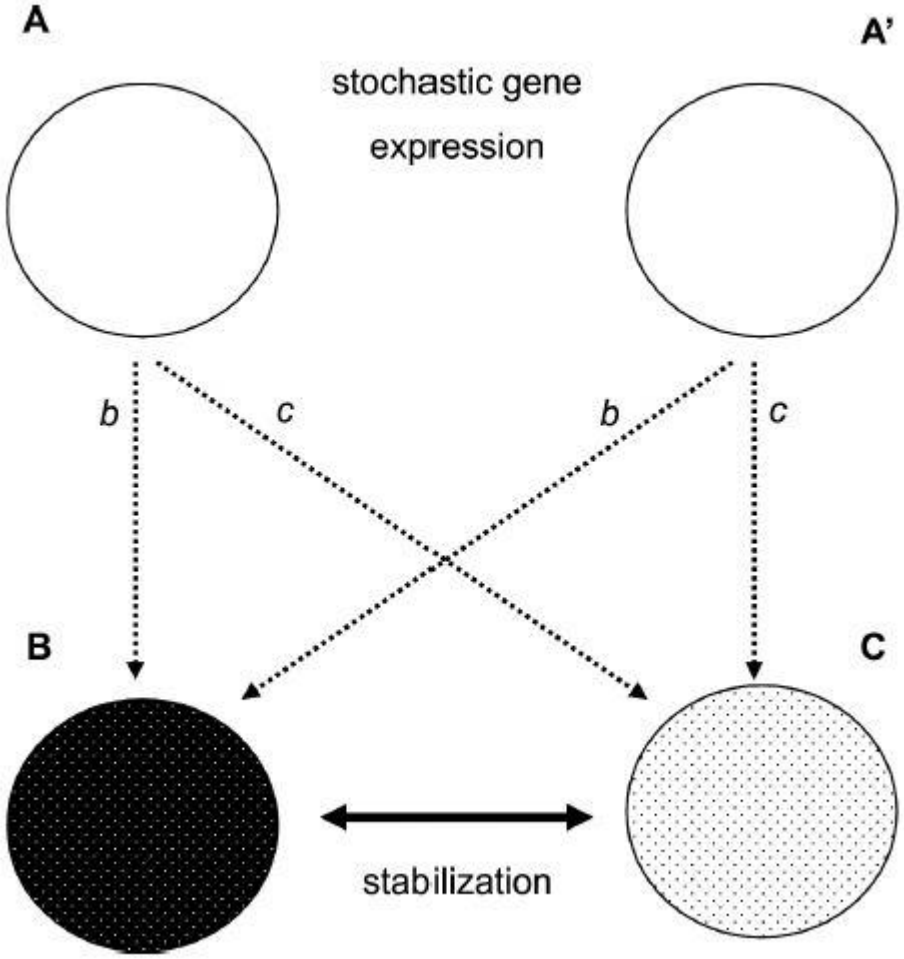


Figure 4 Selection stabilization model (SSM; image by Kupiec 2014a, p. 162). For further detail, see main text.

In the SSM model depicted in Figure 4, we start from symmetrical, undifferentiated cells (A and A') in which gene b or c may be randomly activated. What does it mean that these genes are randomly activated? According to Kupiec (1997), the activation of the genes depends on two main factors: 1) the distance between genes and transcriptional regulators[172] (the closer these genes and regulators are, the higher the chance they interact with each other); 2) the bio-physical affinity between genes and transcriptional regulators (the higher the affinity, the higher the probability that the two kind of molecules will interact). Because of these two factors, gene activation is probabilistic rather than deterministic (the probability of their activation depends on the two variables of distance and affinity).

The SSM model does not only explain the activation of genes via their probabilistic nature. It goes further and postulates that after this random activation, stabilization follows. Referring

[172] Transcriptional regulators are a specific class of biomolecule that control whether a particular gene will be transcribed into mRNA (Wikipedia).

to Figure 4, once genes have been stochastically activated and cells have been differentiated, stabilization occurs through interaction between these differentiated cells (i.e. B and C). We can summarize Kupiec's SSM model by saying that cell differentiation consists of "stochastic variations followed by stabilization". The variation is due to stochastic gene expression and stabilization is due to the interaction between differentiated cells. For Kupiec, this type of model has two very important theoretical and practical advantages:

1. The SSM model considers the fact that biological processes are inherently and fundamentally stochastic. In fact, it frames gene activation not as a deterministic induction but as a stochastic event. Kupiec writes that this model fits with empirical data developed by scientific practice (see Kupiec 2014a, p. 157).
2. This model solves the problem of asymmetry in IIM. In fact, from Figure 4 we can see that both initial cells A and A' are white, that is to say identical. Therefore, the model is effective because it explains cell differentiation starting from two actually undifferentiated cells.

On the basis of what we have said so far it is reasonable to state that the stochastic model has undoubted advantages over the deterministic one (i.e. it fits the data and resolves the problem of asymmetric cell differentiation). I can therefore say I align with the SSM model. However, using this new model brings new issues.

2.3 Ontological work leads to epistemological questions

In Figure 4, the combination of B and C enables cells to stabilize their gene expression and, more generally, to differentiate respectively into phenotype B and C. But if cells A and A' had differentiated into X and Z cells, would stabilization still have taken place? More generally, how can we know that stabilization occurs for B and C cells and maybe not for X and Z cells? Or that stabilization would have occurred for all possible combinations BC, XZ, BX, BZ, CX, etc.? Or that it might not occur for any of these combinations? In other words, what is it that makes stabilization possible for two differentiated cells? Furthermore, what happens when two differentiated cells fail to stabilize? Do they die? Indeed, in this model there seems to be an implicit assumption that *only some stabilization takes place*. What Figure 4 does not tell us, and, more generally, what the SSM does not tell us, is what "stabilization" refers to and what the conditions that permit stabilization actually are. Even Kupiec recognizes this as a weak

point in his model. He eventually tries to circumscribe the problem with the hope that future work will be able to explain in more detail what happens during cellular stabilization:

"In the future, the probabilistic theory of cell differentiation presented *here should be complemented by a theory of cell interactions*, in order to account for the whole process of embryonic development" (Kupiec [1983]2020, p. 85; emphasis added).

In light of this problem, even if he trusts to future work, he himself proposes possible solutions. In his 1997 work, Kupiec hypothesizes that the stabilization between cells is due to the interaction between their membrane proteins. These proteins are one of the products of cellular differentiation. This means that one property of these differentiated cells is that they have developed specific membrane proteins (though Kupiec himself is not clear on this point). Applying his 1997 argument to Figure 4, we can say that cell B synthetizes membrane B proteins; these B proteins interact with C membrane proteins which are synthetized by cell C. The interaction between B and C proteins "triggers the stabilization of gene expression" (Kupiec 1997, p. 204). By contrast, if we had differentiation in two C cells, it could be speculated that C membrane proteins would not interact at all or that their interaction would not result in stabilization. This interaction is therefore species-specific. Not all proteins from different cells can interact.

Kupiec goes even further and attempts to reason why this interaction between C and B proteins actually gives cells stabilization. He does this by taking into account the phenomenon of chromatin modification.[173] With this addition, the scenario becomes very complex and far more difficult to manage. He writes that the interaction between B and C proteins causes a phosphorylation cascade that triggers the modification of chromatin which, in turn, causes the stabilization of gene expression. Kupiec feels the complexity of the picture he is proposing and quotes his colleague Paldi in an attempt to make his proposition more compelling. Indeed, Paldi (2003, 2012, 2018, 2020) elaborates further on the relation between cell differentiation and epigenetic chromatin modifications. To better understand Paldi's work it is first of all necessary to explain epigenetic modifications. Epigenetic modifications are a kind of modification that

[173] Present only in eukaryotic cells, chromatin is composed by proteins and DNA. Proteins called histones bind and pack up DNA, forming the basic structure of nucleosomes. A nucleosome is a segment of DNA wound around eight histone proteins and resembles thread wrapped around a spool (Wikipedia). When the interaction between DNA and histones are weak, chromatin spontaneously fluctuates between accessible and inaccessible states. In this case, there are more opportunities for transcriptional initiation and so for gene expression. This state of chromatin is called euchromatin. When the interactions between DNA and histones are strong, chromatin remains packed longer, and the probability for transcriptional factors to access DNA decreases. This state of chromatin is called heterochromatin.

does not change the DNA sequence but instead affect the state of chromatin and, as a consequence, impact the regulation of gene expression. By affecting the state of chromatin, which can be heterochromatic or euchromatic,[174] they actually influence the possibility of access to DNA by transcription factors and DNA polymerases. Very roughly, we can then say that the interaction between B and C membrane proteins (cf. Figure 4) triggers a signal that influences chromatin epigenetic modifications. This kind of modification create a specific balance between heterochromatin and euchromatin states that allows only certain genes to be expressed, and globally, for gene expression of the cell to stabilize.

Can the work by Kupiec and Paldi be integrated in a unique theoretical picture? Despite the fact that they cite each other, at the end of the day, a satisfying synthesis between their models or, at least, between their frameworks, is lacking. For example, Paldi (2003, 2020) writes about the fundamental role of metabolism in epigenetic chromatin modifications, and, then, in cell differentiation. Kupiec (1997) writes that "in addition to membrane interactions […], *metabolism* could also be involved in the stabilization of gene expression" (p. 205; emphasis added). Nonetheless, Paldi does not enter into detail of cell differentiation, and Kupiec does not enter into detail on the role of metabolism in cell differentiation. Philosophical analysis of a possible compatibility and dialogue between these two accounts is lacking. It is not my task here to build this theoretical bridge, but it is important to underline this issue for one specific reason. This critique of Kupiec serves to show that some philosophical questions end up falling more on epistemological issues, such as the need to create a common conceptual framework between Kupiec's and Paldi's work. This conceptual gap in the work of the two biologists testifies well to how an epistemological elaboration is always necessary in discussion of ontological dynamics. This further substantiates my idea of epistemology's priority over ontology (cf. Introduction, sections 1.3, 1.4). In other words, it is often necessary to understand how scientists work in order to give a coherent picture of the phenomena they study.

3. Random drift

In this section, I briefly present the philosophical debate on random drift, a biological concept that is often associated with chancy phenomena (see Millstein 2016). Specifically, I will focus on one of its specific definitions which considers drift as a deviation from an expected

[174] Chromatin can occur in two states. The euchromatin state is a loose state in which DNA can be joined by transcription factors and RNA polymerase, and gene expression can take place. By contrast, heterochromatin is a state of DNA that is tightly packed (or condensed) with histone proteins and cannot be reached by transcription factors and RNA polymerase, and gene expression cannot take place.

result. More specifically, if natural selection is meant to be the main process in evolution, drift is often conceived as merely (random) deviation from the expected result (in terms of changes to gene frequencies) under the hypothesis of evolution by natural selection. The point I would like to bring to the reader's attention in this section is that we can find a similar conceptualization – which describes a property of very different kind of processes – in cellular and molecular biology. Chance is described as noise in terms of deviation/error with respect to the average or expected result of the process of gene expression (i.e. the amount of proteins produced). In this section, I will argue that: 1) if drift and chance-as-molecular-noise are conceptualized in terms of "deviation from a mean" (i.e. in terms of result vs. in terms of process), both suffer from conceptual and epistemological impoverishment and cannot be explanatory with respect to the target phenomena; 2) it is time to rethink both of these concepts in order to take seriously their potential explanatory power with respect to certain phenomena. If, on the one hand, many philosophical efforts have been made to rethink the notion of drift (e.g. Beatty 1992; Millstein 1996, 2002, 2016), on the other, few rethinking chance at the level of gene expression have been forthcoming.[175] Bearing in mind the account of stochastic explanation (SE) I proposed in Chapter 3, to fill this gap I clear the way for my proposition in Chapter 5. There, I propose an augmented epistemology of chance at the molecular and cellular level that aims to integrate the possibility that chance has an explanatory role. This section can therefore be understood as the seeding bed for the ideas I will grow and elaborate in the following two chapters.

3.2 Brandon and Carson's (1996) conception of drift

Brandon and Carson (1996) state that if evolutionary phenomena were only due to natural selection, then we should conceive of evolutionary theory as deterministic.[176] But because of drift, evolutionary theory is indeterministic/stochastic. In general, drift is stronger when the sample of populations is small (i.e. genetic diversity can randomly and more significantly vary across generations). Conversely, the larger the population, the more the drift effect decreases. This is simply a statistical fact: the more balls I take from an urn, the more representative the result will be with respect to the type of balls that are in it (e.g. a certain relation between red and black balls). If I take fewer balls, the result will be less representative and the sample will *deviate more* from the proportions inside the urn. The assumption behind

[175] With the exception of Bravi and Longo (2015) who provide a deep analysis of the functionality of randomness in biological systems.
[176] With this statement, the authors imply the assumption that natural selection is a deterministic phenomenon.

Brandon and Carson's statement is that if the biological population were infinite, then drift would not take place (and we would therefore only have natural selection). But in a scenario with real populations, "with[in] finite populations selection does not eliminate the drift-effects of sampling error" (Brandon and Carson 1996, p. 325). They reduce drift to a deviation from the expected reproductive success (i.e. deviation from an expected result). The authors make this idea explicit, quoting (and agreeing with) Roughgarden who describes drift with the language of information theory:

"[T]he gene frequency, *p*, is a 'signal' that must be transmitted from one generation's gene pool[177] to the next. The effect of sampling *error is to introduce 'noise'* into the communication channel. Because of this *noise*, the signal that is received fluctuates from generation to generation" (Roughgarden 1979, p. 58 quoted in Brandon and Carson 1996, p. 323; emphasis added).

The sampling error produced by drift "introduce(s) noise into the communication channel". Drift (i.e. noise) is conceived as deviations (i.e. fluctuations) from the expected result (i.e. the signal).

But this way of conceiving drift has been strongly contested. Millstein (2002) criticizes Brandon and Carson's definition of drift:

"[For Brandon and Carson 1996], [n]atural selection is a process whereby organisms achieve their expected reproductive success (based on their fitness in a given environment), random drift is a process which probabilistically *deviates the population from those expectations*" (p. 48; emphasis added).

Millstein explores the implications of conceiving of drift as a simple deviation from the expected results in a natural selection scenario. She notes that conceiving of drift as noise, as Brandon and Carson do, leads to an *impoverishment* of the concept itself. This is due to the fact that

"[n]atural selection is stripped of its probabilistic character altogether, while random drift becomes a mathematical deviation from the expectations of natural selection (*rather than seeing it as a process in its own right*)" (Millstein 2002, p. 51; emphasis added).

[177] The *gene pool* corresponds to the genetic richness (i.e. the set of all genes) of a certain population.

In other words, she underlines the fact that reducing noise to a mathematical deviation rules out the possibility of (1) conceiving of drift in terms of the definitions mentioned (or any other processes) and (2) conceptually distinguishing it from natural selection. By contrast, Millstein stresses that drift is not (only) a mathematical concept but a physical one that describes biological processes (Millstein 1996, p. 32).[178]

I am in line with Millstein on this critique and I even want to push it further. Brandon and Carson's reduction is not only a conceptual pauperization, it can also be dangerous insofar as it would disregard other phenomena which might be described as drift. Indeed, since drift is conceived as noise, that is to say as a mathematical deviation from a mean, no further analysis on the biological processes of interest is needed. This notion of drift-as-noise by Brandon and Carson[179] interests me because it is curiously similar to a particular characterization of chance at the molecular level, specifically with respect to gene expression. In cellular and molecular biology, chance is described as noise, and is often defined and quantified in terms of random variation around the mean value of gene expression (i.e., the average amount of RNAs or polypeptides produced; Chapter 5, section 1.2). As both drift and chance are conceptually reduced to a deviation from an expected value, the term "noise" appears in both cases. This similarity between noise in gene expression and Brandon and Carson's 1996 conception of drift as noise is quite evident and, in my view, quite remarkable. This is because it allows me to extend criticism of drift to this other characterization of chance as noise at the molecular level. In Chapter 5, I will analyze at length the conceptualization of chance. Since in this case it is also understood as noise, it will reveal problems similar to those outlined in this section with respect to drift. The purpose will then be to try to go beyond these issues and rethink chance as an active element in explanation.

[178] More specifically, her main idea is that drift and selection, as processes, can be distinct concepts. Otherwise, as outcomes, they are hardly discernible. Thus, focusing on processes: if the physical difference between organisms is causally relevant to the difference in their reproductive success, natural selection takes place. Otherwise: if the physical differences between organisms is causally irrelevant to the differences between their reproductive success, then their differential survival and reproduction are due to a drift process. (see Millstein, 2002, p. 18, the snail example). This process associated to drift is called an indiscriminate sampling process (Beatty 1984, 1992; 2006a; Dodson and Dodson 1985; Millstein 1996, 1997, 2016).

[179] This is not the only definition of drift developed by Brandon. With the collaboration of other colleagues (Brandon 2006; Brandon and McShea 2020; McShea and Brandon 2010) Brandon develops the principle of drift or biology's first law. Very roughly, if there are no acting forces, a population tends to drift from the equilibrium (Brandon 2002, p. 328). In their 2010 book, McShea and Brandon make this conception even more precisely, calling it ZFEL (zero-force evolutionary law): "In any evolutionary system in which there is variation and heredity, in the absence of natural selection, other forces, and constraints acting on diversity or complexity, diversity and complexity will increase on average" (p. 3). Even if remarkable, this definition is not essential for my argument.

4. Random mutations

In this fourth section, I discuss in more depth Merlin's work (2009; 2010; 2016a) on what it means for genetic mutations to be "random" at the level of the origin of their molecular processes. In developing her account of "weak randomness" Merlin considers the dimension of time, which is rarely considered by either biologists or philosophers when discussing cell processes described as stochastic. Merlin's work, and especially her idea of making the temporal dimension one of the two main pillars of her definition of "weak randomness" (the other pillar is the dimension of space), produces a very relevant theoretical movement as molecular biologists rarely integrate time in their model. By opening any textbook it's easy to see that the sophisticated images describing cellular processes with transcription factors, enzymes, proteins, DNA polymerase etc., are, most of the time, a-temporal. The temporal dimension is not represented, or at least, not explicitly discussed (see Théry 2015). Reintroducing time to molecular biology – and more specifically to reflections on stochasticity – could help with understanding the explanatory power of this property that is often attributed to cellular and molecular dynamics – or at least this is what inspired me in the development of Chapter 7. Indeed, in that chapter, I will try to show that the ways in which stochasticity can operate in gene expression – and in particular in translation, a stage of gene expression in which stochasticity is rarely discussed) – can be better understood when framed by taking into account the timing of biological cellular processes. But let's start by looking more closely at the specific details of Merlin's philosophical project that inspired me.

By giving a rich source of empirical, historical and theoretical arguments, Merlin states that the most appropriate way to characterize mutations is by using the notion of "weak randomness". Very roughly, she argues that the chemical-physical process at the origin of mutations can be conceived as a *discriminate sampling processes variant over time.* Using the notion of "discriminate sampling" she refers to a process where differences in the physical-chemical characteristics of the nucleotides play a causal role in the production of mutations and in which the probability of a mutation occurring at a particular site is not the same at each moment of time under consideration (Merlin 2016a, p. 181). Merlin mentions that there is now copious evidence for *hotspots* (Moxon *et al* 1994, 2006) – specific parts of the DNA sequence in the genome where there is a high probability of mutations occurring (see also Wong 2020, p. 35) as compared to other parts in which there is, by contrast, low probability. Concerning the meaning of "variant over time" in her definition, she underlines that there is convincing evidence to suggest that the rates of these mutations are not the same in all stages of a cell's life

cycle. Pointing out that the probability of mutation varies in space *as well as time* greatly enriches the way we think about these phenomena, at the origin of genetic mutations.

One of Merlin's original insights, or at least, the one which will inspire me in Chapter 7, is the fact that it makes time explicit when talking about random mutation. In biological discussions, it is rare to find deep reflections on time, space, cell, chance, randomness or/and stochasticity. Why? The reasons can be found in the general and still lively (more or less explicitly) commitment to determinism in biological thinking. This is the conception that molecular and cellular processes follow a predetermined program inscribed in the DNA, which functions as an instruction manual that aims to orchestrate the growth, functionality, and, more generally, the life of cells. In this context, it is no longer necessary to discuss time because if we have the certainty that a particular cause will always be followed by a particular effect, time becomes only "a framework within which the 'causes' operate" (Gottlieb 1993). The problem of describing cellular and molecular processes in a deterministic way has been widely discussed by philosophers. Philip Kitcher memorably characterized the fight against genetic determinism (Kitcher 2001) as "battling the undead". Jamieson and Radick (2017) pick up the same theme: "[genetic determinism] should have expired long ago, and yet, in small ways and larger, in the public as well as backstage cultures of science and medicine, it remains not just resilient but resistant to challenge" (Jamieson and Radick, 2017 p. 1). In recent years, certain biologists have also underlined the conceptual disadvantage that genetic determinism can lead to in biological exploration. In this sense, the following insight by Paldi – a biologist – is particularly relevant in this context of talking about time: "[d]eterministic explanations not only ignore the time scale of the biological processes but also have difficulties to incorporate it in the explanatory scheme." (Paldi 2020, p. 71). We see here clearly the fact that a deterministic approach does not allow the introduction of time into the explanation of molecular dynamics, since deterministic frameworks flatten time over causal events. In studying the process at the origin of random mutation, Merlin is one of a few philosophers of biology making explicit the role of time and space at the cellular scale. This is the work that will inspire me in a reflection that aims to admit the temporal dimension to the explanation of stochastic dynamics in translation (see Chapter 7).

Conclusion

The purpose of this "interlude" chapter has been to open a space of reflection and to rework conceptions that have already been established concerning issues relevant to my account of chance in biology (sections 1 and 2). Additionally, it was intended to open up discussion – and to stimulate the reader's curiosity – with respect to what will follow (sections 3 and 4).

In the first section, I highlighted an account that elaborates its conception of chance without using probability as an essential defining element. Montévil *et al*'s (2016) "biological randomness" is developed at the interface of physics and biology and is based on the idea that biological objects are variant under transformation. They write that if probabilities must be taken into account, they will be considered *a posteriori*; the concept is developed regardless of these (*a posteriori*) probabilities. This aspect of Montévil *et al*'s account is very similar to the way my own account of stochastic explanation (SE) is related to probability. In fact, SE is not explanatory in virtue of the fact that it contains probability, but thanks to the notion of chance as present contingency (CPC) (see Chapter 3, section 2.1).

In the second section, Kupiec's work allowed us to see how an ontological investigation of chance is articulated in biology. The differences with my work are clear – such as the fact that Kupiec tries to argue that stochasticity is an intrinsic and fundamental feature of living systems, whereas I focus on understanding its role in biological explanations. Analyzing this work was also helpful in highlighting how ontological work often raises important epistemological questions, such as the need to develop a common conceptual framework between Kupiec's and Paldi's work. This can be seen as further evidence that epistemological work takes priority over ontological investigation.

In the third section, I proposed a comparative analysis between the notion of drift and the notion of molecular noise. I pointed out how some conceptions of these two notions are curiously similar. I also mentioned that, while a rich work rethinking the conception of drift has already been proposed, for chance-as-noise at the molecular level virtually no philosophical reflection yet exists. This demands a philosophical work rethinking chance in this biological field – which is exactly what I attempt in Chapters 5, 6 and 7. Taking my account of SE as a foundation, there I address more specific biological issues. All of these discussions generally aim to re-evaluate the explanatory power of chance in molecular, cellular and developmental biology.

In the fourth section, I briefly discussed Merlin's idea of weak randomness concerning genetic mutation. I underscored that her account is of great interest for the link she elaborates between chance and the dimensions of time and space at the cellular and molecular level.

Inspired by her work, I will develop an idea of chance in translation in Chapter 7 that takes into account the relationship between time and stochasticity.

Part II – Biological-centered analysis

Introduction to the second part of the dissertation

In the first part of this dissertation, I developed a philosophical account of chance in terms of stochastic explanation (SE). In the chapters that follow, my account of stochastic explanation (SE) will no longer be the subject of discussion. Rather, in order to enable greater focus, it will become an assumption[180] underpinning more local philosophical reflections on specific biological processes and/or explanations. In Chapter 5, I will focus on the status of chance as noise in gene expression. More specifically, I will look at the meaning of chance in a case study of alternative splicing. In Chapter 6 I will ask what kind of reduction and reductionism can be embraced in studies of chance. In Chapter 7 I will examine the role that time plays in explaining stochastic dynamics.

1. Notes and functions of this second introductory chapter

As in the first introduction of this dissertation, here I would like to say a few words about the genesis of the three chapters which follow. Although they come at the end of my work, they were in fact written prior to conceiving the idea of researching the role of chance in explanation in terms of abstraction. So, why place them at the end? As already specified in the first introduction, my motivation concerns the structure of my analysis. The following three chapters are detailed descriptions and critiques of precise biological case studies. Putting them at the beginning of the dissertation would have meant an excessively vertiginous zoom into biology for the reader, without at the same time giving a precise theoretical framework with which to properly frame them. In addition, the order of writing does not need to follow the order of elaboration. Having established a bottom-up project in philosophy of biology, it was predictable that the first things to come up would be philosophical reflections focused on the specific cases studies from which they arose. But that does not minimize the fact that, prior to writing, developing a more general theoretical framework in which to collocate these detailed reflections is completely legitimate. So, the order in which the dissertation is read does not have to follow the order of writing, especially with respect to a bottom-up philosophy of science project.

These chapters aim to investigate more local and specific philosophical problems that are, however, always related to each other by the idea at the base of my stochastic explanation (SE) – that chance has an essential role in the explanation of certain cellular and molecular phenomena. These chapters should be thought of as sunrays that depart from a central core/sun

[180] Even though I will show its relevance for most of the specific case study I will analyze.

to explore new lands. Being sunrays, they maintain the essence of the sun from which they emanated: a project in philosophy of biology that attempts to elaborate the role of chance in biological explanation in terms of abstraction. In Figure 1 below, I propose a schematic representation of the structure of my work, from its core idea of stochastic explanation (SE) found in the first two chapters (the sun), to the more local reflections of the final chapters (sunrays).

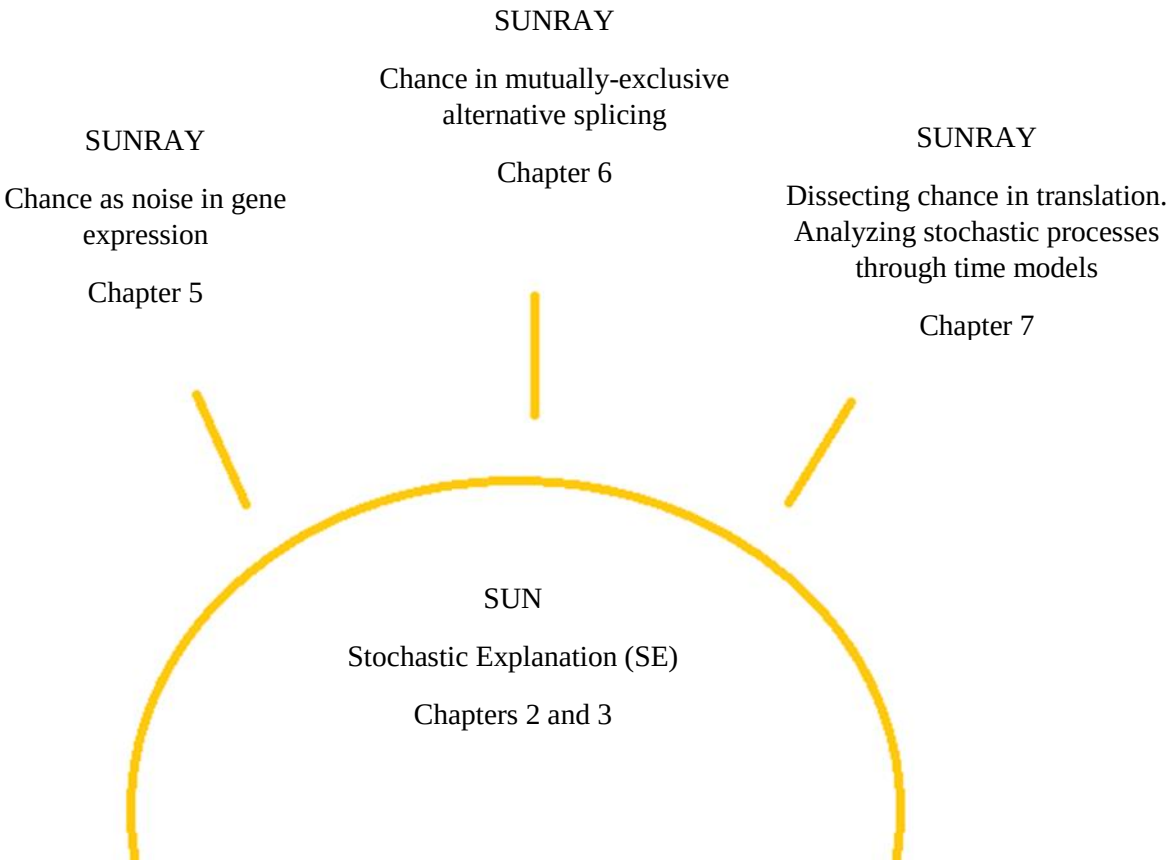


Figure 1 The metaphorical structure of my dissertation's topics.

But let's look specifically at what these chapters will be about. Why are Chapter 1 and 4 not present in Figure 1? Chapter 1 is not included since it addresses meta-philosophical preliminaries which, since they are "meta", do not address the main topic of the dissertation. Nor did I include Chapter 4 because, as mentioned in the first Introduction, it is an interlude in which I give a non-exhaustive overview of the existing studies on chance in biology which appear to me useful for developing these final three chapters.

Chapter 5 addresses the issue of noise in gene expression. In cellular biology it is common practice to refer to chance in terms of noise, that is, as an error, a random deviation

with respect to what is expected in terms of the proteins produced. My argument says that chance is more than just noise, to the extent that it is an essential element in explaining these kinds of molecular processes. I will not resort to the idea of abstraction contained in SE (although it remains the theoretical background of these reflections). However, I will develop three new arguments in favor of the thesis that SEs can be better than more detailed ones. 1) We do not always need a high degree of detail to explain certain dynamics such as those in gene expression; 2) I will show that the proportionality criterium by Woodward (2010) corroborates my thesis; 3) Three historical examples which evidence how less detailed explanations have had – and still have – great value in explaining important cellular dynamics. This chapter is the one in which the presence of normativity within the argument is most pronounced. I provide prescriptive claims (cf. Chapter 1, section 4) in this context along the following lines: since I argue that chance can have a role in biological explanation, then, at least for certain phenomena, it is preferable to use a stochastic explanation than a hypothetical non-stochastic mechanistic explanation. I therefore propose that in certain contexts (e.g. in the study of gene expression), biologists could *rethink* certain explanations in order to provide a greater number of stochastic explanations.

In Chapter 6, I venture into the process of mutually exclusive alternative splicing, one of the most complex phenomena involving processes described as stochastic. Through a careful and pointed analysis with respect to what biologists write about this case study and how they define its stochastic nature, I attempt to develop a precise definition of chance involved in this type of process. Why did I choose this case study and not, for example, gene expression? I have at least three reasons. The first reason reflects the fact that in contrast to gene expression, which is composed of a large number of different processes and events, splicing is a specific process, which means keeping it in mind during the development of the analysis is easier and more feasible. Related to the previous point, the second reason is the fact that splicing is a single process which may encourage finding a specific notion of chance. Generally speaking, in gene expression this would not be possible because its processes are very different, meaning chance could have many different facets. The third reason is that splicing is in fact one of the processes of gene expression. Studying it could therefore be a first step in discerning the various meanings that chance could have in this intricate series of processes and events. By contrast, in the second part of the chapter, I will propose a reflection with respect to the different types of reductionism that can be adopted when talking about chance at the molecular and cellular level.

Chapter 7 is an attempt to link chance with time. Why time? First, because chance is a property of descriptions of certain processes, and processes and events happen over time.

Second, reading the literature in molecular and cellular biology, the overall impression is that time is rarely made explicit in the description and explanation of stochastic processes. This lack therefore encourages a deeper analysis. Could this reflection on the parameter of time contribute to the explanation of some processes described as stochastic? In this chapter I answer in the affirmative. Focusing on the case of alternative start-codon selection, an event in which biologists themselves call for an "inherent stochasticity" (Boersma *et al* 2019, p. 471), I provide a framework that shows that the explanatory role of chance in this case study can be unpacked only if we seriously consider the parameter of time. Providing these models also indirectly reinforces my argument about chance as an abstractor. Indeed, in the case of alternative start-codon selection events, chance provides an *explanans* that holds and synthetizes details in order to account for the *explanandum*. Overall, throughout these final three chapters, I seek to show how original reflections on chance are necessary in light of biological practice, and how more are needed in order to continue exploring the role of chance in molecular and cellular biology.

Chapter 5: From noise to chance in gene expression and cell differentiation[181]

Introduction

The first case study I propose is chance-as-noise in the process of gene expression. I chose it for several reasons. 1) In the 1990s, it became possible for the first time to study stochastic dynamics within single cells (Elowitz *et al* 2002; McAdam, H. & Arkin 1997). From that point onwards, noise in molecular cellular biology became the main notion evoking some signification of chance (see, e.g. Harton and Batchelor 2017), and it is of philosophical interest to unravel its meaning. 2) The notion of noise presents conceptual confusion which is fertile ground for philosophical research and calls for a clarificatory analysis. 3) Chance as noise is conceived as having a negative epistemic value. When in the biological literature a dynamic is labelled as "noisy", it often means that this dynamic cannot be fully understood and explained in a satisfactory way. In this chapter, I highlight that the conceptualization of noise often conflates three different significations, referring to its causes, its consequences and its operational definition. I also argue that, in molecular cellular biology, the epistemological view of chance should be reconsidered in light of new theoretical and experimental results. My aim here is to acknowledge that chance, if we free it from the label of "noise", can also have a fundamental explanatory power and it is crucial to explain certain biological phenomena.

The chapter is structured as follows. In the first part, I outline a non-exhaustive history of the notion of noise in molecular and cellular biology. In the second, I unravel three different meanings that are attributed to noise in gene expression. In particular, I sketch a historical reflection which shows how before the 1960s chance was not associated with noise, and the chancy source of heterogeneity in cell phenotypes was seen as an interesting topic of research (especially with respect to bacteria analysis). In the third part, I argue that in certain contexts concerning the explanation of certain biological phenomena it is better to use a stochastic explanation (SE) than a more detailed one, for example, variability in cell fate and cellular differentiation.

The ultimate goal of this chapter is to suggest that the discussion of chance in gene expression should be broader than its characterization as mere noise.

[181] This chapter is a reworking of an article I published with Francesca Merlin for Elsevier: Casali and Merlin (2020).

1. A brief history of noise in gene expression, from the 1940s to the present day

The historical overview I am about to propose should not be understood as exhaustive. Its aim is only to provide the reader with a rough map for understanding the origin and development of the notion of noise. After this short journey, I describe how, from the 2000s to today, a partial conceptual evolution has allowed noise to begin to be seen and conceived as functional for the cell.

1.1 From the 1940s to the 1990s

Though biologists have been aware of stochastic cellular and molecular phenomena since the 1940s,[182] it was not until the 1950s that the first biological studies were published. These studies concerned morphological asymmetries in bilateral organisms, in particular between the number of bristles in *Drosophila Melanogaster* (Wright 1952; Reeve and Robertson 1953; Thoday 1956; Latter 1964) and in the morphological shape of butterflies' wings (Manson *et al* 1967; Soulé 1982). These case studies were optimal for isolating and studying chance. Indeed, as the differences analyzed were *within* a single organism (for example, between a butterfly's two wings), there were no detectable differences either in environment (the two wings share same environmental conditions) or in the genome (the two wings are orchestrated by the same genome). Therefore, biologists assumed that morphological asymmetries were due to some sort of internal stochastic perturbation they called "noise". In those years, the exact nature and origin of these fluctuations remained unknown.

Conrad H. Waddington (1905-1975) in *The Strategy of Genes* (1957) was one of the first biologists to use the expression "developmental noise" to describe this internal source of variation, unrelated to either genes or to environmental changes (see also Lewontin 2002). He elaborates the famous epigenetic landscape consisting of a three-dimensional surface with valleys and mountains and a sphere that rolls around it (Figure 1).

[182] For example, in 1940 Kramer introduced a stochastic approach to studying chemical reactions (see Turner *et al* 2004).

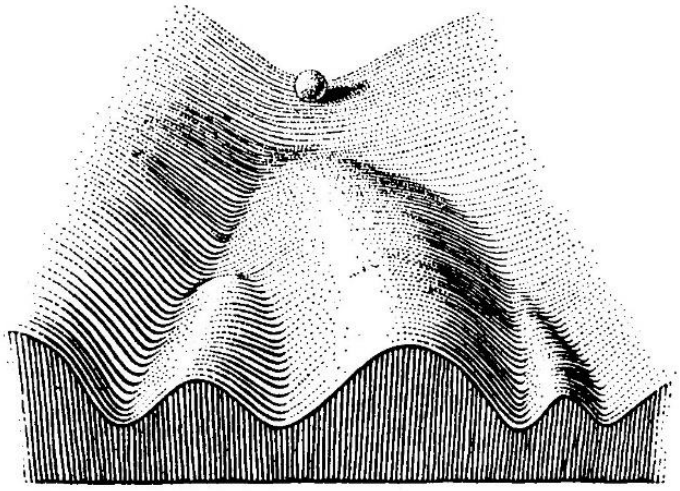


Figure 1 Waddington's representation of the epigenetic landscape (1957).

In this image, the ball is a cell, and valleys are the possible pathways of cell differentiation. What is of interest here is that Waddington made the hypothesis that certain fluctuations can influence the cell differentiation pathways. These fluctuations were not caused either by genetic modification or the environment, but by "intangible internal sources of variation" (p. 39). He describes these internal sources of variation as the ball's own imperfections: "if canalization is represented as a valley in an epigenetic landscape, the noisiness of the system might perhaps be symbolized by the imperfection of the sphericalness of the ball which runs down the valley" (Waddington 1957, p. 40).[183] But what do the sphere's imperfections, these fluctuations, represent? Waddington gives a precise description of the origin of these fluctuations with respect to the metaphor, but does not elaborate on what this implies for biological reality. I do not intend to critique this epigenetic landscape as scholars are already working on these interpretive issues (e.g. Baedke 2013; Gilbert 1991; Perlman 2010; Peterson 2016). Rather, I merely want to establish that Waddington was one of the first to realize that chance plays a role in cellular differentiation (cf. Chapter 4; section 2).

In the 1960s and 1970s, it became clear that the origin of these stochastic fluctuations was attributable to fluctuations within the cellular *milieu*. The influence of physicists in this context was fundamental (Morange 2000) in leading to the recognition that these stochastic

[183] It is noteworthy that Waddington states that developmental noise is related neither to genes nor to environment. In his epigenetic landscape, it is possible to visualize why developmental noise is *not* related to genes: the imperfections of the sphere (the origin of developmental noise according to Waddington) located at the upper part of the landscape (see Waddington 1957, p. 29) do not interact directly with the grips and the roots (the gene network) located in the lower part of the landscape (p. 36). On the contrary, impossible to visualize developmental noise as *not* related to the environment. This because the environment is not depicted at all! Indeed, valleys and hills are not representations of the environment but the possible differentiation pathways that the cell can follow. (on this point see Loison 2022).

fluctuations could be traced back to the thermal agitation of molecules. Physicists' contribution could also be seen as originating of the use of noise as a notion to characterize stochastic dynamics. Noise is a term often used in physics to describe the external or internal perturbation of a system (e.g. thermic agitation of molecules). More specifically, the negative epistemological status of the concept of noise in molecular and developmental biology could be traced back to the influence that the theory of information, and theoretical and applied physics, had on biological studies between the 1970s and 1990s. In the theory of information, a signal represents useful, significant and structured information whereas noise is understood as insignificant, unstructured, irrelevant information which jams the signal.[184] In those same years, biologists were grappling with a tension. While chance was clearly a source of variation, biologists nevertheless found it hard to account for it in their explanations. Morange (2009, p. 21) gives two reasons for that resistance. The first is the "good reason" that biologists lacked the technology to study chance. Indeed, the 1960s-70s was still far off the first single cell study which was elaborated in 2000s. The second is a "bad reason", namely that biologists suffered from "undue admiration for the precision and reliability of molecular processes operating in organisms" (p. 21). This was due to their tendency to remain tied to the idea of a genetic program, which was developed during those years,[185] and which stipulated that DNA already contains all the information for the developmental processes which are therefore considered as already determined (see Moss 1992; Peluffo 2015). I would like to add a third reason for why biologists delayed taking chance into account in cellular dynamics. Nicholson (forthcoming) provides a convincing historical argument that the influence of Schrödinger's *What is Life* (2013[1947]) prevented at least of two generation of biologists from studying and understanding stochastic intra- and inter-cellular dynamics. I already addressed this issue in Chapter 4, section 2.2 and therefore invite the reader to return to that section to fully understand what I mean. Here, following Nicholson, it is sufficient to say that Schrödinger was responsible for propagating the paradigm that the governing principle in biology is not, as in physics, that "order comes from disorder", but that "order comes from order". This led generations of biologists to try to explain molecular processes in a deterministic way, thereby delaying reflection on the intrinsically stochastic nature of molecular dynamics.

[184] For a detailed analysis of the different notions of noise, see Cohen (2005). For an excellent historical reconstruction of the evolution of the concept of noise in biology see Merlin (2009, chapter eight).

[185] The notion of the "genetic program" was used by Jacob and Monod in their famous 1961 article published in the *Journal of Molecular Biology* (see also Morange 2002, pp. 54-55) and, at the same time, by Mayr (1961) in the journal *Science*. Peluffo (2015) hypothesizes that the authors met before these publications and this would explain why they both wrote, at the same time, about the genetic program. For an in-depth look at how genes control development see Gehring (1998).

The 1970s and 80s saw further progress in bringing the study of stochastic cellular dynamics closer, with the addition of new model organisms. As previously mentioned, drosophila was the classical model organism used to visualize macroscopic properties associated with microscopic stochastic dynamics (e.g. number of bristles). While it is true that even today drosophila is one of the most widely used model organisms in genetics, the study of cellular populations, such as bacterial populations, became very popular in these decades and perhaps of more use for investigating microscopic dynamics such as cellular behaviors (Endy and Brent 2001). In this regard, the anecdote that Morange tells in his 2000 book *An History of Molecular Biology* about the first meeting between two great biologists, Max Delbrück (1906-1981, a physicist who turned to biology) and Thomas H. Morgan (1866-1945; a geneticist of chromosomes), is interesting. Morange writes that Delbrück, when he visited Morgan's laboratory, was disappointed to see drosophila as a model because he considered it too complex to reveal "the secret of life" (Morange 2000, p. 42). In line with Delbrück, the new focus on cellular bacterial populations allowed greater precision in controlling the factors that influence cellular proliferation, which helped in better understanding what the origin of noise might be, or at least, what it is not. The hypothesis was that a population of genetically identical cells in a constant environment[186] should all have the same phenotype (i.e. the same differentiation state). However, observation carried out during those years showed this was not the case. Even with same environment and genetic makeup, cells can still manifest diversity. From this it was deduced that noise does not originate either from the genome or the environment, and that there must be something intrinsic to individual cells that produced this phenotypic heterogeneity.

In the 1990s – and still today – the study of noise is mainly focused on gene expression. These studies recognize the fact that gene expression is a series of molecular processes subject to stochastic fluctuations. For example, the process of transcription (the first step of gene expression, consisting of the synthesis of RNA sequences from DNA sequences) is often described as an intrinsically probabilistic process (see Maheshri and O'Shea 2007), because genes are activated or silenced through the stochastic association and dissociation of transcription factors (see also Paldi 2003). Although enormous progress has been made with respect to understanding stochasticity within cells, still today, noise is defined in very

[186] Since cell populations are studied *in vitro*, namely in petri dishes, and all cells share the same breeding ground, cell populations share *virtually* the same environmental conditions. One possible objection to this is that environment is not actually the same for *the whole* cell population because there might be some differences in contingent local events or/and in physical-chemical breeding ground composition. Even if this objection could be seen as conceptually correct, its argument loses its force in the approximation of experimental practice. So, when biologists write about identical environments, they usually intend *significantly* similar environments.

heterogeneous ways and a conceptual clarification of this heterogeneity is still missing. This is why in the next section, I distinguish three of the main ways in which biologists talk about noise.

1.2 Contemporary accounts of chance as noise in gene expression

Let me provide an overview of the way in which contemporary biologists characterize noise. Theoretical and experimental research on noise in biology in the 2000s has focused on its impact on gene expression (Elowitz et al 2002, Klingenberg 2005, Samoilov et al 2006, Maheshri & O'Shea 2007, Raj & Oudernaarden 2008, Pipel 2011). In this context, the term "noise" generally refers to various random microscopic events taking place inside the cell that produce fluctuations in one or more steps of this intracellular process (namely, transcription and translation).

Biologists increasingly point out that chance as noise can play a role in the development of biological systems, sometimes even a functional one, rather than always being a mere nuisance (e.g., Balázsi *et al* 2011; Belson 2021; Ca˘gatay et al., 2009; Eldar and Elowitz, 2010; Feinberg and Irizarry, 2010; Gandrillon *et al* 2012; Gupta *et al* 2011; Holmes *et al* 2017; Losick and Desplan 2008; Maamar *et al* 2007; Mettetal and van Oudenaarden 2007; Meyer and Roeder 2014; Nanjundiah and Bhogle 1995; Nanjundiah 2003; Pilpel 2011; Pujadas and Feinberg 2012; Raser and O'Shea 2005; Raj and van Oudenaarden 2008; Roberfroid *et al* 2016; Sanchez and Golding 2013, Vogt 2015). This opens up the possibility for it to enter into biological explanations alongside acknowledged deterministic processes. In this chapter, I take these recent research results seriously and argue that chance affecting various biological processes involved in cell functioning and organism development may actually play a positive, constructive role, and should thus be conceived as more than just noise (see also Huang, 2009). Rather than being a perturbation represented by an error term, chance can contribute to the explanation of these processes. To order this conceptual heterogeneity, in what follows I will develop three main characterizations of chance as noise, with respect to its causes, consequences and operational definition.

1.2.1 Noise characterized in terms of its causes

The first characterization of noise we can find in the literature refers to its causes. The origin of noise in gene expression is commonly attributed to various physical phenomena that can affect the concentration, localization, and state of the different molecular species involved in this intracellular process. First, molecules inside the cell are in constant thermal agitation,

(i.e. they move and collide with each other continuously). With respect to this simplified idea of molecules colliding with each other in a stochastic way, we need nonetheless to make some clarifications. It is not in fact the case that *all* biomolecules can collide with each other because the intracellular environment is not homogeneous. Cells are structured by numerous compartmentalizations (organelles, nuclei, vesicles etc.)[187] containing different physico-chemical conditions and different macromolecular concentrations. Therefore, the intracellular environment is not a homogenous "open space" in which macromolecules can crash each other free from obstacles but is, rather, a heterogeneous space in which thermal agitation is constrained by these compartmentalizations. With this clarification in mind, we can say that noise in gene expression is therefore often associated to the stochastic movement of biomolecules in thermal agitation. Second, the concentration of molecules involved in gene expression is low (sometimes but a few copies per cell; Pipel 2011). Thus, statistically speaking, the specific behavior of a single molecule could have important consequences for cell behavior. This is called the low number effect (cf. Chapter 4, section 2.2). For instance, the absence of just one transcription factor can lead to the blockage of DNA polymerase, and consequently, blockage in the production of a certain proteins. Nevertheless, we have to be careful on this second point too. Asserting that the concentration of molecules involved in gene expression is low does not necessary means that globally there are a "few" numbers of macromolecules. Miné-Hattab and Taddei (2019) write that less than 3% of the nucleus is occupied by chromatins and the rest of the space is occupied by millions of proteins (p. 105). Furthermore, Ellis (2001) stresses how, even if we cannot speak of the "concentration" of macromolecules in cells (due to their low number compared to the rate of concentrations of molecules, for example in a solution),[188] it is nevertheless possible to designate them as a "crowd" (Ellis 2001, p. 597), which underscores how these molecules occupy virtually the entire cellular *milieu*. This shows that there is no consensus within the scientific community with respect to whether and how the small numbers effect acts at the molecular level. This problem that becomes more complex if

[187] Interestingly, recent works underline that there could be different *kinds* of compartmentalization and not only in the membrane. For example, Miné-Hattab and Taddei (2019) give accurate physical descriptions of what they call membranes' sub-nuclear compartments (MSC). The MSC are localized in the nucleus and are self-organized compartments that "are formed by the ensemble of proteins that bind available chromatin sites" (p. 108). They propose three different models to explain this kind of compartment (binding model, bridging model, and droplet model) that, nevertheless, share the same characteristics, namely being 3-dimensional organizations that delimit compartments without the need of membranes.

[188] It is, however, necessary to specify that authors like Kupiec and Nicholson recognize instead that the law of large numbers can also be applied in biology (see Chapter 4). However, this criticism is not relevant to Ellis' argument because it emphasizes only that the intracellular space is not a billiard table where a small number of spheres move and collide with each other, but a full space in which the stochastic movement of biomolecules is limited by the presence of other particles, organelles, compartments, etc.

different answers are possible, depending on the processes under question. The absence of a single transcription factor can be decisive for the activation (or not) of a transcription process. The other ways around, in massive mRNA production, such as the phenomenon called burst, the single molecule can lose its centrality. Therefore, although the question is still open, we can say that noise is still often associated with, and described in terms of, the small numbers effect. Third, quantum events could percolate up and affect biochemical bonds between molecules. The "percolation argument" (e.g. see Stamos 2001, Stamos 2010) states that since the components of atoms at the subatomic level exhibit random behavior, the macroscopic level can "inherit" the random character of the microscopic level by percolation (Rosenberg 1994, p. 60).[189] In this case, noise is mentioned and described in terms of stochasticity percolating up from the quantum level.

We can therefore say that thermal agitation, the low number effect and quantum effects are the ways in which noise is characterized with respect to its causes. As the source of fluctuations, noise affects the process of gene expression, producing stochastic and unpredictable variation its results (i.e. protein abundance).

1.2.2 Noise characterized in terms of its consequences

In the biological literature on gene expression, noise is also characterized with respect to its consequences. These consequences are not limited to the intracellular level in terms of random variation in the abundance of proteins produced. Noise is also recognized as producing random variation in the phenotype – namely phenotypic differences in isolated individual organisms over time or between members of clonal populations, even in the absence of genetic and environmental changes. There are many biological examples of the effect of noise in gene expression on cell-fate decisions in viruses, bacteria and humans. For instance, noise produces random variation in bacteria's resistance to antibiotics (Kint *et al* 2012) as well as the ability to uptake DNA from the environment and incorporate it into their genome (i.e. bacterial competence state; Maamar et al 2007, Losick & Desplan 2008). It also produces (random) variation in the cell units (called "ommatidia") of drosophila's compound eyes (Johnston and Desplan 2010; for a philosophical discussion see Merlin 2015a, p. 96) as well as in the

[189] This question of whether or not quantum effects percolate up to the biological level has sparked heated philosophical debate. Millstein (2000b) defines – though does not support – the counter-argument, named "asymptotic determinism". Despite proponents of this view agreeing that "it is not in principle impossible that quantum indeterminacy might occasionally alter a biological outcome" (Graves, Rosenberg, Horan 1999, p. 145), asymptotic determinism postulates that percolation does not take place. While the "percolation argument" supports an indeterministic perspective, the "asymptotic argument" supports a deterministic one.

phenotype of cloned animals – such as the case of the first cloned cats (see Raser and O'Shea 2005). Noise in gene expression has also been shown to be involved in the development of various human pathologies, such as different types of cancer (Han *et al* 2016; Huang 2012).

1.2.3 Noise characterized by an operational definition

Another common way to characterize noise is to define it "operationally" (i.e., by the way it can be observed and measured). Noise in gene expression is thus defined as random variation around the mean amount (number or concentration) of proteins produced; in other words, it is conceived and measured as a deviation or error with respect to the average or expected result of the process of gene expression.[190] This way to quantify noise has become its more common definition in the literature. More precisely, noise in gene expression is measured statistically by calculating the coefficient of protein variation or mRNA abundance (i.e., the standard deviation by the mean) in a clonal population of cells at a given point in time (synchronically) or in a single cell during a certain lapse of time (diachronically), in a stable and homogeneous environment. This is considered as "the most direct and unambiguous measure of gene-expression noise" (Kaern *et al* 2005, p. 454).[191]

To sum up, noise in gene expression is commonly considered as having its origin in various physical and statistical phenomena (thermal agitation of molecules, the low number effect, and quantum effects) that can affect one or more steps of this process. It is also characterized according to its narrow and broad consequences (respectively, random variation in the abundance of proteins produced and variations in the phenotype). Moreover, noise is usually operationally defined as random variation around the average result of gene expression, (i.e. the mean abundance of proteins). However, it is pretty unclear what noise actually is. It is not supposed to be identified either with its physical sources or its consequences because it is actually caused by various physical phenomena and variations in protein abundance are its effects. However, in the biological literature, in particular in studies focused on gene expression, one can find various verbal characterizations with no precise ontological or epistemological "flesh". For instance, Kaern *et al* (2005) writes that noise arises from

[190] More explicitly, the expected result corresponds to the one predicted by deterministic models of gene expression, which rely on deterministic chemical kinetics. As a matter of fact, research on noise in gene expression has led biologists to acknowledge the limits of deterministic models of gene expression and to complement them with probabilistic ones (Gillespie 2007) based on stochastic chemical kinetics and stochastic simulations (see also Pujadas and Feinberg 2012).

[191] I should note that the operational definition is strongly related to the characterization of noise in terms of its consequences because, as stated above, noise is defined by the way it is measured at the level of the effects it produces in terms of protein abundance.

"fluctuations in transcription and translation, despite constant environmental conditions" (p. 451); Pilpel (2011) says that "stochastic variation in gene expression levels among genetically identical cells grown under the same conditions is often dubbed 'noise'" (p. 410). Noise is supposed to be fluctuations between its physical causes and its biological effects, but is often ontologically reduced to the first and epistemologically defined and measured in terms of the latter. This mix of ontological and epistemological features blurs the issue of what noise actually is (Figure 2).

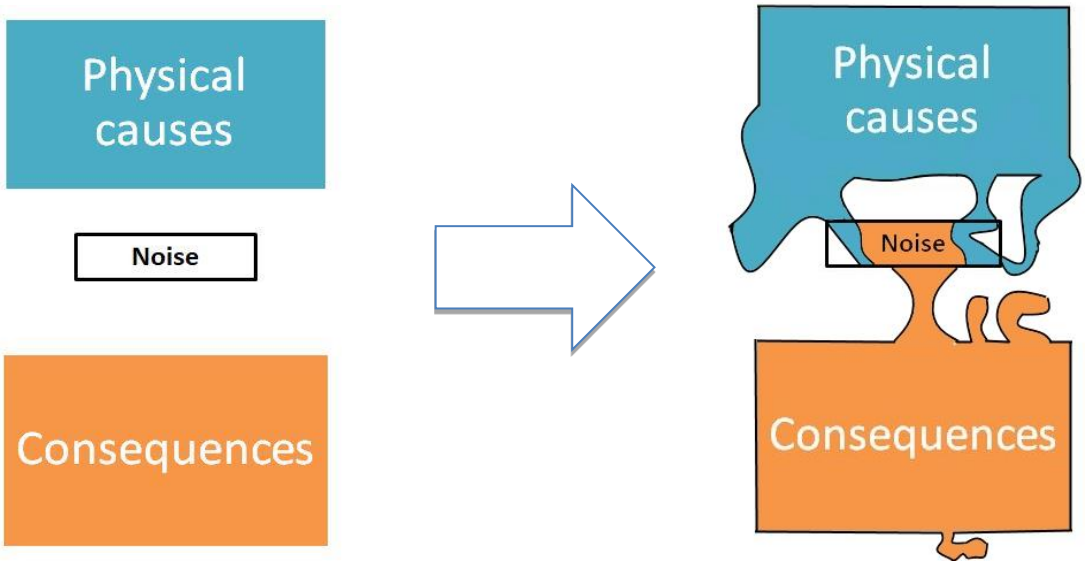


Figure 2 On the left is a visual metaphor of the ideal definition of noise showing its independence from its physical causes and consequences. On the right is a representation of the actual fuzziness of the definition of noise, depicting its mix of physical causes and consequences.

Until to this point, we have briefly reviewed the scientific literature on noise in gene expression and highlighted the way noise is defined and measured, its commonly acknowledged causes, consequences and operational definition. This gives us an idea of the way it is commonly conceived in the recent biological literature, and thus the main features of the epistemology of chance which dominates the biological discourse, that is to say, chance as a perturbation that deviates the process of gene expression from its deterministic course, thus rendering chance a foreign element in explaining the functioning of the cell. However, despite rather broad consensus on this conception, it seems difficult to understand what exactly the term "noise" actually refers to. Is noise reducible to the physical phenomena which affect the various molecular processes taking place inside the cell? Does it correspond to random variation in the

average amount of protein produced due to gene expression? Is it definable only when compared to its consequences? Or is it something else?

Before getting to the meat of my proposal to rethink chance in this specific study context by abandoning the notion of noise and embracing an argument that speaks of stochastic explanations, I offer up one last historical section below. It is devoted to an analysis of Seymour Benzer's (1953) work. Benzer is an interesting figure because he wrote of heterogeneity with an epistemic "freedom" – a conception that was later lost. If Benzer (as well as other biologists of his time) observed heterogeneity in the phenotype of a population of cells, he did not define it as noise but, on the contrary, counted it as a phenomenon deserving attention and further study. Between the lines of my proposal of Benzer's study there is the idea, which could be better developed in a more accurate historical work, that before the influence of physicists in the 1960s and 1970s (mentioned above, section 1.1) and not yet having the paradigm of noise that delayed the study of cellular stochasticity by at least fifty years (see Nicholson forthcoming), biologists were more open to studying stochasticity in the cell.

2. Chance before noise. Benzer's conception of heterogeneity

I propose a rethink of chance which avoids using the notion of noise. In this vein, saying that (some) processes are *chancy (and not noisy)* permits us to open an epistemological research space looking for the role chance can have in the explanation of certain biological dynamics. For example, the definition of chance I proposed in Chapter 3 (i.e., chance as present contingency (CPC)), goes in this direction. Let us recall it:

Given some fixed initial conditions, chance is a property of a description of a process. This property specifies that the described process (or at least some events in that process) could take place, could not take place or could take place in different ways. (cf. Chapter 3, section 2.1).

In this definition there is no negative epistemological reference. CPC is not noise, and does not prevent chance from being explanatory. CPC is a definition lacking evaluative implications, which makes explicit what it means for a process to be characterized as chancy.

This philosophical position, which avoids characterizing chance as noise and instead takes a more neutral notion such as CPC, can find strength in the analysis of a specific historical context in biological research. If we focus on the work of biologists before the 1960s, and thus before physicists introduced the notion of noise to biology, we can appreciate a discourse on chance, but more often on its consequences in terms of heterogeneity, free from value

judgments. Indeed, my claim is that before 1960s, biologists were epistemically "freer" in the characterization of chancy cellular processes. The following analysis of Benzer's (1953) work provides an argument in this direction.

The general scenario with which Benzer (1953) begins is the study of the kinetics of enzyme-induced synthesis in microorganisms. The general methodology used is to add an inducing substance to the breeding ground of cell populations and after that, measure concentrations of the induced enzymes. Specifically, Benzer works on populations of *Escherichia coli* bacteria, the inducing substance is lactose, and the enzyme produced is galactosidase[192] (on the main theme see Merlin F. and Loison 2021). He writes that studying this induced synthesis of enzymes in bacteria suffers from a major problem. Since these studies are performed on cell cultures, "such measurements of the overall activity of the culture do not reveal whether or not all cells participate equally and simultaneously in the synthesis of enzymes" (p. 383). He writes that in principle we have no reason to think that cells in a population have to respond in *exactly* the same way (i.e. each cell producing the same amount of enzyme) when induced by the same amount of inducing substances. Benzer stresses that the average of chemical kinetics does not need to reflect the behavior of any single cell. It could be the case that cells present some degree of *heterogeneity* in their behavior, so that they synthesize different numbers of enzymes in different periods of time, which *cannot be detected* by the average study of kinetics. In the present context, we can reasonably assert that the heterogeneity mentioned by Benzer could be related to chance. Indeed, he underlines "that a given cell might synthesize its maximum amount of enzymes in an abrupt fashion, this transition occurring at *random times* for particular cells" (Benzer 1953, p. 383, emphasis added). We can interpret here that the author is implying that this kind of heterogeneity is caused by events that could be described as happening in a stochastic fashion.

Benzer attempts to create an experiment to test whether it is really the case that some cells behave differently. He proposes measuring the induced production of a given enzyme in a cell population in different conditions, in order to understand whether all cells synthesize the induced enzyme in the same way and at the same time, or not. Leaving aside all the technicalities (e.g. for how bacteriophages are used as discriminants in the test, see pp. 383-384), what is important here is to highlight Benzer's attitude. He is interested in understanding the origin of the heterogeneity of these processes. In exploring what the cause might be, it seems

[192] In the 50s, this phenomenon, often called enzymatic adaptation, was an object of deep interest. Jacques Monod himself wrote a book on it titled *Cybernétique enzymatique*, which remained unpublished until the recent critical edition edited by Merlin and Loison (2021).

to me that Benzer does not rule out that there could have been an important stochastic component. His motto seems to be: "is there any heterogeneity? Could it be stochastic in nature? Well, let's go explore it!". This is a completely inverted attitude with respect to what came later and is still present today in molecular cellular biology (i.e., the notion of noise acting as a limit beyond which we cannot try to characterize, describe and explain the search for stochastic cellular and molecular dynamics). Benzer's paper is itself evidence of the importance that the author gave to understanding and explaining heterogeneity in cell populations, which deserves attention by scientists *as well as* any other biological phenomena under analysis. But Benzer's were not just hypotheses and research suggestions. The results of his experiments showed that in fact, under certain specific conditions, this kind of heterogeneity really does occur (p. 389) and that this kind of cellular behavior requires further study.

In section 1.2.3 of this chapter, I developed a characterization of noise in terms of an operational definition: noise in gene expression is described as a deviation or error with respect to an average or expected result. By contrast, it is clear that for Benzer, heterogeneity in the production of induced enzymes by different cells (treated with the same stimulus) is not intended as "deviation from the mean value" but as a specific cellular behavior that deserves the attention of biologists. More clearly, using Benzer's work, I want to underline that the way in which scientists approach chance in cell studies has not always been negative (i.e. noise), but that there have been biologists like Benzer who approached chance without any underlying value bias. My proposal of chance as present contingency, or more generally, my account of stochastic explanation, seeks to reinstate this attitude (even if in a more philosophical fashion), and to study the description of stochastic dynamics for what they are, without any negative judgment.

Note that, in 1950s-60s Benzer and colleges were not the only scientists mobilizing positive notions of chance. Indeed, a positive view of chance was already present and operative in another biological field focusing on molecular processes taking place inside the cell: immunology. It is interesting to note that the first formulation of this positive attitude towards chance in immunology dates back to the 1950's (Jerne 1955). We suggest this is not accidental because it is exactly the period in which biologists were still free with respect to the possibility of studying chance in cellular and molecular domains, as in the case Benzer. In the following section, I briefly show the acknowledged role that chance plays in some immunological explanations in order to take it as an inspiring example.

3. A positive epistemology of chance in gene expression

3.1. The case of immunology: an example to follow

In this section, I briefly describe how the synthesis of antibodies is accounted for in immunology, and focus attention on the acknowledged role of chance in the explanation of this molecular process. My motivation is to point to the fact that, in immunology, chance has a well-recognized – although not philosophically explicit – explanatory role. This shows that the positive epistemology of chance I argue for in order to make sense of the role stochasticity can have in gene expression is already present and implicitly acknowledged by scientists in other biological fields dealing with molecular and cellular processes. Immunology thus provides us with an example with which to counter the widespread, longstanding resistance in studies of noise in gene expression and cell differentiation, of considering chance as more than just "noise".

Pick up any immunology textbook,[193] and a reader will learn that mature B-lymphocytes produce antibodies, a specific type of glycoproteins, which in turn bind to antigens[194] that, through a cascade of molecular mechanisms, allow the organism to respond properly to the presence of these molecules. Each mature B-lymphocyte can produce only one type of antibody that can bind to only one type of antigen. The entire repertory of antibodies cannot be linearly stocked in the DNA because this information would take up too much space in the nucleus. So, how is this diversity possible? The answer is, by chance! (see also Jaeger and Fernandez 2012).

Present in a developing B lymphocyte are numerous gene segments categorized as *variable* (V), *diversity* (D) and *joining* (J), whose main function is to synthetize antibodies. During B-lymphocyte development, these regions are stochastically brought together in a process called V(D)J rearrangement. Each mature B-lymphocyte thereby contains a unique DNA V(D)J sequence which is involved in the synthesis of a unique antibody through gene expression processes (see Figure 3).

Different pieces of V(D)J are rearranged in two steps: (1) two types of recombinase enzyme called RAG-1 and RAG-2 (Recombination Activating Gene) cut different parts of

[193] There are two types of immunological response: the innate immune system and the adaptive immune system. The former protects the organism from dangerous external agents. It is not species-specific, and comprises of innate immune mechanisms such as skin, mucus, lysozyme, cytokine, transferrin, which respond in similar ways to heterogeneous external agents. The latter protects the organism in a species-specific manner, through mechanisms which respond in specific ways to each antigen. The adaptive immune system is divided into humoral immunity and cellular immunity. The former is mediated by antibodies (immunoglobulin, Ig) produced by B-lymphocytes (a specific type of immune-cell). The latter is mediated by T-lymphocytes. In what follows, we focus on humoral immunity and the processes of antibodies synthesis.

[194] An antigen is any molecule or molecular structure outside the body that is capable of activating the immune response of the organism.

genes at different sites;[195] (2) the DNA repair machinery ligates the DNA strand back together.[196] What is called the "snip-and-fix process" is the process whereby different pieces of DNA (VDJ gene segments) create a unique sequence which can give rise to a single, unique antibody. The variation of these sequences results from: (1) the stochastic rearrangement of (V), (D), (J) gene segments by recombinase enzymes (RAG); and (2) what is commonly expressed as "imprecise joining" between rearranged gene segments. This variation, which is thus considered to be due to partly stochastic molecular processes, is recognized by immunologists as an aspect fundamental to explaining antibody synthesis.

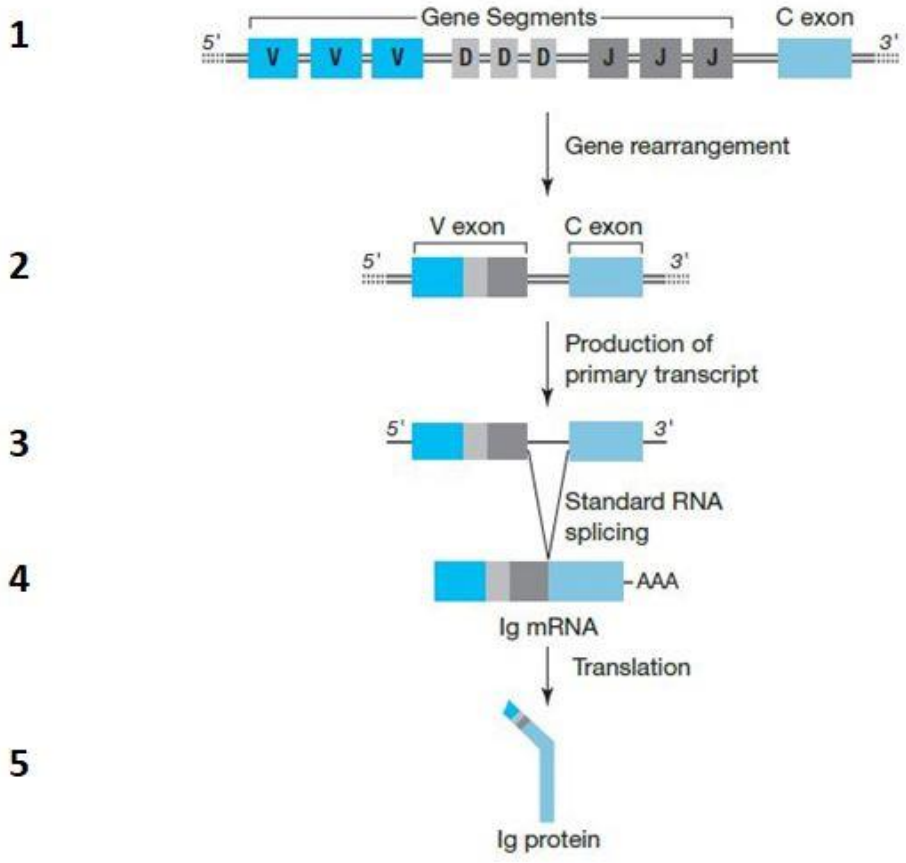


Figure 3**.** A qualitative overview of the process which transforms V(D)J genes segments into antibodies (Ig protein). (1) The DNA strand with segments V, D and J (segment C is not relevant for our purposes) of the immature B-lymphocyte; (2) the DNA after the V(D)J rearrangement in the mature B-lymphocyte. The transition from (2) to (3) represents the first step of gene expression (transcription, DNA → immature RNA); the transition from (3) to (4) represents the maturation of RNA through the splicing process, a mechanism of splitting and lumping of RNA sequences; the transition from (4) to (5) represents the second mechanism of gene expression (translation, mature RNA → proteins). (Figure by Mak and Saunders, 2008, p. 67).

[195] The RAG's enzymes recognize the recombination signal sequences (RSSs) to bind DNA.
[196] For further details, see any immunological textbook (e.g., Mak and Saunders, 2008; Kindt *et al* 2007).

To better understand just how essential chance is in immunology explanation, let us consider two counterfactual scenarios in which no chancy process is involved in the production of antibodies. Either the organism is able to produce just one type of antibody from all mature B-lymphocytes, or the rearrangement of VDJ sequences, which allows the production of different types of antibodies, happens in a deterministic way. In the first scenario, the immune system would effectively be useless for the organism as it would be able to fight only one type of antigen. In the latter, we can hypothesize that the rearrangement process would be far from being an optimal solution in terms of energy expended. Indeed, it would require a lot of energy, and could thus destabilize the general energy balance of the cell[197].

As in the case of immunology, I argue that in other biological fields dealing with molecular and cellular processes, biologists have been aware for decades of the functional and explanatory relevance of chance.[198] They acknowledge that the chancy character of DNA sequence rearrangement is an essential component in explaining antibody synthesis, even if they never make this fact explicit. In fact, it surprises me that there exist no philosophical studies that attempt to make explicit the extent to which chance in immunology is central to explaining antibody synthesis. I will therefore try to do it myself, briefly and with the tools I have developed with respect to my account of chance. I would like to propose that the above-mentioned immunological explanation is indeed a stochastic explanation (SE) in the sense I specified in Chapters 2 and 3. The explanation could be structured as follows:

Explanandum: How can cells produce a virtually infinite number of antibodies (that can, if necessary, bind to any antigen that comes from outside the body)?
Explanans: The random rearrangement of antibody gene segments during the development of B-lymphocytes (cells that will produce antibodies). This random rearrangement allows each mature lymphocyte to produce a different type of anti-body.

Following this structure,
the explanans *of antibody synthesis contains a notion of chance a present contingency (CPC). This CPC provides an* explanans *of a stochastic explanation (SE) through the process of LY*

[197] More general considerations of energy metabolism and optimization can be found in Paldi (2020). For more details on the relationship between the precision of a biological process and its energy costs, see Lestas *et al* (2010), which are also mentioned in Paldi's 2020 work.
[198] Another more recent example in which biologists *make explicit* a functional role of chance is the study by Landau *et al* (2014; see also Swanton and Beck 2014 for a review) in which chance/noise is described as having a strong and clear commitment to a functional role in cancer cell evolution.

abstraction (PLA) by making the synthesis of ***cumulative presentations*** *of a variety of possible ways in which the* explanandum *can be realized. More precisely, the stochastic*[199] *rearrangement of (V), (D), (J) gene segments (CPC) holds and synthesizes in a cumulative representation all the different ways (tending to a very large number) in which these segments of genes can be rearranged (PLA).*

Showing then that in the *explanans* we have CPC (which, among other things, is also very similar to the third example in Chapter 3 in which we talked about "random rearrangement" of microfibrils) allows us to identify this explanation as a stochastic explanation (SE). Obviously, this is just one way in which it is possible to make explicit the power of chance in immunological explanation. In fact, it would also be interesting to deepen the argument and understand in which historical moment, in which circumstances and within which theoretical frame such importance has been attributed to chance. But we leave this line of research for future work, more focused on the history of molecular biology.

We can conclude then that in immunology we can find an excellent example of chance with a constructive and explicative role in molecular and cellular processes. By contrast, research studies on noise in gene expression are still embedded in the deterministic framework which have characterized molecular biology since its origin in the 1950s. The idea of the "genetic program" is a central feature of this view (Jacob & Monod 1961, Mayr 1961). It postulates that the relation between genotype and phenotype is deterministic: the same genes give rise to the same phenotypes, and chance can play no more than a disruptive role in genetic processes (see Peluffo 2015). Research advances since the end of the 1970s have shown that this strictly deterministic view is no longer tenable because of the complex set of processes involved in gene expression, protein synthesis and, more broadly, organism development (for historical and philosophical analyses of genetic determinism, see Sarkar 1996, Kay 2000, Keller 2000a, Hall 2003, Kitcher 2003, Rheinberger & Müller-Wille 2018). Nevertheless, the most common view of chance in the literature on gene expression still describes it as a mere "nuisance" for the deterministic process of gene expression,[200] which consists of running the

[199] Even if in the immunological textbooks we find the word "random", in this context I prefer to use the notion of "stochastic" in light of the fact I am talking about a process (cf. introduction of the thesis, section 1.6.1).

[200] Examples of the negative role of chance (i.e., noise as a nuisance, hence not explanatory of the functioning of biological systems) in the literature are: (1) the study of cell identity, where noise is conceived as an obstacle for correctly measuring the identity of cells (e.g., see Birnbaum and Kussell, 2011; Morris 2019; Reiter *et al* 2011; Efroni *et al*, 2015; Grün *et al* 2016; Various 2017); (2) the study of the alternative splicing process (e.g., Melamud and Moult, 2009, p. 4873; Jin *et al* 2017, p. 11), characterized by the (more or less explicit) assumption that if alternative splicing is due to noise, *then* it is not functional (cf. Chapter 6).

genetic program. While immunology and molecular biology both study processes taking place at the same level of biological organization (namely, gene rearrangement and gene expression), the explanatory role of chance is perfectly acknowledged in the former, and tends to be denied in the latter. Why is there widespread resistance to according any epistemological relevance to chance in the study of gene expression?

It could be argued that whereas chance is functional for immunological response, it is actually detrimental for gene expression, and this is why it is conceived differently in these two areas of biological research. However, this counterargument fails when faced with recent empirical evidence in molecular and cellular biology that calls for the positive epistemology of chance already present in immunology. Indeed, an increasing number of studies of noise in gene expression, like those described in section 1.2, clearly show the role of chance in molecular processes, in particular those taking place inside the cell (e.g., genetic recombination and gene expression). We thus need a positive epistemology for molecular and cellular biology that admits the possibility of chance being explanatory to cell functioning. It is worth noting that this does not mean that all research fields in biology have to include the concept of chance in their epistemology or, to put it differently, that chance is essential in accounting for all biological processes. But we want to underline how, in certain cases, it is necessary to adopt an epistemology that admits chance as an important and even necessary part of biological explanation, such as (but not necessarily limited to) my account of stochastic explanation (SE). Moreover, such a positive epistemology could allow for new questions to be posed, which cannot be addressed from the mainstream perspective of a negative epistemology which views chance as a mere nuisance. For example, can chance play a functional role in development? Can it be selected and passed on across generations? How could its functional role (if any) be explained in an evolutionary framework?[201]

3.2 The reduction problem and the counterfactual argument

Let me now present in further detail the main features of the positive view of chance I propose here for molecular and cellular biology, and highlight the theoretical advantages it brings about for both philosophical reflection and biological research. In this context, chance is not always seen as a perturbation, described with an error term, but can also be an integral part of the explanation of biological systems. However, this does not necessarily imply an

[201] It is noteworthy that similar questions are not completely absent in literature. There is much biological work that already gestures to a positive attitude of chance with respect to explaining stochastic cellular and molecular dynamics (see, e.g., Heams 2014, p. 5 and Pavé 2007 p. 191).

ontological commitment to chance as a specific and irreducible biological phenomenon. Rather, chance that characterizes biological systems can in principle be reduced to the physical processes at its origin even while these processes are constrained and canalized (Bravi and Longo 2015, p. 10; Buiatti and Longo 2013, p. 20) by highly evolved and organized systems (namely, developing organisms) that bias their behavior and main features.[202] In this last part of the chapter, I give philosophical reasons for preferring a stochastic explanation at the biological level. This argument implies an epistemological anti-reductionism of physical explanations – so that when we talk about chance, the explanations we call into question are non-physical but biological in fashion.[203] In the following chapter, I will go on to discuss various types of reductionism. Deepening the analysis, we will realize that a certain idea of epistemological anti-reductionism can, at the end of the day, be compatible with a certain idea of ontological reductionism.

Let me first illustrate with a biological example why, if we do not embrace the positive epistemology of chance I defend and thus do not acknowledge the role it plays in certain biological systems, we leave a hole in the explanation of their functioning. The example, taken from Meyer and Roeder (2014), is about the various steps of cell differentiation and development in animals. The general schema proposed by these two authors is the following:

[202] This is the main difference between my epistemology of chance and, for instance, Kupiec's view who, from my perspective, does not stress enough that organisms are highly evolved and organized biological systems constraining chance. Kupiec is also metaphysically committed to chance and ascribes to it a preponderant and even leading role in the origin of biological organization (see Kupiec 2019), which I resist. More precisely, he argues that molecular interactions lack specificity, and are intrinsically noisy, such that the genetic programming theory is no more valid. More broadly, according to him, all biological processes, and in particular cell differentiation, necessarily need chance in order to take place (see Kupiec and Sonigo 2003). Unlike Kupiec, in this chapter I argue in favour of an epistemological role for chance in biological explanations, but without any commitment as regards its ontology (for further reflection on Kupiec see Chapter 4, section 2).

[203] Here I echo the famous anti-reductionist argument by Putnam that claims that "the fact that the behavior of a system can be *deduced* from its description as a system of elementary particles does not follow that it can be *explained* form that description" (Putnam 1973, p. 131; emphasis in original). More generally, Kaiser (2015) writes that "what holds for ontology need not also apply to epistemology" and that "you can have ontological reduction without, at the same time, having epistemic reduction" (p. 51). In the same vein, Weslake (2010) gives interesting arguments for the thesis that nonfundamental explanations (e.g. biological explanations) "can be deeper than explanations in fundamental physics" (p. 273; cf. Chapter 3, section 1.2). By contrast, Weber (2005, chapter 2; 2008) defends what he calls "explanatory heteronomy" according to which in certain models that describe biological phenomena, "the explanatory relevant generalizations" (Weber 2008, 995) could be made only using physico-chemical laws. Biological reductionism is a rich debate that it was, and still is, approached in different fashions. For a more detailed discussion of reduction and reductionism that specifically concerns my account, see the next chapter, section 2.

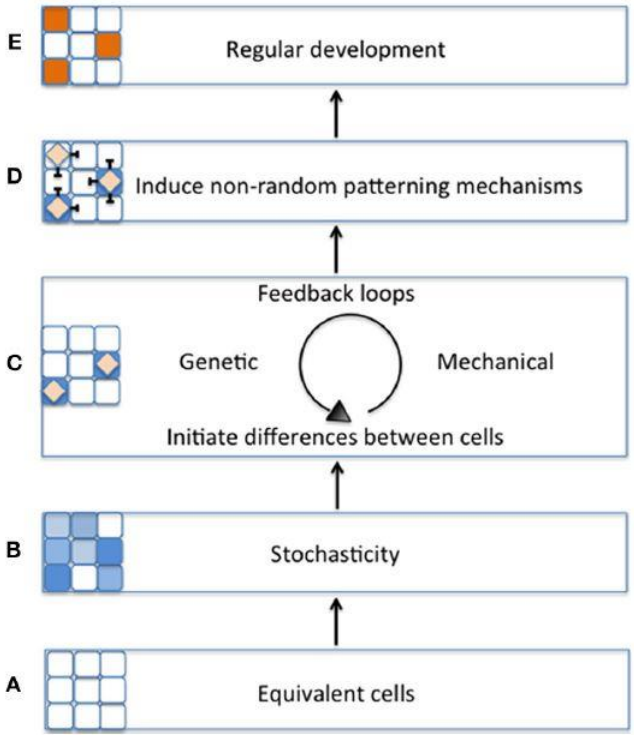


Figure 4**.** Schematic model of the importance of stochasticity in promoting regular plant and animal development. (A) In early development, cells are morphologically equivalent. (B) Differences between morphologically equivalent cells arise through stochastic fluctuations in gene expression. (C) Differences between cells are stabilized by regulatory mechanisms such as feedback loops. (D) Once stabilized, lateral inhibition mechanisms take place. (E) Regular development takes place in a regular manner (taken from Meyer & Roeder 2014).

In Figure 4, the second step of cell differentiation in animals is "stochastic" (step B). This is not the case in the following step (step C), during which "[g]enetic and mechanical feedback loops can enhance and solidify these differences [caused by stochastic processes] to begin cell differentiation" (Meyer and Roeder 2014, p. 1). Let us focus on step B. Which processes are stochastic? And above all, how can we account for their stochastic character? Could it be the case that what we call "a stochastic step" is not chancy at all but rather a complex set of deterministic and unidentified mechanisms?

These questions recall the philosophical debate about the origin of the stochastic character of evolutionary theory (for a review, see Malaterre and Merlin 2015). It can be roughly summarized as follows: is the stochastic character of evolutionary theory due to our ignorance of some "hidden variables" (Rosenberg 2001) or is it an ontological feature of the biological world, in particular of biological evolution (Brandon and Carson 1996)? This metaphysically ambitious question remains unanswered. According to Millstein (2000b), the debate is doomed to a philosophical dead end because we have not yet enough knowledge to argue for the

deterministic or the indeterministic character of the evolutionary process. In light of Millstein's warning, and to avoid entering in this sort of debate, I take the example of cell differentiation and development in animals in order to show and argue for the fact that stochasticity (and more generally, chance) can play an explanatory role in the study of these processes. It can thus be seen as a specific biological explanatory element, which does not imply any ontological commitment concerning its irreducibility to physical phenomena or its eliminability by hidden variables yet to be discovered.

In order to show more precisely how stochasticity plays an explanatory role, let us return to Meyer and Roeder's paper and analyze in detail their model of cell differentiation in animals. Lateral inhibition (Figure 3, step D) is a mechanism that enables neighbour cells to communicate with each other in order to synchronize their differentiation during, for example, neural development (Morrison *et al* 2000). In this way, cells inhibit or activate the expression of specific genes in a coordinated manner. In animals, two trans-membrane proteins, the receptor Notch and the ligand Delta, are involved in this mechanism. When the receptor Notch of one cell interacts with a *trans*-Delta ligand (i.e., the Delta of another cell), it triggers a phosphorylation cascade that activates the expression of Notch genes (*trans*-activation). If this does not happen, the interaction between a Notch receptor and a *cis*-Delta (i.e., within the same cell) produces a *cis*-inhibition in the cell (i.e., no Notch genes are activated). Moreover, when a cell has a stabilized higher level of Notch receptors than Delta ligands, it is in a "receiving state", in which it is "waiting for" an interaction with *trans*-Delta to activate its own Notch genes. By contrast, when a cell has a stabilized higher level of Delta ligands than Notch receptors, it is in a "sending state", in which it is "waiting for" a cell with *trans*-Notch to activate. What is the mechanism that allows cells to be in a sending or waiting state? The answer is, regulatory mechanisms such as genetic or mechanical feedback loops (Figure 3, step C) – but what determines the concentrations of ligands and receptors in cells that, once stabilized, are responsible of the sending/waiting cell states? Sprinzak *et al* (2010) show that small differences in the concentration of Delta and Notch factors are caused by stochastic fluctuations in the expression of genes involved in the production of these proteins. In light of these molecular details, it is no longer possible to deny that stochasticity is a relevant feature in the source of differential gene expression (Figure 3, step B), which constitutes a precise step in cell differentiation, and so of random variation in the state of cells. Put differently, if we reject the explanatory role of stochasticity here, we cannot account for what step B gives rise to (namely, differences between isogenic cells of a developing embryo/organism), and we thus have a gap in the explanation of cell differentiation and development.

Let's examine this point more closely in order to argue for it in a counterfactual manner. With no stochastic fluctuation, the following scenario should be imagined in order to explain what happens, at step B, during cell differentiation: cells can regulate gene expression in a fine-grained manner, which here means the concentration of Notch and Delta proteins. Otherwise, with neither stochastic fluctuation nor a deterministic mechanism present, cell differentiation could not take place. This counterfactual scenario shows why stochasticity should be seen as contributing to explaining how cell differentiation takes place and should thus be understood as an explanatory feature of the developmental process: genes are expressed in a stochastic manner, thus producing random fluctuations in protein concentration that are stabilized, giving rise to various lineages of differentiated cells.

At this point it might be objected that a stochastic element is used in the explanation, as I do in SE, simply because the processes are too complex to list out and describe. I do not deny that there may be the possibility of developing more accurate explanations, but my claim is only to say that SE has an explanatory power that is *independent* of whether it is possible to develop these more detailed explanations. The fact that SE has the value of holding and synthesizing all possible ways in which the *explanandum* can relate to all possible *explanantia* is independent of the fact that other explanations can provide us a more detailed description of a certain phenomenon described as stochastic.

3.3 Three reasons for favoring biological explanations in terms of chance

At this point, it is important to note that I began the first section of this chapter with the aim of showing how chance can have a role in (some) biological explanations. But I have realized that in order to answer that question, we have to formulate it in a different, more sophisticated way. To explore the possibility that chance could have a role in biological explanation, we have to ask why a stochastic explanation (SE) could be seen as a better explanation than a fine-grained one. Referring to what I wrote in Chapter 3, this means that I have to show why a stochastic explanation (SE) with a notion of chance as present contingency (CPC) can be better than[204] a more detailed explanation of the same phenomenon under

[204] It is important to note how the choice between explanations (i.e. stochastic explanation vs. possible other fine-grained explanations) has to be a choice between adequate explanations, that is to say between explanations that work. On the contrary, choosing (e.g.) higher level explanations *because* a finer-grained one has not yet been developed is not a real choice, because it is forced. We are not interested in this hypothetical scenario because it does not permit us to highlight the philosophical *reasons* showing why we should use/prefer a higher stochastic explanation than a finer-grained non-stochastic one. We need another scenario in which there is the *real* possibility of choosing between different explanations. In this way, we can highlight how these reasons are *decisive* in favoring stochastic explanations over more fine-grained and non-stochastic ones. Furthermore, these philosophical

consideration. If in Chapter 3 I went into detail with respect to the philosophical account, in this and the following sections, I provide new arguments in favor of this hypothesis by closely comparing specific case studies, more precisely, two related reasons and three instances of historical evidence.

3.3.1 The first argument in terms of the level of explanation

The first general reason to favor explanations in which chance plays a role is that a more detailed explanation, that is non- stochastic and at a lower level, does not always guarantee a higher explanatory power than a higher level, less detailed one.[205] Think of Borges' story of an empire in which cartography has reached a level of accuracy such that it only produces very detailed maps on the same scale as the empire itself (Borges 1962). One can ask whether such maps would be useful to anybody. By analogy, would a complete description of the physico-chemical processes involved in gene expression and protein synthesis be wholly explanatory of cell differentiation and development? The *hubris*[206] of completely reductionist explanations[207] makes us forget that it is possible to abstract away fine-grained details in order to produce better explanations of the phenomenon being studied. Of course, in our case, an explanation mentioning all the physico-chemical details of gene expression and protein synthesis would be useful for explaining the origin of variability in protein concentration. However, what is at stake here (i.e. our explanatory purpose) is not to explain how it comes to be that the concentration of proteins varies in a random manner from cell to cell, and even within the same cell over time. Rather, I investigate how such variation in protein concentration, which turns out to be well described in a stochastic way (i.e., in terms of probability), contributes to the process of cell differentiation and, more broadly, to development. In other words, if the *explanandum* is the origin of stochastic fluctuations in gene expression (i.e., stochasticity itself), the *explanans*

---

argumentations (implicitly) reject a certain idea of subjectivism of explanations, which are very popular in the reductionism/anti-reductionism debate (e.g. see Sober 1999).

[205] On this point, see also Jackson and Pettit (1992) who write "there is no reason to think that finding smaller and smaller levels of causal grain means better and better explanations" (p. 16).

[206] It is interesting to note how the reductionist tendency for complete explanations is not always due to *hubris*. For example, Woodward (2010, p. 295) claims that adopting a "reductive direction" can come from the need to find a more "stable causal relationship" in a scenario of complex distal causal relationships. Nevertheless, he claims, even if the necessity of stability drives us towards a reductionist approach, it could be the case that macro-relations are more stable than micro-ones, and that macro-relations satisfy other conditions better (e.g. proportionality, see later in this paper).

[207] Rosenberg (2006) writes that an explanation is complete only if it explains the phenomena in molecular terms: "explanations in functional biology need […] [to] be corrected, completed, or otherwise made more adequate by explanations *in terms of molecular biology* (2006, p. 26; emphasis added). He adds that "there is such thing as a complete and correct explanation independent of contexts of inquiries' questions" (2006, p. 44; see also Weber and Van Bouwel 2007 for a critique). By contrast, in following pages I will show how context can instead play a key role in the development of explanation (cf. my discussion of van Fraassen, Chapter 3, section 1.1).

should be formulated in terms of physico-chemical causes. Otherwise, if the *explanandum* is how cell differentiation takes place, the *explanans* should be identified as stochasticity in gene expression and the consequent random variation in protein synthesis.[208]

This reasoning emphasizes how adding details does not always make an explanation better. To make this first general argument stronger, I would like to propose two sub-arguments that try to go deeper on this point:

1 - As Baedke suggests (2018), adding details to an explanation does not necessary imply its improvement. Indeed, it could be the case that, considering other detailed elements (e.g. other causal dependencies), the explanation changes its original structure becoming, thus, *another explanation*:

"this strategy [of adding details] leads to a new, yet related, explanation addressing a *different* explanandum, not necessarily to a more fine-grained explanation of *one and the same* explanandum" (Baedke 2018, p. 186, italics in original)

In the same vein, referring specifically to causes and effects, Ylikoski and Kuorikoski (2010) write:

"[…] knowing more about how the cause brought about its effect often *broadens* the explanatory landscape rather than improving the original explanation" (p. 216; emphasis added).

Let us come back to Baedke for a moment. To make his argument more convincing, he proposes a concrete example showing that this shift can in fact occur if we add more details Y to a certain explanation X. He uses causal explanations as an illustration. The inheritance of certain DNA methylations (DM) are one of the causes of the obesity inheritance (OI). I can then explain the OI by using the description of DM (Baedke 2018, p. 186). To make things even clearer with respect to the author's example, one more element should be introduced, which is the fact that the explanation could be seen as a why-question such as the following (van Fraassen, 1980):

[208] Note that my claim joins Garfinkel's argument (1981) for the autonomy of macro (higher-level) explanations. They do not have the same object and explanatory target as micro (lower-level) explanations, so that they should be preferred in accordance with the explanatory context.

(A) DNA methylation patterns (DM) → obesity inheritance (OI)
(*Explanans*) (*Explanandum*)

Following (A), "why do we have obesity inheritance (OI)?", "because there are DNA methylation patterns (DM)". This is a causal explanation settled as why-question. But say we were not satisfied with this explanation (A), and wanted to go deeper and ask, what is the complex molecular process that permits DM to give OI and how would the explanation be modified? Would this new question lead to (just) an enrichment without a substantial change in structure of (A)? Baedke would answer this question in a negative fashion. Indeed, his answer might be the other way around: changing the elements transforms one explanation in another. This process is what he calls "the shifting of an explanation" (Baedke 2018, p. 186). Indeed, if the question is to gain understanding and explain all the molecular details of OI, the explanation would become as follows:

(B) Fine-grained processes of DM → obesity inheritance mechanisms (OIM)
(*Explanans*) (*Explanandum*)

With explanation (B) we can structure the following *new* how-question: "how can we explain obesity inheritance mechanisms (OIM)?". The answer could be "we can explain it depicting the fine-grained process of the DNA methylation patterns (DM)". This is obviously a legitimate and interesting explanation but it is different because the *explanandum* of (A) and (B) are different.[209]

2 - The second point concerns errors. Gervais (2013) notes that the "more details one gives, the greater the chances of error" (p. 109). He provides two different parameters to measure the goodness of models and explanations (cf. Chapter 3, section 1.1): plausibility and richness. The former refers to "the degree of probability that a model [or/and an explanation] is accurate, in the existence of, and distinctions between, the various entities and activities it postulates" (p. 105). The latter "concerns the degree of detail a model [or/and explanation] provides in its description of a mechanism's entities and activities" (*Ibid.*). He asserts that a high degree of plausibility does not necessarily entail a higher degree of richness (p. 109) and that the two

[209] Nevertheless, even if Baedke stresses the difference between these explanations, he does not deny their possible similarities. On this point, he specifies that "the relatedness of the two explanations stems from the fact that both might share *explanans* variables; their *explananda* might overlap as well" (p. 240, note 20).

elements are in some degree independent from each other. Furthermore, he highlights an asymmetry. While plausibility is necessary for explanation, richness is not. Indeed, we could have a highly plausible explanation without the counterpart of a high degree of richness in terms of details. The example that he provides concerns a functional model of facial recognition elaborated for the first time by Bruce and Young (1986). In this model, they propose a functional approach to the brain's complex facial recognition process. More specifically, in their model the various functional components have no counterpart with respect to where or how they relate to the brain[210] (quoted in Gervais 2013 p. 122 that refers to Bruce and Young 1986 p. 306). In other words, they do not provide molecular/physiological details with respect to the model in question. Even nowadays, this model is still recognized as having a high degree of plausibility, even though it was developed without any details of the brain.[211] What Gervais suggests is an example of an explanation with a high level of plausibility but a low degree of richness. In conclusion, the author remarks that plausibility could constrain richness: "[f]requently, considerations of plausibility in the form of experimental evidence constrain the richness of a model [or/and an explanation]: add any more information, and the risk of getting it wrong increases. Again, I hold that while at least some plausibility is necessary for a model to be explanatory, richness is only required with respect to a mechanism's activities" (p. 114). This study provides me with a further sub-argument in favor of the fact that adding details (i.e. increasing richness) could not only increase the risk of changing the explanation (cf. Baedke above), but also of decreasing its plausibility.

3.3.2 The second argument in terms of Woodward's proportionality criterion

The second reason I have for favoring higher level explanations in which stochasticity intervenes as an explanatory element, resides in what Woodward (2010) calls the "proportionality criterion". According to Woodward, proportionality is one relevant criterion we should use to choose an appropriate level of explanation. Roughly, a good level of explanation is one in which causes and effects are proportional, so that no irrelevant detail at the level of causes is part of the explanation of the effects.[212] For example, as claimed by Woodward himself (p. 297), if an explanatory target is the generation of spike trains by

[210] Gervais nonetheless specifies that, even if not explained in details, these components "are not drawn by random: they are drawn to accommodate experimental evidences" (Gervais 2013, p. 115).
[211] It is interesting to note that this model – years later – found numerous neurobiological confirmations (Gervais 2013).
[212] Note that Garfinkel (1981) also claims that we should favor macro-explanations because micro-explanations often do not tell us what would have been otherwise. This is due to the fact that different situations at the micro-level can bring about the same result at the macro-level (what he calls, "redundant causality").

individual neurons in response to incoming signals, it could be useful to study the detailed temporal features of this process for each neuron. However, if the aim is to explain the neuronal response to external stimuli, these details may be explanatory irrelevant. Only the overall firing rates should be mentioned in the explanation.[213] In order to better understand what he means by proportionality between causes and effects, Woodward proposes an interventionist counterfactual scenario. If interventions concerning certain details were made at the level of cause and no changes followed at the level of effect, then these details should be considered irrelevant for the explanation because they make no difference for the effects, and thus play no role in their explanation.[214]

Starting from my *explanandum* of cell differentiation (more specifically, variability in cell fates), let us now try to choose the appropriate *explanans* by using Woodward's proportionality criteria.[215] The *explanans* could be at different levels. It can correspond to the physical processes at the origin of stochasticity in gene expression (e.g., thermal motion; quantum effects, etc.) or it can be stochasticity in gene expression itself, which is measured as the random variation of protein concentration in a population of cells. Which one of these two *explanans* better accounts for my *explanandum*? In order to answer this question, the following counterfactual situations can be imagined (see Figure 5).

[213] In the same vein Kaiser (2015, p. 168), distinguishes causal factors from explanatory factors saying that "the natural world exhibits a particular causal structure, which can be discovered and represented. The role of the concept of explanatory relevance is to specify which causal relationships in a phenomenon's long and complex causal history are explanatory relevant" (see also Ylikoski 2007, p. 29). This argument indirectly critiques Rosenberg (2006). Indeed, Kaiser argues that Rosenberg implies that "what is causally relevant must also be explanatory relevant" (Kaiser 2015, p.61).

[214] The other way around works too: if one intervened on certain details at the level of causes and this change provoked corresponding changes at the level of the effects, then these details should be considered as relevant for the explanation because they make a difference for the effects, and they thus play a role in their explanation.

[215] It is worth noting that we are not just looking for a more suitable level of explanation but also arguing that this level is the one where chance plays an explanatory role.

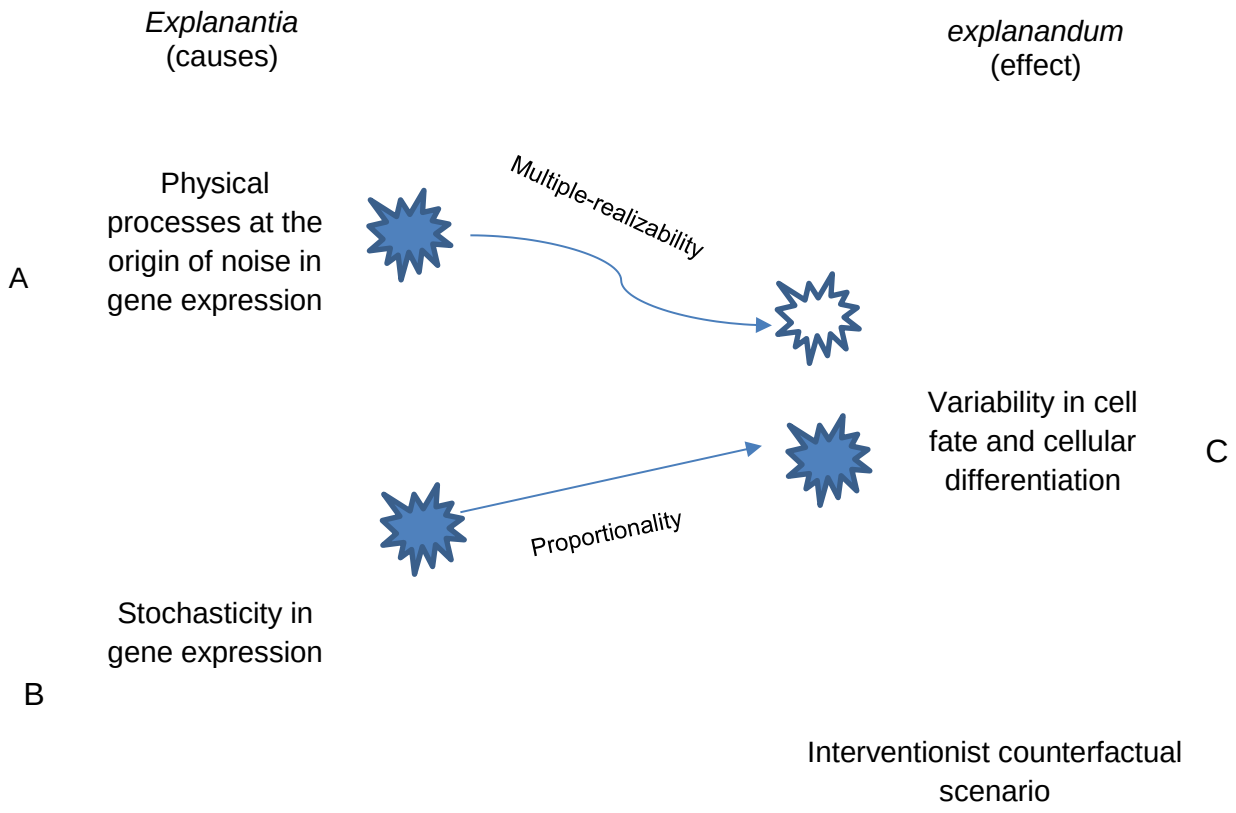


Figure 5. Schema of an interventionist counterfactual scenario. On the left, two possible *explanantia*, A and B, of the *explanandum* C (on the right). The arrows represent the causal connections between *explanantia* and *explanandum*. In the case of A→C (multiple-realizable causality), if one intervenes on A (blue star), C *could* fail to change (white star); in the case of B→C (proportional causality), if one intervenes on B (blue star), it is *very likely* that C changes too (blue star). For the implications of this difference, see the main text.

First of all, what would happen if one intervened to change the physical processes at the origin of noise? Would this provoke some detectable change in the effect, that is, in the process of cell differentiation? In light of the complexity of cell dynamics, a cautious answer could be "maybe yes". But this does not fit Woodward's proportionality criteria because it is likely that some changes in the physical processes at the origin of noise (the causes) "*fail* to be associated with changes in the effect" (Woodward 2010, p. 298, italics in original) – which is to say, in the distribution of various fates in a population of cells. In other words, because of the multi-realizability, at the physical level, of the processes of gene expression and protein synthesis, no proportionality exists between these physical processes and variability in cell fates. We can thus conclude that the physical processes at the origin of noise can be considered causes of variability in cell fates, but that they are explanatorily irrelevant. Second, what would happen if one intervened to change stochasticity in gene expression, and thus the random patterns of protein

concentration in a population of cells?[216] More specifically, what would happen if we intervened to make this random *pattern* non-random? Would this provoke some detectable change in the effect, that is, in the process of cell differentiation? The answer is yes. Actually, if one changed the patterns of protein concentration (which is the way noise in gene expression is measured), it is very likely that the distribution of various cell fates in a population of cells would change. Thus, in this case, the proportionality criterion between causes and effects is fulfilled, which reveals which causes are relevant in accounting for our *explanandum*.

We must add a note here with respect to multiple realizability. Following Figure 5, I said that acting between A and C is possible, that is, even if I intervene on A I can still have C because the latter is multiply realizable. By contrast, I have stated that between B and C, multiple realizability does not exist: if I intervene on B I can have something different to C. In this case, the Woodward criterion is respected and my argument favoring stochastic explanation over more grained and non-stochastic explanation is granted. Nonetheless, an objection could be that if we intend *explanans* B and *explanandum* C as an instantiation of a stochastic explanation (SE), then *we do have* a multiple realizability, even if, I could say, at a different level. As already said, the SE *explanans* holds and synthetizes the different ways in which the *explanandum* can be realized. This is tantamount to saying that C can be realized in multiple ways! Is this multiple realizability that is inherent to the SE *explanans* (i.e. in B) incompatible with Woodward's proportional criterion? The objection goes that if I have multiple realizability at the level of *explanans* B, any possible intervention would be "reabsorbed" in its multiple realizability and we would have the same C. In this case, my argument using the criterion of proportionality by Woodward would fail. Is there an inconsistency here? I think not.

Everything can be settled if we specify better what I intervene on. In fact, if I intervene on the synthesis of a single gene X, this intervention could actually be reabsorbed in the multiple realizability of B, permitting C to be the same. This is the case where the Woodward principle could not be applied. But I propose intervening at the level of the *pattern* of gene expression. If we modify the fact that this *pattern* is random, for example making a certain number of genes expressed in a non-stochastic and regular way in time, then the *explanandum* C can actually change too. The intervention is not reabsorbed by the multiple realizability of the SE *explanans* because what changes is its very structure – it is no longer stochastic. The fact that it is not in fact stochastic can change the way in which cells differentiate. With this argument I consider

[216] On this point see Viñuelas (2012).

the apparent incompatibility between the multiple realizability of SE and the use of the criterion of proportionality by Woodward to be dissolved.

To conclude, it seems reasonable to say that the explanation to choose in order to account for variability in cell fate and cellular differentiation should be put in terms of stochasticity in gene expression. This argument which I have just elaborated is perfectly in line with my account developed in Chapter 3, since the proposed explanation – explaining the variability in cell fate and cellular differentiation through stochasticity in gene expression – is, in fact, a stochastic explanation (SE). The term "stochasticity" in the *explanans* (see Figure 5, B) shows us that genetic expression can occur both in cumulative and disjunctive ways: it can be expressed gene X or gene Y, or can be expressed gene X and gene Y and gene Z, for example. These possibilities are held and synthetized in the word "stochasticity" that thereby acquires the value of chance as present contingency (CPC). Woodward's proportionality criterion tells us that if *explanans* B changes, *explanandum* C should change too if the two are proportionally related. But if something changes in *explanans* B, following Woodward's criterion, does this mean that the explanation is no longer the same, that is to say it is another explanation? No, because according to the notion of chance as present contingency (CPC), the term "stochastic" referred to in this context already holds and synthesizes all the possible ways in which the *explanandum* can be accounted for. So, all the possible variations of *explanans* B are already included by CPC in the stochastic explanation, while avoiding the danger of shifting towards a new and different one (cf. section 3.3.1 about Baedke).

3.3.3 The third argument in terms of historical evidence

Let us finally look at three historical moments which favor my claim that a more detailed explanation, at the lower-level, does not always guarantee a better explanation than a higher-level, less detailed one.

First, following the rise of molecular biology (1940s-1960s), cell biologists were afraid that their research would be reduced to molecular studies. This fear stemmed from the evidence that vesicles, transport proteins, organelles and organelles activities (i.e. all subjects of research in cell biology) are made of the same macromolecules molecular biologists use to work with. In this scenario, molecular explanation of objects and activities studied in cell biology could have been considered a better explanation than cellular-level explanation, which could have produced the end of cell biology. However, this did not happen. Indeed, in the 1980s, cell

biology had its "golden age" (Morange 2000, p. 244) and has proven fertile ground for scientific research up to this point.[217]

Second, as already mentioned in section 1.1, physicists played a major role in the rise of molecular biology by introducing physical ideas to biological studies, such as the notion of noise. Nevertheless, there is no need to interpret this event as a movement towards reductionist thought. Indeed, one of the fathers of quantum mechanics, Niels Bohr (1885-1962), in his August 1932 lecture "Light and Life", which he gave at the International Congress on Light Therapy at Copenhagen, underlined that physicists were never able to explain organism functions through *only* their own discipline-specific competence. Instead of replacing biology with physics, Bohr's suggestion was to do what Morange (2000) calls "epistemological transfer", namely to "see how the new vision of the physical world changed perceptions of biological word" (p. 72). Even if physicists approached biology with their own technology and competences, they should remain aware of the irreducibility of organisms to matters of physics. Bohr compared this irreducibility to quantum mechanics. Just as the latter could not be reduced to classic physics,[218] biology has its own "quantum" that cannot be reduced to the description of matter.

Thirdly, the rise of molecular biology in the 1940s also arose thanks to the rapprochement of two previously separate disciplines, namely classical genetics and biochemistry. The former was founded on Mendel's laws (rediscovered in 1900) and focused on genes, namely the pair of factors that are transmitted over generations. The latter, which was first tested by experiment in 1897 (see Morange 2000, p. 11), studied the chemical properties of molecules such as proteins and enzymes. This was pursued by scientists such as Hermann Emil Fischer (1852-1919), the author of the lock and key enzymes metaphor and Linus Carl Pauling (1901-1994), who first classified atomic bounds as covalent and weak. A paradigmatic and simple example of the interaction between these two disciplines is the identification of genes (as epistemic units) within the physical-chemical structure of DNA (as material units). What matters here is to underline how, even if classical genetics was "more abstract"[219] than biochemistry, the former was not reduced to the latter but allowed the birth of molecular biology in the 30s and 50s.

[217] Note that molecular approaches to cell biology coexist with strictly cellular ones; for example, see Alberts *et al* (2008).

[218] Einstein tried all his life to discover the "hidden variables" that would enable the rejection of indeterminacy in quantum mechanics and frame it in classical physics. His famous phrase, "God does not play dice with the universe" testifies this well.

[219] "More abstract" in the precise sense that classical genetics does not refer to molecular reality but only to unspecified factors of inheritance.

These events in the history of biology show that lower level, reductionist approaches can sometimes fail to prevail over higher level, non-reductionist ones. I see them as providing a strong historical case, progressing in the same direction as my specific epistemological antireductionism as regards the explanatory role of stochasticity in cell differentiation and development.

Conclusion

In this chapter I explored the case study of noise in gene expression and other examples derived from developmental biology. This analysis aimed to provide a convincing argument for the necessity of viewing chance as an important element in explanation. The first point I developed during this analysis is that a fruitful way of exploring the relation between chance and biology is by focusing on explanations, analyzing in what ways chance could be seen as an explanatory element and how this matter in choosing explanations in practical science. Specifically, I argued in favor of according an explanatory role to chance in biological explanation and showed why we have good reasons for preferring stochastic explanations (SE) over fine-grained ones.

Chapter 6: Chance (CPC) in alternative splicing and the issue of the (anti)reductionism

Introduction

In the previous chapter, I argued for the need to rethink the way in which chance is conceived, from an epistemological perspective, in gene expression. I proposed reframing chance, which is often conceived as noise, instead as an essential element in certain biological explanations. In the course of writing the chapter, I realized that there were still philosophical questions to answer. I considered whether I might develop a definition of chance that is specific to a molecular process in addition to my general definition of chance as present contingency (CPC). Following that, I wondered if the discourse on reductionism with respect to chance was satisfied by simply acknowledging an ontological reductionism and an anti-epistemological reductionism. It seems to me this was not the case. Either way, I told myself that there was more work to be done, and this chapter aims to elaborate on these two further points.

On the first point, I will turn to the case study of mutually exclusive alternative splicing and the context in which it is studied which concerns the explanation of neural wiring (a key step in the development of the animal brain) and Dscam receptors (a transmembrane protein that plays a central role in the interaction between different kinds of neurons). But why did I choose this particular case study and not, for example, genetic expression? I have three reasons for this choice:

1) Mutually exclusive alternative splicing is one of the most controversial case studies with respect to chance because still today, biologists wonder whether this process is inherently stochastic or not. If it is stochastic, they question whether this stochasticity is a mere error or actually functional for cell activity. If it is not, they question whether this process can be seen as deterministic. In this second scenario, they also question whether using probability to describe and explain splicing merely admits to ignorance of the underlying (fine-grained) causes: since the process cannot be described and explained in its entirety, we use probability to explain it. This scientific controversy contains interesting conceptual issues from a philosophical perspective, namely what is meant when we say that splicing could be described as a stochastic process.

2) The second reason is related to the previous point and concerns the fact that splicing, being a single process, may encourage the identification of a *specific* notion of chance. Generally, in genetic expression this would not be possible because the processes that this phenomenon

embodies are numerous and heterogeneous, and chance could have many different facets. However, mutually alternative splicing is a *single process* in which pre-RNAs are matured into mRNAs by cutting and pasting exons and introns. This means that conceptual research should refer only to that process, to avoid getting lost in the vast phenomena of gene expression.[220]
3) Splicing is in fact one of the processes of genetic expression, and its study could therefore be seen as an initial step in discerning the various meanings that chance could have in this intricate series of processes and events.

Having covered the reasons why I chose this case study, let's examine my goals: 1) to contribute to the conceptual clarification of the stochastic character, if any, of this particular phenomenon; 2) to show that the study of chance underlying evolution is not only possible but necessary, the case study of splicing bearing tangible evidence of the fact that there are still many molecular and cellular processes in which the role of chance has not yet been fully brought into focus.

The second part of the chapter is devoted to a reflection on the kind of reduction and reductionism[221] that can be embraced when approaching chance in cellular and molecular biology. In the previous chapter, I acknowledged that chance in gene expression is, ultimately, due to physical phenomena such as thermal agitation and quantum effect. I therefore agree with an ontological reductionism concerning chance in cellular and molecular biology. Nonetheless, I also specified that from an epistemological point of view, a biological explanation cannot always be reduced to a physical one. The explanatory role of chance in biology cannot always be reduced to physical explanations. This is the reason for my argument against epistemological reductionism. Although the difference between admitting ontological reductionism while rejecting epistemological reductionism might be intuitively clear, many questions still remain. How can we formulate a clear and satisfactory framework for these two kinds of (anti)reductionism? What does "reductionism" mean in the two cases? When we talk about ontological and epistemological reductionism, what we are talking about reducing to what? How can I argue for an epistemological anti-reductionism without committing to an ontological anti-reductionism too? The purpose of this clarification is to clarify my account of stochastic

[220] Being composed by these other processes of cutting and pasting, alternative splicing could be seen not as a single process. However, biology textbooks always discuss it as an epistemic unit. To put it another way, it is referred to, analyzed and explained as a single phenomenon. For a philosophical discussion on processes at the molecular and cellular level see Baptiste and Dupré (2012) and Dupré (2015). For the historical and philosophical importance of studying RNAs see Guttinger (2021).
[221] I refer to "reduction" as the act of reducing one to another. By "reductionism", however, I refer to philosophical positions that advocate certain kinds of reductions. Although they might be used interchangeably, these two notions do not refer to the same thing.

explanation concerning reduction and reductionism and specify the kind of reduction that is admissible and that which is not.

The structure of the present chapter is simple. In section 1, I will describe and explore the case study of mutually exclusive alternative splicing to propose a concept of chance specific to this kind of process. In section 2, I go deeper concerning the reduction and reductionism that is admissible in my project on chance in cellular and molecular biology.

1. Mutually exclusive alternative splicing: a case study for a definition of chance

To talk about mutually exclusive alternative splicing, I need to introduce two other key elements: neural wiring and the Dscam receptor. Neural wiring explains the way neurons develop. The Dscam receptor is a specific receptor fundamental for the explanation of that process. First, I analyze work by Schmucker (2007) on the Dscam proteins in *Drosophila melanogaster*. I then propose work by other biologists to highlight that the notion of stochasticity concerning this case study is problematic, and no consensus exists regarding in what sense splicing can be qualified as stochastic. Third, I propose a definition of chance to characterize mutually exclusive alternative splicing that aims to create some conceptual order in the description of this process.

1.1 The context: neural wiring and Dscam receptors

The Down syndrome cell adhesion molecule, or Dscam, is a single-pass transmembrane protein of the Immunoglobulin (Ig) superfamily that is present on the surface of developing neurons, mainly on dendrite and axon surfaces. It is synthesized from Dscam genes, which are composed by different domains (i.e. signal peptide, 10 Ig-domains, 6 fibronection III domains, transmembrane segment TM, and cytoplasmic domain CYTO) and that, thanks to alternative splicing, permit the coding of up to 38,016 different mRNAs.[222] From these mRNAs, up to 18,048 different isoforms of Dscam are possible.[223] Since each neuron has particular Dscam isoforms, Schmucker (2007) makes the hypothesis that they "provide a unique cell surface identity" (p. 916). In biochemistry "isoforms" refers to proteins that have the same function despite partially differing in their structures. In the present context, the principle hypothesis is that even if a single neuron has different isoforms, a majority share of isoforms of the same

[222] Schumcker (2007) writes that the number of possible mRNAs is 36,096, but more recent work, such as that by Jin *et al* (2018), specifies that this number amounts to 38,016. I do not want to enter in these technicalities because they are irrelevant for my argument. Nevertheless, when necessary, I refer to the most recent work on this issue.
[223] The role of the Dscam receptors in neural wiring is investigated in flies, *Caenorhabditis elegans*, mice, and humans (see Schmucker 2007, p. 915).

group characterizes its "identity surface" (see after; see also Hiesinger and Hassan 2018, p. 580). We can therefore simplify by saying that each neuron will have its own peculiar and unique Dscam isoforms which allow it to have a unique surface identity.

In the last decade, it has been recognized that Dscam plays a major role in neural wiring (Hattori *et al* 2007; Jin and Li 2018; Santos *et al* 2018; Schmucker *et al* 2000). By "neural wiring" biologists mean the development and growth of neurons forming the neural web that is one of the main features of the central nervous system of so many living things. Schmucker's hypothesis is that the Dscam receptors "are likely to regulate the cytoskeleton" (p. 915), and consequently, the movement of dendrites exploration. This movement is mediated by the interaction of Dscam-Dscam isoforms. The author stresses that Dscam receptors can only mediate homophilic bounds. This means that binding between receptors is possible only between the same Dscam isoforms. When two Dscams of the same isoform (which therefore belong to the same neuron) interact, reciprocal repulsion results. Somewhat counterintuitively, it is precisely their *binding* that causes their mutual *repulsion*. This kind of interaction is called "self-avoidance" and enables sister dendrites (i.e. dendrites of the same neuron) to not overlap. Instead, when different Dscam isoforms of different neuron interact, "no-match of isoforms" results, and consequentially no repulsion takes place. This enables some degree of "tolerance" to different dendrites of different neurons overlapping, but not to dendrites of the same neuron overlapping. Both self-avoidance between sister dendrites and tolerance between dendrites of different neurons are functional for the correct development of the neuron network. The former is important for the reason that the overlap between sister dendrites would be deleterious for the neurons' morphogenesis; the latter is important because the mature neurons' network would otherwise be *made only* by overlaps between different neurons (see the following Figure 1).

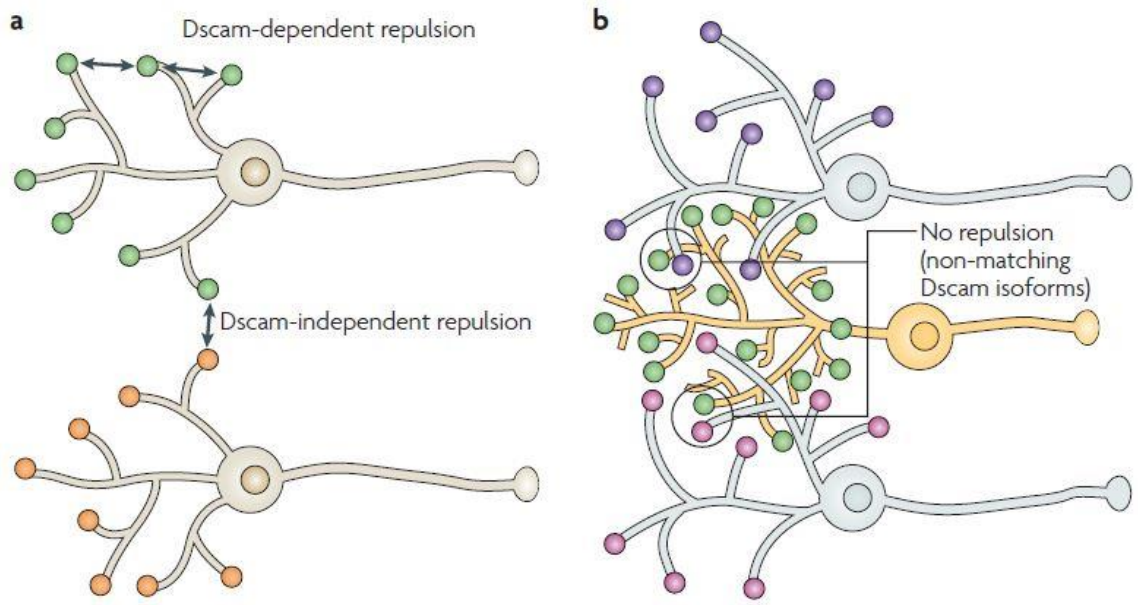


Figure **1** A representation of different kinds of Dscam-Dscam interactions between dendrites of the same neurons and between dendrites of different neurons. a) Having the same Dscam isoforms, when sisters dendrites interact (green points), there is a reciprocal repulsion ("Dscam-dependent repulsion", namely the self-avoidance); whereas dendrites of different neurons, having different Dscam isoforms, do not match and self-avoidance does not take place. The repulsion, if any, is a "Dscam-independent repulsion" (green and orange points); b) Having different Dscam isoforms (and in the absence of other kinds of Dscam-independent repulsion), dendrites of different neurons overlap, permitting correct neural wiring (image by Schmucker 2007, p. 919).

Since there are many neurons in a developing neuronal system, and since each neuron has to have its own surface identity, it is necessary for there to be a very large number of Dscam isoforms (2007, p. 916). Indeed, if there were no diversity in Dscam isoforms, the neural wiring would be incorrect. In this counterfactual scenario, this is because all neurons would exhibit the same Dscam isoforms and so there would be repulsion not only between sister dendrites (a normal process in neural wiring), but also between dendrites of different cells (a deleterious process in neural wiring).

After pointing out the essential role of this diversity, Schmucker wonders: where does the diversity in Dscam isoforms originate from? (Schmucker 2007, p. 917). The answer he proposes is: mutually exclusive alternative splicing. We have then, finally arrived at describing the case study of my interest, a process by which a single premature mRNA[224] is processed in a large range of different mature mRNAs that, in turn, permit the synthesis – through translation

[224] Premature mRNA, or pre-mRNA, is a molecule of RNA that is obtained immediately after transcription. It needs to go through various chemical processes (capping, polyadenylation, splicing) to become mature mRNA.

and post-translational modifications – of a large range of proteins. Indeed, it is this kind of alternative splicing that enables the diversity of Dscam isoforms to be produced. The mutually exclusive alternative splicing of Dscam is an extreme case of alternative splicing. Indeed, in this case, the number of possible results is much higher than all other versions of alternative splicing (such as exon skipping, intron retention, selection of alternative donor and acceptor splice sites, see Jin *et al* 2017, pp. 1-2). To understand in detail this kind of process we first of all have to describe the arrangement of exons and introns in Dscam gene.

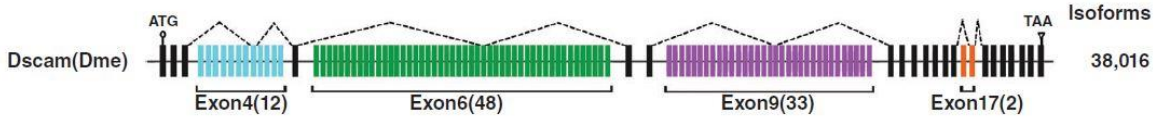


Figure 2 A representation of the Dscam gene. Colored boxes represent alternative exons, black boxes represent constitutive exons, lines between boxes represent introns, and dotted lines represent alternative splicing events. Group of colored boxes of the same colors are mutually exclusive exons (MXEs) clusters that are labeled with a name (e.g. "exon6") followed by the number of exons present (e.g. (48). To the right we have the number of possible isoforms (for further details see the main text; image by Jin *et al* 2017).

Figure 2 represents the Dscam gene. Alternative exons are depicted as boxes with different colors, the constitutive exons are black boxes and the introns are represented by lines. Mutually exclusive exon (MXEs) clusters, namely groups of alternative exons, are clustered together (blue, green, violet and orange lines). For example, the first cluster from the left is the "Exon4(12)", where 4 refers to the name of the cluster and (12) refers to the number of alternative exons present. The colored exons are called "alternative" because in the process of mutually exclusive splicing, only one exon from each cluster is spliced into mature mRNA isoforms (see Jin *et al* 2018, p. 1). This means that in a putative mRNA sequence there will be only one exon from the cluster Exon4, only one exon from the cluster of Exon6 and so on. Therefore, "mutually exclusive" refers to the fact that in a mature mRNA there will be only one exon present from each cluster. This process permits the creation of a huge number of isoform possibilities.

As already mentioned, in principle this process provides the possibility of 38,016 separate mRNAs. This number comes from multiplying the number of mutually exclusive exons clusters, namely 12x48x33x2 = 38,016. How could the mutually exclusive alternative splicing process give rise to this huge number of results (i.e. huge number of different mRNA sequences)? Where does this diversity stem from? Is it a finely regulated process or does it show some degree of stochasticity? If the latter, what does it mean that this kind of splicing can

be described as stochastic? Although biologists continue to ask these questions, still no unanimous answer exists, and confusion about what is meant by stochasticity in the context of splicing remains. Even when mentioning alternative splicing,[225] Kupiec (2019, p. 124) does not investigate it deeper, leaving the conception of chance for this process somewhat vague.

1.2 A critical analysis of explanations of mutually exclusive alternative splicing

While biologists have at least asked the questions, from a philosophical perspective, no work yet explores the role of chance with respect to this case study. This is unfortunate because the problem might be conceptual rather than methodological and/or empirical in nature. It may fall precisely to the philosopher to try to clarify its meaning – or at least this is what I attempt to do in the present chapter. I will therefore try to contribute to the study of the present subject by beginning first of all by looking at what biologists say about it and considering what the disagreement consists of. We will see that confusion, or at least heterogeneous claims, exist around these issues. My purpose is, therefore, to propose a definition of stochasticity for the splicing process that can help disentangle this conceptual heterogeneity.

1.2.1 If splicing is stochastic, then it is noisy and not functional

In some articles, when biologists approach the possible stochastic nature of splicing, they speak in terms of noise (cf. previous chapter). Melamud and Moult (2009) write that:

"[t]he results [of our work] strongly support the hypothesis that most alternative splicing is a consequence of *stochastic noise* in the splicing machinery, and has no functional significance" (p. 4873, emphasis added).

What they say in this quote is of considerable importance. The authors state that most splicing processes are the result of stochastic noise in the splicing machinery,[226] and are therefore non-functional. Echoing the notion of noise in gene expression depicted in the previous chapter, in the present context we can also see how noise is referred to as something (e.g. fluctuations) that disturbs the correct unfolding of a certain process, and is therefore incompatible with any notion of function. Thus, with this work we find ourselves again trapped in the flattening of chance to mere noise.

[225] One of the pioneers of the idea that chance plays a major role in cells (e.g. Kupiec 2009; cf. Chapter 4, section 2).
[226] The splicing machine is referred to as the molecular protein-enzyme assembly that allows the cutting and stitching of the different pieces of pre-RNAs during the splicing process.

However, some works try to push the envelope, opening up the possibility of reflection, albeit with poor conceptual advancements. Jin and Li (2018) ask:

"[i]t would be interesting to determine whether this preference for stochastic isoforms has any functional significance or is merely cell-specific stochastic expression" (p. 6).

This is a bit of a game changer. Two things in this quote are important: the focus on results (i.e. on the Dscam isoforms produced), and the question of whether the heterogeneity of isoforms produced by splicing is actually functional or just the result of "cell-specific stochastic expression". Indeed, in this case we even have the classical dichotomy of whether a process – or rather, the result of a process – is functional *or* stochastic. What's important is that the authors see these two elements as incompatible: splicing – or in this case, its results – *cannot be functional if it is stochastic.*

1.2.2 The alternative splicing event can be real, and then functional, or unreal, and then noisy

Jin *et al* (2018) explore the process of alternative splicing in depth, although what results is an even more confusing conception of the relationship between chance and splicing. The biochemists write:

"[w]hen questioning how many alternative splicing events are real or "noise," we believe almost all mutually exclusive splicing events are functional" (p. 11).

Two points must be extrapolated from this quote. 1) Most alternative splicing is believed to be functional; 2) Splicing can be real or "noise". While the first point might prove useful insofar as the authors more directly link alternative splicing to the idea of functionality, the second nonetheless is terribly confusing as it contrasts a real event with noise, which is not considered by the authors as real. This is of philosophical interest because it seems that the category of what is real is here contrasted with what is not, and "what is not" is associated with noise. Although the "non-existence" of noise might upset the reader, there is actually a plausible explanation as to why the authors use this category of "non-existence". Very roughly, Shannon's information theory postulates that an input and an output exist and information passes from the former to the latter. However, noise in this theory is not integrated into the possibilities of the input, but only as a disturbance in the channel affecting the output. So, in a certain way, the information theory states not only that noise is not information, but also that it

does not exist at the source. This idea of noise as non-information may have influenced biologists so deeply that we can still find it in biological works and, as in the case of our quote, "non-existence" of biological processes.

1.2.3 Splicing is described as a probabilistic process

In spite of the reluctance to explain or consider splicing as a stochastic process, many articles describe it in probabilistic terms, by using probability distributions. If we grant that probability is conceived as a measure of chance, then there exists awareness – that is more or less implicit – of the role of chance in explaining these dynamics. If the process is described with probabilities, from a certain initial condition X, more results are possible, each thus having a probability of less than one. That is, this process is stochastic. Nonetheless, labeling a phenomenon as probabilistic does not help us understand its ontological status. In fact, probabilities can be applied because we are not able to reach levels of detail that would allow us to explain the phenomenon deterministically and because the world presents a certain level of indeterminism, of which probability is a measure – or a combination of these. Thus, if we have no additional information with respect to the process itself, we cannot infer from the probabilistic results what kind of process we have on our hands. But at a more theoretical level, the reason why a definition of chance cannot be extrapolated from probability is because they are not synonymous. Normally probability is defined as measuring chance – but this does not necessarily mean that a definition of chance can be extrapolated from it.

1.2.4 Splicing process is biased, which is to say, regulated

Let's go back to Schumucker's (2007) paper. The author stresses that "assuming that all exon combinations are possible, this allows for the generation of up to 36 096 [38 016] different coding mRNA" (p. 915) and that the process responsible for this mRNA synthesis (i.e., the splicing) is spatially and temporally regulated, even if the extent to which this regulation takes places is unknown (see p. 916). Very curiously, when Schumucker refers to this hypothetical "regulation", he writes about a "bias" which regulates the process of splicing. Usually, "bias" is spoken about in terms of a process that is influenced by factors that are not part of the process, but that are inherent to the entities that participate in it. For example, when we say that a coin has a bias, we are referring to the fact that its physical-geometric conformation is of a certain type and does not allow for a 0.5 chance of having heads or tails. In this example, the bias refers to the coin. But in the case cited, what exactly is "biased"? It seems that the author is referring to the fact that the process itself is biased. What kind of process are we talking about? Are we

talking about an almost deterministic process which is "biased" and therefore acquires a stochastic component? Or, rather, are we talking about a stochastic process that gives equiprobable results, but since it is biased, the property of equiprobability is lost?

Schumucker does not answer these questions directly but wonders what kind of regulation, that is to say what kind of bias, we can have in the case of splicing. He proposes two possible scenarios to explain the relation between the role of this regulation and the diversity of Dscam receptors produced: (1) the first asserts that the way in which alternative splicing is regulated provides the production of specific mRNAs sequences different from cell to cell; (2) the second proposes that, even if alternative splicing is a regulated process, it nevertheless gives rise to the stochastic expression of mRNA synthesis which is similar for all cells. In other words, the first scenario implies that *between cells*, the possible combinations of mRNAs are not equiprobable because each cell is regulated differently; the second scenario implies, on the contrary, that these possible combinations are equiprobable *between cells* because the type of regulation is the same for all cells. Nonetheless, in proposing these two scenarios, the author refers to the *results* of the process,[227] interpreting them differently (equiprobable, not equiprobable) depending on the type of regulation the process has undergone. This could undoubtedly be helpful, for his task is to understand what kind of regulation, that is to say "bias", could influence the products of the splicing process. However, regardless of which of the two scenarios is correct, it is not helpful for my task, which is instead to develop a precise definition of stochasticity with respect to the splicing *process*. Indeed, starting from the (even partial) results of a process, we cannot deduce whether stochastic or non-stochastic processes produced them. Perfectly deterministic processes could lead to stochastic results and stochastic processes could lead to perfectly deterministic and therefore reproducible results (e.g. see Bertin 2012; Franceschelli 2012; Glennan 1997; Werndl 2009, 2012, 2016). So even in this case, Schumucker's analysis does not suffice for going deeper into the stochastic nature of the splicing process.

[227] Here I must specify that these results are partial because they refer to which mRNAs are produced and not to which isoforms of the Dscam receptor will then be synthesized (I remind the reader that mRNA is translated into protein through a process called translation).

1.2.5 Splicing process is biased, that is to say it depends on biochemical affinities between molecules

In a more recent work, Schumucker and other biologists specify what a splicing process *is not*. Indeed, Sun *et al* (2013) assert that alternative splicing is not a process that produces equiprobable results[228] since this process implies "stochastic splicing events" are biased by interactions between RNA-elements and splicing factors. These interactions follow different probability distributions that depends on biochemical affinities between molecules.[229] They call this kind of stochasticity a "biased stochasticity", referring to the authors (Neves *et al* 2004) who first used these terms. More specifically, Neves *et al* (2004) write that:

"the *Dscam* mRNA profile of a given cell arises from a series of stochastic alternative splicing events for each *Dscam* transcript. The probability of sectioning each individual alternative exon *is a function of the splicing factors expressed by each cell type*" (p. 244; emphasis added).

Both Schumucker and Neves share the same general idea that the splicing process is a stochastic process that is biased. But one must be careful because the term "biased" now no longer refers to the fact that splicing is regulated in some (unknown) way. Here, it is indicated instead that splicing is constrained by the chemical-physical affinity between different molecular components (e.g. mRNAs and splicing factors) involved in it.

1.2.6 In summary, the meaning of stochasticity concerning splicing is heterogeneous and unclear

The analysis just provided shows us that the significance of stochasticity with respect to splicing is heterogeneous and often unclear. The critical work just proposed has been particularly strenuous. Biological publications usually lack direct references or conceptual reflections to help with orientating the meaning of certain notions. The concept of stochasticity in reference to splicing is no an exception. Despite this, I have made more evident the problems related to chance and splicing. First, when stochasticity in the splicing process is conceived as

[228] In order to highlight how the "Dscam gene locus was almost five times larger than what may be considered as necessary" (Jin *et al* 2013, p. 2034) Hattori *et at* (2007) postulated that the potential isoforms are randomly sampled with equal probability. Jin *et al* write that this is preliminary assumption by Hattori *et al* (2007) is an "oversimplification" (*ibid.*). The results cannot be randomly sampled with equiprobability because they depend on the interactions between biomolecules.

[229] This kind of dependence recalls what Kupiec (2019) writes about the fact that probability does not come from "heaven": "[Stochastic processes are] governed by probabilities that do not fall from the sky. They are constrained by the material elements that participate in the stochastic process" (note 2, p. 23; on this point see also Kupiec 2019, pp. 118-119).

noise, it is not considered functional. Second, stochasticity in splicing conceived as noise is not even considered as a real event. Third, in the context of splicing, the term "biased" can mean either that the process is regulated in some way, or that it depends on the physico-chemical affinity between the biomolecules involved in the process. In light of these sticky points just raised, the need to make a philosophical work that can help with this heterogeneity is even more evident and urgent. In the next section then, I will try to propose something positive, a definition of chance that can help clarify the notion of stochasticity in the process of alternative splicing.

1.3 Extrapolating the concepts of unpredictability and causal dependence from biological practice

The work with which I would like to begin proposing my idea of stochasticity in splicing is by Hiesinger and Hassan (2018). In it, the authors propose an interesting conceptual differentiation between stochastic processes and noisy processes. In order to avoid any confusion between the meanings of stochasticity, stochastic explanation (SE) and noise that I provide in the previous chapters and the meanings that these authors propose, when I talk about Hiesinger and Hassan's ideas, I will add a double "H" (i.e. HH stochastic process; HH noisy process). An HH stochastic process is, therefore:

"[a] process in which at any point in time the precise state or value is […] unpredictable and neither depends on […] previous values or states of the system, nor on […] an inherent bias of the system." (p. 578).

An HH noisy process is:

"a process in which at any point in time the precise state or value is unpredictable. However, in contrast to a truly random [HH stochastic] process, there might be a dependency on previous states of the system or an inherent bias[230]" (*ibid.*).

They propose an example of an HH stochastic process: the meiotic segregation of Y and X chromosomes in which the 0.5 possibility of having a male or a female depends on processes that give equiprobable results (see Avner and Heard 2001; Heard 2005). Nonetheless, this example is doubtful because no empirical evidence exists that having a male or a female actually

[230] Note how in this case "bias" is described in a different way to sections 1.2.4 and 1.2.5. Here it is something intrinsic to the process itself which renders it in some way dependent on the prior conditions of the system.

depends on processes that can give equiprobable results. In fact, the absence of evidence in favor of stochastic HH processes affects not only meiotic segregation, but most of the molecular/cellular processes:

"Notably, there are factors that bias [the] outcome [of meiotic segregation] and [more generally] *there remains doubt as to what extent biological processes in general are truly stochastic*" (*ibid*.; emphasis added).

By contrast, the two authors are more positive about noisy processes and hypothesize that (some) biological process could be seen as HH noisy. To corroborate their idea, they give an example that can be considered as a HH noisy process: the process of mutually exclusive alternative splicing in Dscam receptor synthesis. They write that in the example of neural wiring, there are no HH stochastic processes, only HH noisy processes. More specifically, they recognize two kinds of HH noisy processes: 1) the phenomenon of filopodial exploration in which the "noisy part" is the fact that the directional growing (e.g. the exploration of a dendrite) is stochastic (i.e. equiprobability in the direction of growing) but it follows precise developmental rules (e.g. self-avoidance; other cell-to-cell communications mechanisms such lateral inhibitions, etc.); 2) alternative splicing, since it is an unpredictable process that nevertheless presents a "dependency on previous states of the system or an inherent bias" (p. 578). Let us focus on the second example of alternative splicing which corresponds to my case study. Even in this case, the authors write that "Dscam are not predictable, but may not have equal probability for all variants" (Hiesinger and Hassan 2018, p. 578), and that "specific [Dscam] isoforms are expressed noisily in a cell-autonomous fashion" (p. 580). To go deeper with respect to the question of splicing as a noisy process, the authors propose the two following notions: 1) causal dependence i.e. "[…] there might be a dependency on previous states of the system" (hereafter CD), and 2) unpredictability, i.e. "[…] a process in which at any point in time the precise state or value is unpredictable" (hereafter U). However, in elaborating these two notions, the authors do not go deeper with respect to their relationship. If the HH noisy process is characterized by both causal dependence (CD) and unpredictability (U), how can these two notions fit into a single defining theoretical framework? The authors do not offer us this kind of clarification and it is up to philosophy to work on the issue. So, let's try to propose two possible relationships between (U) and (CD), one more intuitive and the other more philosophical, and see which works for us. In addition, I believe working on this definition can help bring to light an appropriate definition of stochasticity in the splicing process.

1.3.1 The "intuitive" relationship between causal dependence and unpredictability

The first "intuitive" relationship between (CD) and (U) could consist of saying that the causal dependences driving the (evolution of the) process under attention are obscure for us and this is the reason why we are not fully able to predict them. More explicitly, in this scenario the relation between (CD) and (U) would be as follows: in light of the limits of our knowledge concerning the causal steps of the process (i.e. we do not know all the causal dependencies between different steps of the process), *then* this process is in some degree unpredictable. I would like to specify right away that I disagree with this first intuitive relation because its logic drives us back to a metaphysical idea of determinism. If one day we knew how the causal steps of the phenomenon under attention depend on each other, we would be able to predict its evolution in a precise and complete way, no matter its complexity.[231] Clearly, the possibility of predicting *in principle* contains a strong form of determinism: since the same causes correspond to the same consequences, knowing *all* the former means to having the possibility of predicting *all* the latter.

In this scenario, we therefore postulate that the relation between (CD) and (U) present in the notion of HH noise processes contains a commitment to determinism. Assuming that this notion is used to develop models of certain processes (like splicing), the following question could be asked: would the notion of HH noise processes used to develop that model affect it, not only the way in which we think about prediction, but also in the kind of explanation we can provide? My answer is positive. Indeed, following the deterministic idea, at this point we can only develop incomplete and "palliative" explanation since more detailed information about phenomena *are not yet* available.[232] However, I refute this deterministic picture because it inevitably drives me towards what I call *metaphysically blind alleys*, in which metaphysical implications are hard to resolve with respect to an *epistemological* account of chance in biology, such as the one I am developing in the present work. Therefore, I stress again that: (1) I want to avoid statements on the deterministic or indeterministic status of the biological world; (2) stochastic explanations are not merely "palliative" but proper explanations whose "dignity" is recognized, which is to say, an explanation which has an irreducible epistemic value (cf. stochastic explanation, Chapter 3) and cannot be developed just by waiting for better and more detailed explanations to come along. So, given this first "intuitive" relationship between (CD)

[231] Nonetheless, it should also be noted that very good predictions are also possible using probability (see for example, studies of quantum mechanics).

[232] These metaphysical implications arising from this first relationship between (CD) and (U) have already been addressed in section 3.3 of the previous chapter in which I tried to refute them using different arguments (Cf. Borges' story, Baedke's shift, Woodward's proportional criteria).

and (U) is unsatisfactory, I would like to propose another original relation drawing on, and inspired by, important works in biology and philosophy by Stephen Jay Gould, John Beatty and Derek Turner.

1.4 The "philosophical" relationship between causal dependency and unpredictability

In trying to provide a satisfactory relationship between causal dependence (CD) and unpredictability (U), which might prove useful to clarifying the meaning of chance in splicing processes, I will dive into a well-known debate within the philosophy of biology which concerns Gould's notion of contingency. In fact, Hiesinger and Hassan (2018)'s ideas of (CD) and (U) bear some similarities to a reflection in paleontology proposed by Gould in 1989 which was subsequently reworked by two philosophers, Beatty and Turner, in the 2000s. Although their research subjects are quite different (Hiesinger and Hassan study molecular splicing while Gould reflects on the evolutionary history of living beings), both try to define processes through the use of the two notions of causal dependency and unpredictability that I can generally call stochastic (which Hiesinger and Hassan call HH noisy and Gould, contingent). While I am far from proposing that the paleontological level is extrapolated to the molecular level or vice versa, I remain convinced that this parallel can be useful in finding a satisfactory relationship between (U) and (CD) to allow for the development of an original idea of stochasticity with respect to the alternative splicing process.

According to Gould (1989), the main driver of evolution is contingency. To explain what he means by contingency, he proposes the famous thought experiment: if we could replay a tape of life, the outcomes (i.e. the historical present) would be always different. Therefore, the history of life is contingent. This definition attracted the attention of certain philosophers of biology in the 2000s who tried to clarify some of the ambiguities produced by this thought experiment. In his 2006 analysis, Beatty argues that Gould actually proposes two thought experiments that reflect, in turn, two different notions of contingency:

1. In the first thought experiment, we rewind the tape of life and start it from exactly the same initial conditions as those that actually occurred. In this scenario, Gould hypothesizes that we will have different results (Gould 1989, p. 278). Beatty calls this kind of contingency "contingency as unpredictability", meaning "we get different, unpredictable outcomes from the same indistinguishable priori states" (Beatty 2006, p. 339).

2. In the second thought experiment, we rewind the tape of life and start it from (even slightly) different initial conditions (Gould 1989, pp. 51, 289). In this case too, Gould argues that we will have different results: very tiny differences in the initial conditions could be seen as the cause of the difference concerning the results. Beatty calls this kind of contingency "contingency as causal dependence", meaning later conditions are sensitive to even small differences in the earlier conditions of the history of life.

In more recent years, another philosopher, Turner, has been interested in clarifying Gould's idea of contingency. He tried to delve deeper into the first notion of contingency clarified by Beatty. Turner (2011b) states that Beatty's (2006) notion of contingency "as unpredictability" is "infelicitous" (Turner 2011b, p. 67) since the unpredictability of an event depends on "contingent facts about the human beings doing the predicting, especially facts about their background knowledge, their inferential abilities, and so on" (Turner 2011b, p. 67). What Turner tries to show in this passage, if I understand correctly, is that unpredictability is related to the cognitive limits of human beings rather than being an ontological feature of the world. But Turner has the impression that this is not what Beatty is proposing. Indeed, Turner argues that Beatty "interprets Gould as saying that in many cases, priori conditions do not guarantee the occurrence of any particular outcome" (Turner 2011b, p. 67). In light of this, Turner proposes changes to the notion of "contingency as *unpredictability*" proposed by Beatty to "contingency as *causal insufficiency*". Turner does this to emphasize that we are not talking about a cognitive limitation but rather about a causal insufficiency. In other words, given a set of causes, we cannot be sure it will be the case that particular outcomes will occur. Turner pushes on, confirming with further arguments why causal insufficiency "is not a tantamount to any mysterious sort of indeterminism" (Beatty 2006, p. 345) either. Turner suggests that even if we are not talking about cognitive limitations, we do not have to fall into metaphysical discussions either. Instead, he proposes a middle ground. According to him, saying that the prior states of the entire universe are insufficient to guaranteeing a particular outcome is, no doubt, a statement in favor of indeterminism. However, if we talk about the causal insufficiency of a single phenomenon, no metaphysical notion is brought into play. For example: "the presence of moisture in the soil is not sufficient for a seed to germinate. Here the priori state is just one part of the total sufficient cause[s] of the later outcome" (*Ibid.*). This is to say that certain considered initial conditions X are perhaps not enough to bring about a certain result Y since other necessary causes can play a role. This does not imply any form of indeterminism nor a declaration of unpredictability. It is just a description of a certain state of affairs. I call

this a "middle ground position" which, to sum up, consists of stating that certain clusters of causes would not be sufficient in bringing about the outcomes. We can therefore reasonably conclude that the concept of contingency as unpredictability (intended as causal insufficiency) is neutral concerning (in)determinism (Cf. Turner 2011b, p. 72-73). I propose the following Table 1 to make this idea even clearer:

Metaphysical position	**Middle ground position**	**Epistemic position**
(in)determinism	Causal insufficiency	Cognitive limitations
From an initial condition X, different unpredictable results can occur due to an intrinsic indeterministic feature of the world	Certain initial conditions of X (e.g. certain clusters of causes X) are not sufficient to bring about a given outcome. No metaphysical or epistemic dimensions are considered.	From an initial condition X, different results can occur. These results are unpredictable since the agents are not cognitively able to bring them about.

Table 1 The three positions which concern contingency

The philosophical work I want to propose in the paragraphs that follow uses the relationship between causal dependence and unpredictability proposed by Beatty in the quote below in order to clarify the notion of stochasticity in the alternative splicing process. This work will take seriously the variation proposed by Turner in changing the notion of "contingency as unpredictability" to "contingency as causal insufficiency" (Turner 2011b, p. 67). Let's start by looking at the way in which Beatty conceives of the relation between (CD) and (U) with regard to his notion of contingency:

[2C] "A historically contingent sequence of events is one in which the priori states are *necessary* or *strongly necessary* (causal dependence version) but *insufficient* (unpredictability version) to bring about the outcome" (Beatty 2006, p. 340, emphasis in original).

I first propose declining this definition [2C] by tying it to elements in the alternative splicing case study:

1. In the first part of Beatty's quotation [2C] we can read: "the priori states [of a historical sequence of events] are necessary or strongly necessary". In the case of the explanation

of splicing, this could be linked to evidence that the processes implied in splicing are not entirely random (i.e. they do not give equiprobable results or completely unexpected results) but each step depends, to a certain degree, on the "previous states" of the systems. These previous states could be intended as all those conditions characteristic of the intracellular *milieu*, for example, bio-physical interactions between molecules, micro-environmental conditions, thermal agitation, etc. Recognizing these previous states, and more generally, stressing the relevance of the intracellular *milieu* for molecular processes, highlights that when I state that certain processes are "chancy/stochastic", I am not talking about a complete state of disorder, with anarchic[233] results from processes (e.g. equiprobability results, unexpected results etc.) or processes themselves (e.g. indeterministic processes). Taking the influence of the cellular *milieu* seriously allows us to understand that the probabilities attributed to molecular processes depend on, and are constrained by, the biological possibilities of the cell (e.g. "canalized chance" used by Bravi and Longo 2015, p. 10; Buiatti and Longo 2013, p. 20; see also Moreno and Mossio 2015). For example, within the nucleus, the chance of a transcription factor binding to DNA is greater than zero. In the nucleus, both DNA and transcription factors are present, so their encounter is possible and measurable with probability. By contrast, the chance of a transcription factor meeting a protein in the cytoplasm is almost zero (unless the transcription factor is transported out of the nucleus, but that is very unusual). This is because the membrane of the nucleus provides a biological limit, a constraint (see Umerew and Mossio 2016). In my scenario then, causal dependence resides in the fact that molecular processes depend on, and are constrained by, the conditions in the cellular *milieu*.

2. In the second part of Beatty's quotation [2C] we can read "[These priori states, just mentioned in point 1, are] insufficient [...] to bring about the outcomes". In the case of splicing, this means that the initially considered conditions X are not enough to bring about a certain result Y. It is in this part of the definition that it is necessary to emphasize the importance of the middle position proposed by Turner. Indeed, if we integrate Turner's notion of causal insufficiency, this means that such insufficiency is due not to our temporary ignorance of the causes of splicing, nor to the indeterministic nature of this molecular process. Rather, it is grounded in the assumption that certain clusters of causes can be insufficient in bringing about outcomes.

[233] For a quite different perspective, see Kupiec (2019, p. 233).

More generally, what I am trying to do is reformulate Beatty's [2C] definition with Turner's help, by saying that chance in splicing is related neither to any metaphysical inference nor to any supposed cognitive limitation (cf. Table 1). Let us now try then to reformulate [2C] in order to apply it to the process of splicing itself:

[2C splicing] mutually exclusive alternative splicing is a process in which the priori states of the cellular *milieu* X are *necessary* or *strongly necessary* (causal dependence) but *insufficient* (causal insufficiency) to bring about the outcomes, that is to say the spliced RNAs.

In light of this version of [2C], how shall stochasticity in splicing be conceived? By using the notion, I call "insufficiency of causal dependence". I argue that this synthesizes the concepts of causal dependence and causal insufficiency – developed thanks to Gould's, Beatty's and Turner's work – into a single notion. More precisely, to say that splicing is stochastic means that it is characterized by some insufficiency of causal dependence. For example, prior states of the process such as cellular *milieu* are necessary but not sufficient to bring about the outcome, in this case determine the kinds of mRNAs that are produced.

Again, the value of this conception of stochasticity (in splicing) resides in the fact that it does not necessarily oscillate between cognitive limitations (e.g. since I don't understand what is happening, I use terms such as "stochasticity", "chance") and metaphysical judgment (e.g. biological processes are intrinsically indeterministic). This definition is evidence that a third way can be found, which says that causal dependence is (sometimes) insufficient for explaining the performance of some processes and their outcomes. The advantage of this middle ground (see Table 1) is that, freed from accounts of metaphysical and cognitive limitations, we can ask a new kind of question. Why do we describe splicing with a notion of the insufficiency of causal dependence? For example, it may be due to the fact that certain biologists have considered, and can and/or want to consider, only a certain number of properties in the initial splicing conditions. It may even be that, in some research contexts, it is preferable to choose a few initial conditions that allow the explanation to be cognitively easy to handle.

Where does this analysis leave us? If alternative splicing is described in such stochastic terms, does it really matter for philosophy or biology? Yes, it does. The exploration of new case studies in molecular, cellular and developmental biology is crucial to expanding the study of chance beyond the classic evolutionary framework (cf. Introduction of the present dissertation, section 1). More specifically, this case study allows us to demonstrate how this kind of study can be enriched by considering two previously unremarked planes. The first states that an

interesting reference framework for chance is the study of the physiology of the individual cell which literature shows as having many processes described as "stochastic", even while definitions of this last notion remains most of the time up of the air. Second, the study of splicing has allowed us to understand that studying chance with respect to processes – rather than just results – can be more fruitful in enabling explanation of the behaviors of these processes.

1.5 Stochasticity as insufficiency of causal dependency (ICD) is an instance of my general concept of chance as present contingency (CPC)

The reader may wonder about the connection between this specific definition of alternative splicing and the definition of chance as present contingency (CPC) that I developed in the first part of my dissertation. Indeed, both definitions refer to biological processes, but it is not obvious how they are related. Speaking generally, I can affirm that the insufficiency of causal dependence is a particular instantiation of CPC. If CPC specifies that a process is described as stochastic if it can be otherwise, and the insufficiency of causal dependence specifies why the splicing process can be otherwise, then the notion of insufficiency of causal dependence is an instance of CPC. To unpack this connection even further, we must say that in the case of splicing, the insufficiency of causal dependence refers to the fact that this molecular process can behave one way, or in other possible ways. This behavior is also reflected in the results of the process, since mRNA1 and/or mRNA2 and/or mRNA 3 and so on can be produced.[234] The insufficiency of causal dependence thus reflects the fact that the process can occur in cumulative and disjunctive ways.

Let's spend an additional moment on what I mean when I claim that chance as an insufficiency of causal dependence "instantiates" CPC. If we take CPC as a general definition of chance that attempts to avoid a commitment to ontology, the interpretation of probability and cognitive limitations, the specific definition of stochasticity in splicing demonstrates how CPC can be realized in the description of a specific process. The general definition of CPC can be made more substantial with notions that arise from the careful analysis of cellular and molecular processes. Metaphorically, it is as if by studying a specific case, CPC fits into the processual details and gives way to a regional definition of chance – a notion that specifically applies to that process. Indeed, we have to consider how each individual study of stochastic processes in biology can bring about a specific definition of chance, each of which is one of the possible instantiations of CPC.

[234] Note that an alternative splicing process can produce a single mature mRNA and/or more than one.

2. Ontological and epistemological (anti)reductionism: a clarification

The kind of reflection proposed in the previous sections implies an epistemology that understands chance as irreducible to physics and an active element in the explanation of biological dynamics. However, the discourse on reduction and reductionism has not yet been sufficiently addressed in the present work. In the previous chapter I wrote that from an epistemological point of view, chance should not be reduced to a physical explanation but that this does not prevent the possibility of recognizing that, ontologically speaking, chance is reducible to physical dynamics. In the following section, I try to be more specific with respect to this point. Specifically, I focus on what it means to be epistemologically anti-reductionist while, at the same time, maintaining a commitment to ontological reductionism concerning chance in molecular, cellular and developmental biology.

Very generally, I am in line with Putnam's well-known work (1973) in which he discusses higher-level and lower-level laws. A reductionist assumption says that lower-level laws can be used to deduce higher-level ones. Putnam disagrees because he is convinced of precisely the opposite, that it is not possible to deduce "higher-level" laws from "lower-level" ones. Why so? He justifies this by talking about explanations: by reducing explanations to "lower-level" ones, we lose important structures from the first, "higher-level" ones. Even if he does not specify in detail what he means by these "structures" (p. 136), we can speculate that he is referring to how, for instance, if we try to explain evolution by looking at individual level dynamics, we lose the "structures" of the theory of evolution, including fundamental concepts such as natural selection and drift, as well as the phenomena of multiple realizability that enables us to explain why, even with different processes, we can nonetheless have the same results.[235] In this section, I want to extend Putnam's reflection. Indeed, I maintain that it is necessary to foreground an epistemological (anti)reductionism precisely because "higher-level" explanations can have different structures with respect to "lower-level explanation". In the specific context of my account of stochastic explanation, I argued in Chapter 5 that higher-level stochastic explanations in biology are preferable to nearly-deterministic, finer grained physical explanations that share the same *explanandum* (cf. previous chapter, section 3.3). This clearly appears to be epistemological anti-reductionism in that it denies the possibility of reducing

[235] "[The reductionist] claim that higher-level laws are deducible from lower-level laws and *therefore* laws are *explainable* by lower-level involves [two] mistake[s]. It involves neglect of the *structure* of the higher-level explanations which reductionists never talk about at all, and it involves neglect of the fact that *more than one* higher-level structure can be realized by the lower-level entities [i.e. the multiple realizability argument against reductionism, for a critique see Sober 1999]" (p. 135-135; emphasis in original; see also Putnam 1973, 1975a, 1975b). Furthermore, Putnam reinforces his argument with the famous example of the peg and holes (e.g. see Putnam 1973, section 1; 1975a pp. 295-296).

biological explanations involving chance to mere physical explanations. But many questions still need to be answered. What does "epistemological (anti)reductionism" mean exactly? What is my position with respect to a hypothetical ontological reductionism? Why do I choose to stick to certain kinds of reduction rather than others? In what follows, I propose answering these questions through a brief sketch of the history of reductionism and by pointing out more precisely what kind of reduction and reductionism my account can deal with.

First, in the wake of Kaiser (2012; 2015, p.10, p. 84), I underline how Nagel's original ideas about reductionism no longer present a good way to approach reduction, and how an "epistemic transparency" is actually necessary. Or, to quote Kaiser, "a descriptive correspondence between philosophical theories about science and scientific practice" (Kaiser 2015 p. 10 that refers to Love 2012, p. 179).[236] Second, I briefly describe the main kinds of reductionism discussed in the literature, underlining how ontological reductionism is the least problematic kind of reduction with respect to biology. Third, I describe the way in which Kaiser develops her idea of *reductive explanation*. Finally, using these preliminary points as a basis, I specify in more detail my own epistemological anti-reductionist position concerning chance in biology.

2.1 A few notes about the history of reductionism

As clearly pointed out by Kaiser (2015, p. 43), we can locate the start of the debate on reductionism in philosophy of biology in the 1960s-1970s, and the effort to apply Nagel's classical model of theory reductionism (1961) to biological sciences[237] (see e.g. Schaffner 1969). Nagel's model is a formal model of reduction inspired by the DN (deductive-nomological) model of explanation (Hempel and Oppenheim 1948). For him, reduction is a special case of explanation. If in the DN model the *explanandum* is deduced from the *explanans*, in Nagel's account, theory A is deduced from theory B. Very roughly, the two core assumptions of his account are: (1) the units of reduction are theories; (2) reduction relations are the relations of logical derivations. One of the first attempts to apply Nagel's account in biology was the effort to reduce classical genetics to molecular genetics (see Hull 1974; Rosenberg 1985; Waters 1990; see Kaiser 2015, p. 67 for an overview). Kitcher (1984) was one of the first philosophers of science to clearly show how Nagel's project to reduce classical biology to

[236] On this point see also Hüttemann and Love (2011). See also Keller (2000b).
[237] It was also one of the debates that delineated philosophy of biology as a distinct and independent discipline (Kaiser 2015, p. 67; Griffiths 2007).

molecular biology was not feasible.[238] But there is a problem underlying this debate that always remains. What do we mean by reductionism? Do we always have the same conception in mind – are we always talking about the same thing?

Ayala (1974) was the first philosopher to distinguish different kinds of reductionism, which until then formed a messy part of the debate. His three kinds were theoretical reductionism, methodological reductionism, and ontological reductionism (see Kaiser 2015, p 66). The first kind concerns theories: theory A can be reduced and explained by theory B. Supporters of this kind of reductionism might in fact advance the idea that classical genetics can be reduced to molecular biology. The second type concerns methodology, the idea that in order to have greater explanatory success, it is necessary to focus and develop explanations at the more fundamental (for instance, molecular) level. An example of the tension between methodological reduction and anti-reduction approaches in biology is the case of cancer cell studies, with two major theories characterizing the field. On the one hand, the somatic mutation theory (SMT) focuses on possible cures at the molecular level by stating that DNA mutations cause the disease. This theory assumes methodological reductionism: successful explanations of cancer cells are explanations at the molecular level. On the other hand, according to the tissue organization field theory (TOFT), the relevant causes of cancer have to be found at the level of cell communication and tissue development.[239] This second approach is typically seen as taking a methodological anti-reductionism position in that it seeks to develop explanations by focusing on levels of biological organization that are higher than molecular ones. Methodological reductionism and antireductionism share a strongly normative aspect: they state how biologists *should* pursue scientific research.[240] The third type of reduction proposed by Ayala (1974) is ontological reduction. Today, in philosophy of biology there is (almost) no ambiguity

[238] Kitcher (1984) argues that since *there are* independent levels of causality, *then* we can develop different levels of explanation too. Ruphy (2016, p. 43-44) raises an objection, stating that Kitcher does not provide valid justification for the ontological-metaphysical claims that *there are* independent levels of causation, and thus falls into a paradox: while he claims that the actual state of science is incomplete and susceptible to change, he *nevertheless* wants to claim, from this state of affairs, a universal ontological statement concerning biological worlds (i.e. the existence of independent levels of causation).

[239] For a philosophical analysis of this example see Kaiser (2015, p. 73, p. 201); for a comparison of SMT with the stem cell model, see Morange (2015, p. 33).

[240] Kaiser (2015, p. 71) writes about the vagueness of methodological reductionism: what does methodological reductionism consist of? She claims, in the wake of Wimsatt (2006, p. 445), that proponents of this kind of reductionism do not specify *how* to pursue it. On the other hand, Morange (2000) stresses how at the edge of molecular biology (1930s-1960s) methodological reductionism was the most common strategy in pursuing biological research. "The members of Rockefeller's foundation [that financed numerous important biological research project between 1930s-1940s] who encouraged so many physicists to study biological problems with methods and techniques derived from modern physics, were also convinced that the reductionist approach was the most efficient for studying biological phenomena" (Morange 2000, p. 246).

concerning ontology. All scholars agree that – for example – molecules, cells, lions, and frogs are objects that exist in the world independently from our minds. Furthermore, in this field of studies, all agree on the fact that all biological objects can be reduced (in the end) to physical objects. This kind of ontological reductionism is called token-physicalism (see Kaiser 2011, p. 457 and Kaiser 2015, p. 52). As Rosenberg (2006) wrote "we're all physicalists now" (p. 4).

With this first philosophical distinction between different kinds of reductionism, it is evident that when we talk about Nagel, we talk about theory reductionism, in which the units of reduction are theories, and the relations between these theories are logical derivations (i.e. deduction). It took a great deal of philosophical work to apply Nagel's model to biology.[241] Nevertheless today, no one uses Nagel's theory of reduction in philosophy of biology since it does not mirror biology in practice – specifically, the reduction (if any) that can be found in biological literature. In the wake of Nagel's theory, numerous works on reduction and anti-reduction in biology were proposed, but until today, no "unitary account of reduction" has yet been established, leaving the situation in a "polyphonic disunity" (Kaiser 2015, p. 81 that refers to Wimsatt and Sarkar 2006, p.697).

Kaiser points out that in addition to Ayala's three types of reductionism (theoretical, methodological, ontological), a fourth kind, explanatory reductionism, is of great philosophical interest. In different parts of her work (2012, p. 464; 2015, p. 192), she writes about two specific kinds of reductionism related to explanation. On the one hand, a reductive explanation could be seen as the relation between two explanations: a higher and a lower level explanation of the same phenomena (see Rosenberg 2006). On the other hand, a reductive explanation could be conceived as a relation between the phenomenon to be explained and the relevant factors necessary for its explanation. Indeed, in this case, "reductive explanations […] bridge at least two levels, namely the higher level of the *explanandum* phenomenon and the lower level(s), on which the explanatory relevant factors cited in the *explanans* are located" (p. 192)[242]. She specifies that this second kind of reductive explanation is the most interesting kind because it is what can normally be extrapolated by focusing attention on biological work. She further specifies how this kind of reduction is located *within* a single explanation, and not between two or more. In other words, she proposes finding the reductive character of an explanation by looking for the relation between the *explanandum* and the *explanans*. If the *explanans* is at lower level than the *explanandum*, we have a reductive explanation.

[241] For a critical analysis of Nagel's model, see Kaiser (2012) and Kaiser (2015, p. 84).
[242] See also Wimsatt (1976 p. 689).

It is interesting to underline that Kaiser's account of reduction does not have reductive explanations fixed on some sort of "fundamental level" generally recognized as the physical level[243] (see, e.g. Rosenberg 2006), but rather depends on the relation between the *explanans* and *explanandum* of the explanation under attention. In ecology, for instance, individual-based models (IBMs) explain the behavior of a community in terms of individual organisms and interaction between them. Kaiser (2015, p. 201) claims that this *is still* a case of reductive explanation, even if we do not find the *explanans* at the fundamental/physical level. With this example, she concludes that the fundamental level of reductive explanation (i.e. physical level reduction) is only one specific case among numerous possible reductive explanations.

The description I have just provided of Kaiser's account is fundamental insofar as it allows us to understand that one kind of reductionism in explanations is *relative*, meaning it does not depend on the presence of a fixed fundamental level but rather on the relationship between the *explanandum* and *explanans* of an explanation.

2.2 Which kinds of reduction can be present in reductionism concerning chance?

In this section, I will examine in depth the two types of reductionism that I think are most relevant for the study of chance in biology – ontological reductionism and epistemological anti-reductionism (the latter mostly refers to explanation). But before proceeding, the question arises, why these kinds of reductionism and not others? (See, for example, those proposed by Ayala 1974, in the previous section.) The answer again resides on the "practice turn" (Ankeny *et al* 2011; Soler *et al* 2014). Indeed, I am in line with Kaiser in saying that the study of reductionism has to be the study of reduction in *biological practice*, which is to say, analyzing which kinds of reduction, if any, biologists embrace in the course of scientific practice.[244] By taking this advice seriously and analyzing scientific work, I realized that if my aim is to provide a deeper philosophical analysis of reduction and chance in molecular and cellular biology,

[243] We have to be careful because between the "physical level" and "molecular level" often hides an ambiguity. When philosophers refer to the fundamental level, they can label it both as *molecular* level and *physical* level (or in certain cases also as bio-physical level and chemical-physical level). We have to be careful about these different labels. Indeed, if one calls the molecular level fundamental, it could be implied that one *still* refers to biology, since most of the time "molecules" is understood and meant as "bio-molecules", which are molecular biology's objects of study. But there is also another possible interpretation. Indeed, it could be the case that when one says "molecular" level, one refers *instead* to the physical level, namely the fact that these molecules, or bio-molecules, could be explained by physical laws. This interpretation can be traced back to Crick who famously wrote that "the ultimate aim of the modern movement in biology is to explain all biology in terms of physics and chemistry" (Crick 1966, p. 10). In light of these confusions concerning the label "molecular level", when I refer to the fundamental level, I will label it only as the "physical level". When I refer to the "molecular level", unless otherwise specified, I still refer to biological objects that can be explained in biological terms.

[244] More strongly, Wimsatt (2007) writes that philosophy itself should not be afraid to do science (p. 26).

ontological reductionism and explanatory (anti)reductionism are the most relevant. My goal, however, is not only to specify the kinds of reductionism relevant to the study of chance, but also which are allowed and which forbidden in a context that seeks to enhance the role of stochasticity in biological explanation. So, a normative component comes into play here – but of what kind? It is an in-practice metanormativity[245] with prescriptive claims: biologists *should* recognize that with respect to the study of chance it would be better to embrace some types of reductionism (ontological) and reject others (explanatory).[246]

Ontological reductionism. As already pointed out in section 2.1, there is not much discussion of this kind of reduction in philosophy of biology. Everybody agrees that all biological phenomena are, in the end, reducible to physical phenomena, and that biological objects are, in the end, physical objects (see Rosenberg 2006; Brandon 1996). In line with Kaiser (2015), I referred to this kind of ontological reductionism as "token-physicalism". In the previous chapter, I specified how I can easily acknowledge that chance in gene expression can be reduced to its physical underlying causes (cf. section 3.2 of the previous chapter). Even if Kaiser remains cautious of the idea that "we're all ontological reductionists now. Case closed" (Kaiser 2015, p. 50) and underlines how the study of ontology is far from complete, nevertheless, this dissertation is focused on stochastic explanation and so I leave ontological questions of reduction and chance in biology to another occasion.

Explanatory reductionism. I objected to this reductionism when I said that, in explaining chance in biology, biological explanations cannot be reduced to physical explanations because they can lose the possibility of seeing chance as an essential explanatory element (see section 3.2 of previous chapter). Nevertheless, this kind of reductionism referred to a single explanation in a category that contains, at least two other sub-types of reductionism. In what follows, I would like to specify whether we are talking about epistemological reduction *within* biology or reduction *between* biology and physics. Note that the two categories I will consider, namely explanatory relative reductionism (ERR) and explanatory chemical-physical reductionism (ECR), should be thought of as two sub-categories within the category just discussed that I have named explanatory reductionism (see Table 2).

Explanatory relative reductionism (ERR). Inspired by Kaiser (2015), I depicted this kind of reduction in the final part of the previous section. I suggest conceiving of this account of

[245] Recall that in-practice metanormativity refers to "[a] philosophical theory about a certain feature or element of science E contains only such normative claims about E that take into account, are drawn from, or *are informed* by factual claims about E" (Kaiser 2019, p. 44; emphasis added). For further detail, see Chapter 1, section 2.3.
[246] That said, I do not deny that other types of reductionism may be relevant to the study of chance in biology or more generally to analysis of the different types of reductionism that we can extrapolate from scientific practice.

reductionism as *relative*. In this way, reduction no longer depends on a specific fundamental level of explanation or biological organization (e.g. the physical level), but instead depends on the relation between *explanandum* and *explanans* in a single explanation. The only difference between ERR and Kaiser's proposal is the fact that the first takes place within the domain of biology. Indeed, Kaiser's account of reductive explanation (sketched in the final part of the previous section) identifies the reduction as being the relation between the *explanandum* and the *explanans*, however, she does not specify "where" the *explanandum* and the *explanans* fall into concerning the scientific field. I would therefore like to add their "localization". ERR is a reduction within a *single* explanation of a *single* field of study (i.e. biology; see after Figure 3). If in the previous chapter I defended a general idea of epistemological anti-reductionism, now that we have this kind of reductionism clear, I can say that my account of stochastic explanation is actually compatible with ERR. Referring to the examples that I mentioned in the previous chapter, cellular differentiation (i.e. the *explanandum,* at the cellular level) can be accounted for by the phenomena of lateral inhibitions (i.e. the *explanans,* at the molecular level), in this case, the interaction between the receptor Notch and the ligand Delta (see section 3.2 of previous chapter). More specifically, stochasticity at the cellular level (i.e. stochastic diversification of equivalent cells, see Figure 3 in the previous chapter) is explicable by molecular processes. In this case, we can see how the *explanans* is at a lower level as compared to the *explanandum*. Therefore, we have here reduction in terms of ERR. Furthermore, in Figure 5 of the previous chapter (section 3.3), we argued in favor of explaining an *explanandum* such as (B) "variability in cell fate and cellular differentiation", with an *explanans* such as (C) "stochasticity in gene expression". Even this explanation can be conceived as another example of reduction in ERR terms because the *explanans* (C) is at a lower level compared to *explanandum* (B). Protein concentration refers to the molecular level, whereas the variability of cell fates refers to the cellular level. What I am trying to underline here is that the examples of SE I used in the previous chapters can admit ERR, (i.e. explanations that have an *explanans* at a lower level of biological organization than the *explanandum*). A stochastic explanation (SE) characterized by an explanatory relative reductionism (ERR) is only possible if the following two necessary and sufficient conditions are met: 1) the *explanans* must belong to a biological discipline (e.g. molecular, cellular, developmental biology, etc.); 2) the *explanans* – even if located at a lower level than the *explanandum* – must have a chance element in it (e.g. CPC). If these two conditions are satisfied, the explanation will be stochastic and the role of chance is preserved.

ERR debunks the general idea that, in today's biology, *we must all be anti-reductionists*. This general idea of anti-reductionism, which to counteract the genetic, molecular and informational reductionism that since the 1940s has exploded in molecular biology, must be clarified in each specific context in order to understand exactly which kind of reductionism we are against (cf. Kaiser 2015, p. 44). ERR shows us that while not all types of explanatory reductionism should be forbidden, a more pointed analysis of a reductionism within biology is necessary and perhaps helpful in further clarifying the role of chance in molecular and cellular biology.

Explanatory chemical-physical reductionism (ECR). I use this term to refer to a reductionism structured by an *explanandum* that refers to a biological domain and an *explanans* at the physical/fundamental level. Furthermore, it refers to the assumption that the *explanans* at the physical/fundamental level is the best for explaining the *explanandum* at the biological level.[247] In this case, we are actually considering an absolute level of reduction. This level is understood as "fundamental" and generally associated with the physical-atomic level. We have then a reduction between biology and physics, two different domains of research: I explain a biological process X with a physical explanation Y. I do not want to say that I am against ECR *in general* – indeed I leave open the possibility of explaining some biological dynamics with physical processes (although I believe it to be problematic). However, I specifically argue against ECR in the case of the study of chance in molecular and cellular biology. I cannot accept that the stochastic descriptions of certain biological processes can be reduced to mere descriptions of physical dynamics. This is because placing the *explanans* at the physical level often leads to the loss of the stochastic element present at the biological level which, I argued, is explanatory. For instance, if I explain the spherical shape of cell wall through the random rearrangement of microfibrils, I use an explanation with CPC in the *explanans* (cf. Chapter 3). In this case, the explanatory role of chance is conserved and we talk about a stochastic explanation (SE). By contrast, if I insist on developing an explanation that aims to describe all the intricate causal chains (or better, intricate webs) that occur at the physico-chemical level in the development of a single microfibril, the biological role of stochasticity is lost and we return to a deterministic approach in which explanation is either developed in a deterministic fashion or a probabilistic explanation in which chance is described as physical phenomenon (i.e. if all the causes can be described, everything about the phenomenon will be known). In Figure 3 below, I provide a visual representation of these two latest kinds of explanatory reductionism:

[247] One of the historical supporters of this kind of approach is Rosenberg (2006) who, even if he does not speak explicitly about the physical level, still supports a molecular reductionism that comes very close.

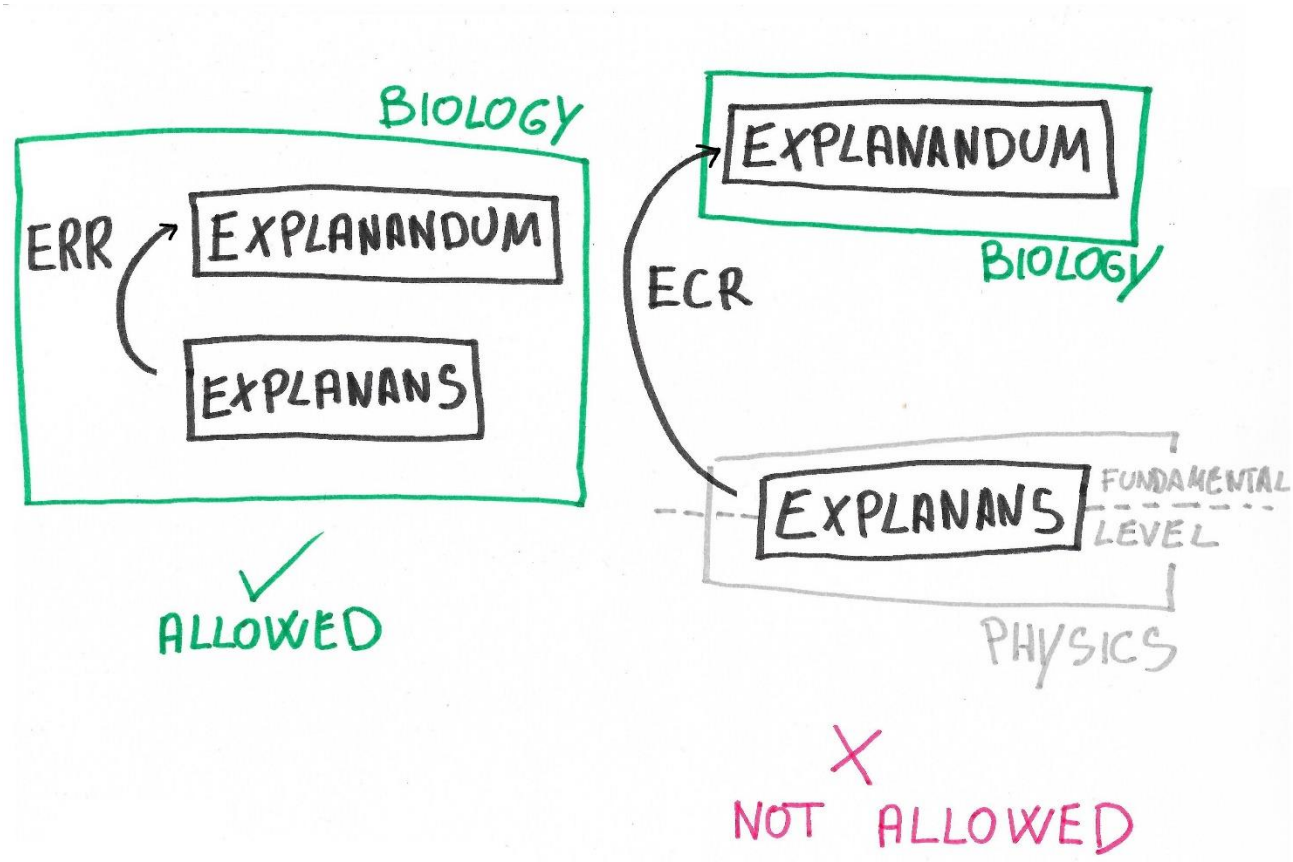


Figure 3. A visual representation of the two kinds of explanatory reductionism proposed.

On the left, we can see ERR in which the *explanans* at a lower level of biological organization explains the *explanandum* at a higher level. This kind of reduction within biology is allowed in my epistemological account of chance. At the right, we can see ECR in which the *explanans* at the fundamental/physical level explains/account for an explanandum at higher and biological level. This second kind of reduction between different disciplinary domains is not allowed in my epistemological account of chance.

2.3 The distinction between the accounts of ERR and ECR matters

I would like to address two additional advantages in highlighting these last two categories of explanatory reductionism, one with respect to the discussion of reductionism in general and the other with respect to my account of stochastic explanation more specifically.

First, the theoretical advantage of developing ERR, echoing Kaiser (2015), is to characterize reductionism as a relative rather than an absolute property. If the intuitive idea of reductionism is indeed that any kind of biological phenomenon has a physical explanation, ERR testifies to the contrary: reductionism is relative, not absolute, because it depends on the relation between the *explanans* and the *explanandum*. If the *explanans* is at a lower level compared to

the *explanandum*, we have reduction in an ERR fashion; if we don't, we do not. By contrast, Rosenberg (2006) collapses ontological and epistemological reductionism (in my terms, ontological reductionism, ERR and ECR) into ontological reductionism because, he says, their difference is an "unstable equilibrium" (p. 7).[248] Therefore, in his view, ontological reductionism reduces both biological objects to physical objects, and biological explanations to physical explanations. But flatting all these kinds of reductions in this way prevents distinctions between when the reduction is between explanations (representations of the world) or ontologies (different kinds of entities in the world). Nor does it permit recognition of important features of reduction – for example, when we are referring to reduction between the *explanandum* and the *explanans* in a single explanation within a single field (see Figure 3, on the left), or when we are referring to reduction between the *explanandum* and the *explanans* in a single explanation but in different fields (see Figure 3, on the right).

The distinction between the ERR and ECR categories helps me clarify which kind of reductionism is permitted and which not in my epistemological account of chance. As I specified above, I am open to ERR but argue against ECR. Table 2 below shows which categories are of interest to my project, their definitions, and my normative judgments concerning them with respect to case studies in molecular and cellular biology.

[248] For a discussion see also Brandon (1996).

Kinds of (anti)reductions		Definitions	The relation with chance	Examples
Ontological reduction		Reduction between phenomena/objects. For instance, biological phenomena/objects can be reduced, in the end, to physical phenomena/objects.	I do not argue against the idea of reducing chance at the biological level (e.g. molecular and cellular level) to its physical causes (e.g. thermal agitations, quantum effect).	Chance in gene expression is caused by thermal agitation and quantum effects.
Explanatory reductionism	**Explanatory relative reductionism (ERR)**	Reduction within single explanations in a single field of study (i.e. biology). More specifically, the reduction resides in the relative relation between the *explanandum* and the *explanans*.	I can agree with reduction in the ERR sense. Indeed, a stochastic element can be located in an *explanans* that is at a lower level than the *explanandum*.	*Explanandum* (at the cellular level): cellular differentiation/heterogeneous cell fates *Explanans* (at the molecular level): lateral inhibition, i.e., interaction between the receptor Notch and the ligand Delta
	Explanatory chemical-physical reductionism (ECR)	Reduction that is established within single explanations, but between different fields (i.e. biology and physics). For instance, a biological *explanans* has to be explained by a physical *explanandum*.	My position against ECR applies when chance has an explanatory role in biological explanations and the desire is to reduce them to physical explanations.	*Explanandum*: chance in gene expression *Explanans*: physical noise

Table 2 The different kinds of reduction I developed in the present section. For further details. see the main text.

With this differentiation of possible types of reductionism, we also concluded that my account of stochastic explanation (SE) can admit reductionism in terms of ERR but categorically rejects all types of ECR reductionism.

Conclusion

In the first part of the present chapter, I explored the case study of neural wiring, in which a particular case of splicing – mutually exclusive alternative splicing – takes place. Splicing is fundamental for the production of a diversity of Dscam receipt isoforms. This receptor is a key element in the development of neurons. In this scenario, I highlighted that how this kind of splicing is conceived as stochastic is not clear in biological literature. In light of this, I proposed an original analysis with the aim of developing a specific idea of stochasticity for this kind of process. I proposed that stochasticity in the splicing process can be defined as "insufficiency of causal dependence" – which is a specific instance of my general definition of chance as present contingency (CPC). This notion strongly resists the tension whereby the unpredictability of the splicing process is viewed as a result of cognitive limitation or as indicating a form of indeterminism inherent in the process itself.

In the second part of the chapter, in line with Kaiser, I articulated reductionism through three different categories: ontological reductionism, explanatory relative reductionism (ERR) and explanatory chemical-physical reductionism (ECR). I specified the relationship between these categories and my account of stochastic explanation. In fact, I argued that there is no philosophical problem in recognizing that, ontologically speaking, biological chance is eventually reducible to physical phenomena such as thermal agitation and quantum effects (i.e. allowing ontological reductionism). In addition, an important result of my analysis is to have recognized the possibility of being reductionist, at least in the ERR sense. Indeed, many of the examples I proposed in the previous chapters, even if they tend to enhance the explanatory role of chance in molecular and cellular biology, are examples of ERR in which the *explanandum* is accounted for by elements at a lower level of biological organization, such as the case of explaining cell differentiation through the molecular dynamics of lateral inhibition. Finally, I showed how my epistemological account of chance is incompatible with reductionism in terms of ECR, which tries to explain chance in the cell through exclusively physical categories. This is because it places the *explanans* at the fundamental/physical level and consequently loses the explanatory chance element which characterizes stochastic explanations and the higher, molecular, biological level.

Chapter 7: Dissecting chance in translation by analyzing stochastic processes through models of time

Introduction

In this chapter I will focus on one last biological process at the level of translation in gene expression: alternative start-codon selection. If in the previous chapters I favored a conceptual analysis of the meaning of chance – for example in the splicing process – in the final part of this thesis, I propose a different approach focused on the models used to describe and explain the case study of alternative start codon selection. More specifically, I will open up discussion on the parameters present in these models, in particular time, to show how analyzing these parameters can greatly help in understanding where stochasticity intervenes in the explanation of the phenomenon under study[249] and how it can even be essential to account for it. A process described as stochastic can be characterized/modeled by using many different parameters (including time, space, energy, etc.) and, depending on the parameter under consideration, we can identify different ways in which chance plays a role in the explanation of this process. Why then do I choose time?

Intuitively it is possible to say that chance is a property of a description of certain processes *which take place in time* (cf. Chapter 3). Each process is composed of events, each of which occurs over a finite interval of time (see Huneman 2017). So, it seems like a good idea to focus on this parameter – but as a reason, it does not feel sufficient. Indeed, can I continue to ask why I chose this parameter rather than space, energy, or any other? Cellular biology, especially molecular biology, often describes and explains processes though complex colored drawings, without making explicit the temporal elements present within them. For example, the drawings that we can find in any molecular biology textbook rarely include a precise temporal framework. At the very most, temporality is mentioned glancingly, without a careful analysis of how the temporal dimension may influence the very dynamics of the processes themselves. If a reflection on time could be generally useful in the clarification of many of these processes, it acquires additional, particular relevance when we speak of processes described as stochastic. The temporal parameter could in fact be useful in specifying in what sense certain processes can be conceived as stochastic.

[249] To be clear, when I say that stochasticity "intervenes" I do not mean directly on the phenomenon but in the models and parameters developed to explain this phenomenon. My entire dissertation calls for an epistemological investigation and this final chapter is no an exception.

This case study also aims to make my analysis of chance at the cellular and molecular level more comprehensive. In Chapter 5, I focused on transcription; in Chapter 6 on alternative splicing. I now finish the picture by analyzing one of the final phases of gene expression, translation. Obviously, this conceptual map of chance in molecular cellular biology (cf. conclusion of the present dissertation) should not be understood as exhaustive but only as the beginning of a reflection that could be further extended.

1. The context of my epistemological analysis

Since the 1990s, researchers on developmental noise have mainly focused on the process of gene expression. Let us recall that, in this context, noise is operationally defined as random variation around the mean value of a given gene expression parameter. It constitutes deviation from the average amount (number or concentration) of transcripts/RNAs or polypeptides/proteins produced, in a cell in a homogeneous environment over time or in a population of isogenic cells in a homogeneous environment at time t (cf. Chapter 5; see also Levine *et al* 2020). Noise in gene expression is thus defined in quantitative terms, by referring to the way it is measured.

Particular attention has been given to analysis of the quantitative effects of stochasticity in the first step of gene expression (i.e. transcription), which Chapter 5 dealt with. At the same time, it has been highlighted that these quantitative effects can in turn have qualitative effects at higher levels of biological organization – from the cellular level, where stochasticity can affect cell fate (Maamar et al 2007, Losick & Desplan 2008), to the organism level, where it can have an impact on, for instance, morphological traits (Raser and O'Shea 2005). A few studies have dealt with stochasticity at the translation step and, more specifically, with how stochasticity can also have qualitative effects at the molecular level of gene expression, for example, stochasticity can explain the production of different kinds of proteins. Since transcription and translation processes share similar physico-chemical features,[250] there is no *a priori* reason, theoretically speaking, that stochasticity could not influence both. This is why, in this chapter, I have decided to focus on the process of translation. Moreover, there is no reason either to exclude that stochasticity can produce both quantitative and qualitative differences in the outcomes of these two molecular processes. This is why, in what follows, qualitative effects form the core of my epistemological investigation.

[250] Indeed, these two steps of gene expression are processes in which macromolecular protein base complexes interact in their cellular environment with nucleic acid polymers (RNA and or DNA) in extended form.

Actually, there are already a number of studies on the quantitative effects of stochasticity in the step of translation (e.g., see Salari *et al* 2012). By contrast, very few analyses have looked at how stochasticity can affect translation in qualitative terms. Why so? Several reasons can account for this and I highlight just two of them. Firstly, it is a technical issue: comparatively fewer techniques exist to measure stochasticity in translation compared to in transcription (though it must be added that in recent years, important new techniques such as ribosome profiling have been developed, see Ingolia 2014). Second, translation is classically conceived as a deterministic or nearly deterministic process that faithfully produces the polypeptide sequence corresponding to a given transcript. Putois *et al* (2015) write that translation is usually seen as "a linear conversion with one predictable, unambiguous outcome" (p. 1). Such a view makes it difficult to even raise the question of whether translation from a given transcript allows for heterogeneous outcomes in terms of the type of product produced (Pilpel 2011). Despite that, and thanks to recent technological innovations (e.g. see Blake and Wu 2019), new research has begun to explore the qualitative effects of stochasticity in translation. In particular, Boersma *et al* (2019) elaborate their single mRNA images by means of a statistical analysis to show that the synthesis of alternative proteins from the same mRNA reflects an "inherent stochasticity" (p. 471). More specifically, they show that stochastic events in translation, called alternative start-codon selection events, influence the diversity (type) of polypeptide sequences produced. We want to take this important work into account for two reasons: 1) it is one of the first studies showing that stochasticity can have a direct qualitative effect at the molecular level; 2) it focuses on the qualitative effects of stochasticity in translation, to go beyond the classical idea of one-transcript-one protein. For these two reasons, this kind of study presents a new and exciting research direction that calls for philosophical investigation. What are the ways in which stochasticity intervenes in the description of this crucial step of gene expression? Does stochasticity have, in this context, an explanatory power?

The general aim of this chapter is to disentangle the various ways in which stochasticity can be studied at the molecular level of gene expression, considering not only its quantitative but also its qualitative effects. More specifically, my aim is to show that stochasticity can intervene differently in the various parameters (e.g. time, see after) used to describe/model gene expression and, depending on how these parameters are represented in the models, it can have qualitative consequences at the molecular level. In this way, it is possible to show in a (more) precise way the need for stochastic explanations in order to account for this molecular process.

With both our general and more specific aim in mind, this chapter is structured as follows. In the first section, I introduce a distinction between two sorts of stochasticity. I define

them in terms of the sort of effects they produce and I respectively call them "quantitative stochasticity" (QTS) and "qualitative stochasticity" (QLS). In the second section, inspired by the pioneering work by Boersma *et al* (2019) (see also Brunet *et al* 2018, Lyon *et al* 2019), I introduce our case study of alternative-start codon selection events. These events represent noncanonical translational events (see e.g. Firth and Brierley 2012) that have been shown to be particularly influenced by stochasticity. I use this case study since qualitative effects are involved here: stochasticity is at the origin of switches in translation initiation from standard to alternative ORFs – which share the same localization in the genome but give rise to the synthesis of different polypeptides. In the third section, I use two simple time models in order to show two ways in which stochasticity can operate in alternative start codon-selection events, and, more generally, in translation. In the fourth section, I draw the consequences of these models in terms of the temporal dynamics of protein production in order to show why the various ways in which stochasticity can intervene in the explanation of translation processes are experimentally relevant. In this context, I open up discussion about whether stochasticity can have an explanatory power in the explanation of this molecular process and suggest arguments in favor of an affirmative answer. In the fifth section, I will distinguish and elaborate on the philosophical consequences that my work in this chapter has brought to light.

Although I analyze a specific event in translation (alternative start-codon selection) as regards a specific parameter (time), my grid to analyze stochasticity might prove useful in opening further reflections and analyses on stochasticity in gene expression more generally. In particular, it could help with exploring the different ways in which stochasticity intervenes in models of transcription, far beyond the current focus on stochasticity's quantitative effects at this step of gene expression. My framework is also relevant to study of the stochastic character of other parameters used to model the steps of gene expression, such as energy and space.

2. Quantitative (QTS) and qualitative stochasticity (QLS)

In the biological literature on "developmental noise", the term "stochasticity" refers to fluctuations affecting the molecular and cellular processes involved in cell development. In other words, the continual and unstructured changes of the values of variables used to describe these processes. Two sorts of stochasticity should be distinguished in terms of the possible effects of these fluctuations on the processes affected – in particular gene expression. On the one hand, I suggest using "quantitative stochasticity" (QTS) to refer to stochastic fluctuations affecting the *amount* (number or concentration) of products produced (in the case of gene expression, the amount of genetic products, namely transcripts and/or polypeptides sequences

synthetized from the same DNA sequence). On the other hand, I suggest using "qualitative stochasticity" (QLS) to refer to stochastic fluctuations which affect the *type* of product produced (in the case of gene expression, the type of transcripts and/or proteins produced). Note that this is a conceptual distinction, which means that, empirically speaking, QTS and QLS mostly coexist, at the same or different levels of biological organization, and can even influence each other.

An example of a phenomenon already developed in Chapter 5 which concerns both QTS and QLS is the switch between states of competence and noncompetence in the bacteria *Bacillus subtilis*. First, stochastic fluctuations in ComK protein abundance are the quantitative effects of stochasticity at molecular level. Second, the switch between the state of competence and noncompetence is a qualitative change, as long as what varies is a phenotypic trait or behavior (meaning, the capacity of the bacteria to uptake DNA fragments from the environment and integrate them in their own genome). Thus, in this scenario, the quantitative effect of stochastic fluctuations at the molecular level (i.e. QTS) produces, in turn, a qualitative effect at the cellular level (i.e. QLS).

Although the relationship between QTS and QLS can be complex and complementary, I want to focus attention mainly on QLS in translation. This is because my aim is to highlight the qualitative effects of stochastic fluctuations at the interface between the RNAs and polypeptide produced. Paying special attention to time, I dissect two of the ways in which stochasticity can intervene in this parameter, by showing its possible effects on cell activity. Finally, I argue that QLS can *itself* be seen as explicative, that is to say it can be conceived as an effective element in the context of certain biological explanations (cf. section 1).

Several "noncanonical translational events" (see, e.g., Brunet et al 2018; Lyon et al 2019) can be said to be affected by stochasticity, and in particular by QLS. These include alternative start-codon selection, frameshifting, stop codon readthrough, and internal ribosome entry site initiation (IRES) (Blake and Wu 2019, p. 3; see also Sonneveld *et al* 2020, p. 609). As already mentioned in section 1, I focus my attention on alternative start-codon selection events because they have been more explicitly associated with stochasticity as a way of accounting for differences in the type of protein produced (see Boersma *et al* 2019; Sonneveld *et al* 2020; Cf. intro). More specifically, five reasons motivate my choice of case study: 1) Alternative start-codon selection concerns translation, a step of gene expression in which stochasticity is rarely discussed from a philosophical point of view but that, by contrast, is emerging as an area of great interest to biologists; 2) As one of the final steps of gene expression, translation can have

a crucial impact on mature products, which few of the other steps are capable of doing.[251] Alternative start-codon selection can be one of these crucial translational events; 3) Alternative start-codon selection is one of the most common non-canonical translational events;[252] 4) While other studies have only put forward hypotheses, a recent experimental study on alternative start-codon selection describes this event as reflecting an "inherent stochasticity" (Boersma *et al* 2019, p. 471), showing that QLS has a direct impact on translation by directly influences the type of proteins produced; 5) The fact that alternative start-codon selection events can be analyzed by focusing on a single cell makes it a particularly suitable case study for analyzing stochasticity across the parameter of time. Indeed, these events naturally lead to being modelled as sequences of events (each event leads to the production of a certain type of protein) taking place over time in a single cell (as compared to a statistical (a-temporal) analysis at the level of cellular populations). Moreover, the temporal analysis of alternative start-codon selection has been an integral part of the demonstration of its stochastic character. Let us now introduce our case study and the various instances of stochasticity affecting it.

2. Alternative start-codon selection events

2.1. Different instances of stochasticity in ORF scenarios

The first thing we need to do to properly understand the dynamics of alternative start-codon selection is define ORFs. ORFs (Open Reading Frames) are sequences of DNA transcribed into mRNA and defined by a start codon (the codon from which a ribosome begins translation), and the first stop codon in the same frame (the codon at which a ribosome finishes translation). These sequences are read by ribosomes, and usually give a single, defined polypeptide. Sequences are termed "standard" or "canonical" because the start/stop codons they are delimited by are unique and, as such, usually annotated in nucleic acid/protein databases, which has led to their recognition by the scientific community. Growing evidence, however, points to the existence of "alternative" ORFs. The situation is more delicate with alternative ORFs as there is no common agreement on their existence, relevance or function. Very roughly, we can say that alternative ORFs are delimited by start/stop sites that are either not included in databases (for example, because they are considered sequence inaccuracies) or that are not

[251] Post-translational modifications can be an example of these other few steps playing a central role in the modification of the synthesized polypeptide.

[252] "The most common [non-canonical] translational mechanism is the use of an alternative start codon" (Pavesi *et al* 2018, p. 9; cf. next section).

deemed to be relevant. Their annotations are recent and/or not universally recognized (Brunet et al 2018; Pavesi et al 2018).

We can identify two main instances in which biologists usually directly invoke or imply stochasticity with respect to alternative start-codon selection events. Most of the time, when referring to the acknowledged *randomness* of the biochemical mutation process involved in the evolutionary origin of an alternative ORF from a standard one, stochasticity is implied. In fact, it is widely accepted that novel ORFs may originate within a given sequence by the process of genetic mutation. If in the case of standard ORFs their stochastic character mainly resides in the randomness of mutation with respect to evolution, for alternative ORFs arising from a random mutation, the situation is more complex. Other sorts of claims can be found in the literature on their stochastic nature.

The second instance of stochasticity in alternative start-codon selection events becomes evident by looking at how they are usually modelled. Stochastic events are described as concerning a single RNA within a single cell. This instantiation is crucial for my argument since it enables me to investigate chance within a single cell, concerning a single process. This scenario will allow me to dissect its further explanatory power and for this reason, I dedicate the entire following section to it.

2.2 Modeling and explaining alternative start-codon selection

Biologists refer to an alternative start-codon selection event when a ribosome starts the translation process from an ORF different to that which is standard (hence, alternative). The *activation* of alternative ORFs, which is actually its *translation*, can be attributed to stochasticity (Boersma *et al* 2019; Brunet *et al* 2018). More specifically, ribosome switching from the translation of a standard ORF to an alternative one could happen stochastically (see Figure 1, i.e. the ribosome switch between $start_x$ and $start_y$; see after). The effects of these dynamics are stochastic in the QLS sense, because different types of protein result (Figure 1, proteins X or Y).

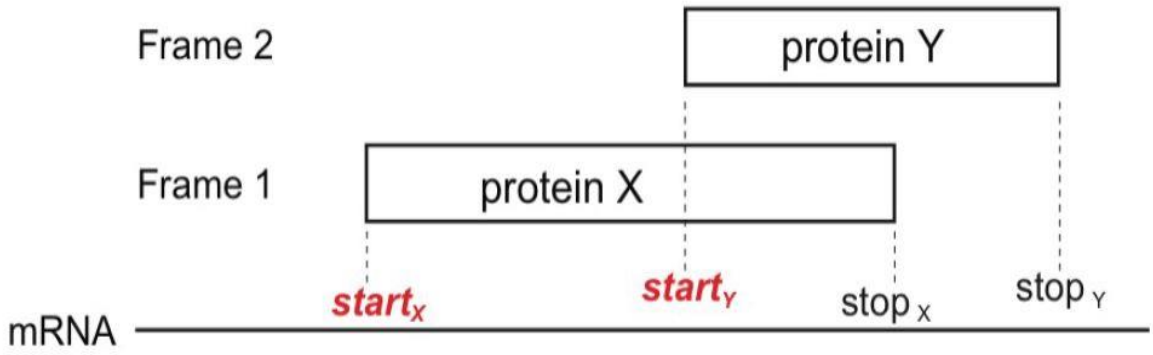


Figure 1 A model of alternative start-codon selection in the case of so-called "overlapping genes" for which the coding frame differs (from Pavesi *et al* 2018, p. 10).

Figure 1 depicts the translation of two different ORFs from the same mRNA. Let's say that $start_x$ is the start codon of the standard ORF and that $start_y$ is the start codon of the alternative ORF. Depending on which start codon the ribosomal subunits bond with, two completely different proteins will be produced (i.e. protein X or Y), as a result of the two different cases' reading frame. The switch between the start codon in frame 1 and the one in frame 2 could happen stochastically. Consequently, the production of protein X or protein Y will also be stochastic. This model is meant as an example of what I have in mind as QLS: starting from the same mRNA sequence, stochastic phenomena (selection of one of the two alternative start codons, selection of frame 1 or 2) influence the diversity (type) of genetic product (protein X or Y).

But what exactly does it mean that this switch happens stochastically? It means that each single ribosome can start translation from a standard or an alternative ORF in a non-regular fashion. This answer merely considers a single event of alternative start codon selection. However, since translation is a process taking place over time, in order to understand more deeply the way in which stochasticity operates in this molecular process, we have to consider an entire sequence of events. This is why, in the next section, I will introduce into the picture the parameter of time, thus focusing on several successive events of different (standard or alternative) start-codon selection. Metaphorically speaking, we do not ground our reflection on a "picture" but rather a "movie"[253] of (alternative) start codon selection events. This enables us

[253] Moore (2012) underlines that efforts by structural biochemists to develop "movie[s] of proteins synthesis" is misplaced since "all the functionally significant movements of the ribosome, both internal and external, are biased random walks, and it is most unlikely that any given ribosome will ever do exactly the same thing twice as it elongates some polypeptide" (Moore 2012, p. 8). Being aware of this critique, our use of the notion of movie does

to assess the possible biological implications of temporal sequences of stochastic events in a single cell cycle.

3. Stochasticity in time

How stochasticity can have a role in alternative start codon selection can be differently evaluated depending on the parameters considered (e.g. intracellular space, time at the level of the cell cycle, energy in terms of ATP used, etc.) and the ways in which they are represented in the model. For the reasons just mentioned above, we will focus on the parameter of time,[254] in particular on two simple models of it: time as order and time as duration[255]. The purpose of this section is to show how considering these two linear models of time and applying them to sequences of (alternative) start-codon events can be relevant for appreciating possible ways in which stochasticity can be explanatory in translation. More precisely, I talk of "time as order of events in a sequence" to refer to the order of (alternative) start-codon selection events over time, and of "time as duration between events in a sequence" to refer to the time interval between successive (alternative) start-codon selection events. These two temporal features of the phenomenon of alternative protein synthesis can both be affected by the stochastic character of biomolecular events. I remind the reader that the fact that these events reflect an "inherent stochasticity" (Boersma 2019, p. 471) is not just a hypothesis but has been demonstrated by experimental work (see also Lyon 2019; Blake and Bin Wu 2019). The authors therefore describe an intrinsically stochastic phenomenon which, we will see in the last part of the chapter, is highly likely to be explained through a stochastic explanation (SE).

3.1 Time as order of events in a sequence

Time modeled as order of events in a sequence refers to the way events follow one another in a linear process (e.g., A → B → C → D → E). Applied to our case study, this time model describes the linear sequence of different ORF activation events (as a matter of fact, start-codon selection events) by ribosomes, namely the temporal order in which ORFs are selected to initiate translation and synthetize proteins (e.g., ORF1 → ORF2 → ORF2 → ORF1). The temporal dimension of translation reveals an important way in which stochasticity can

not have to be intended as a revendication of any sort of regularity and/or repeatability, but rather as an attempt to take into account the temporal dimension of stochastic molecular processes.

[254] It must be kept in mind that my analysis of time exemplifies what can be done when considering other parameters involved in translation (notably, space, energy, etc.), which can help dissect stochasticity in the explanation of this molecular process and, more generally, gene expression.

[255] In my choice of these two models concerning time I have been inspired by Nicoglou (2017).

intervene in this molecular/intracellular process: it can be the source of the *alternative* orders in sequences of start codon selection events, thus meaning the order could randomly be one of the following:

[1] ORF1 → ORF2 → ORF1 → ORF1 → ORF2 → …

or

[2] ORF2 → ORF2 → ORF1 → ORF2 → ORF2 → …

or

[3] ORF2 → ORF1 → ORF2 → ORF2 → ORF1 → …

(and so on)

In other words, having sequence 1 or 2 or 3 (or any other), each characterized by its specific order of events, is a matter of chance. Why is this important? Looking at the biological literature invoking chance, it appears that the notion of stochasticity, and more generally the notion of chance, are both often used without fully spelling out which elements of the models or explanations are considered as behaving in a chancy manner. Stochasticity is often intended as a vague attitude of the system: "the system behaves stochastically". By contrast, by spelling out which elements, in which different situations, the concept of stochasticity refers to, I want to be more specific in this chapter. In addition, in the introduction to the dissertation, I provided a straight definition of chance according to which a process/outcome is chancy when it "could be otherwise". My objective here is to identify, in the context of protein synthesis – and more specifically in translation – which are the processes/outcomes that "could be otherwise". Considering this first linear model of time as order of events in a sequence enables us to go straight to the point: the order of start codon/ORF selection events over time is what "could be otherwise".

Nonetheless, one problem characterizes this first model. It is linked to the fact that time, when modeled just as the order of events in a sequence, is likely to be reduced to *causal* order. Referring to Gottlieb (1993), Nicoglou claims that time modeled as order could become "a framework within which the 'causes' operate." (Nicoglou 2017, p. 379). Even more explicitly, within the linear order of events "causation could operate and make […] time analysis [disappear] in favor of space analysis" (Nicoglou 2017, pp. 385-386). Following Nicoglou, I maintain that we need further models of time in order to avoid this kind of flattening because my aim is the exact inverse: I want to investigate how chance can affect translation by dissecting the possible ways in which it can intervene in the timing of this molecular process. I thus

propose a second model of time that could be seen as an elaboration of the first, in as far as it allows another way in which stochasticity can intervene in sequences of alternative start-codon selection events to be made explicit.

3.2 Time as duration between successive events

I now turn to a model of time that adds another element to that of time as order of events in a sequence. This new model accounts for the time that elapses between contiguous successive events. I call it time as "duration between successive events". More explicitly, this second model takes into account both the linear order of events and the relative relations between events along a fixed ΔT. In this section, I consider that stochasticity in translation can also modify time intervals between successive start-codon selection events in a sequence.

Let us start by focusing on Figure 2 below.

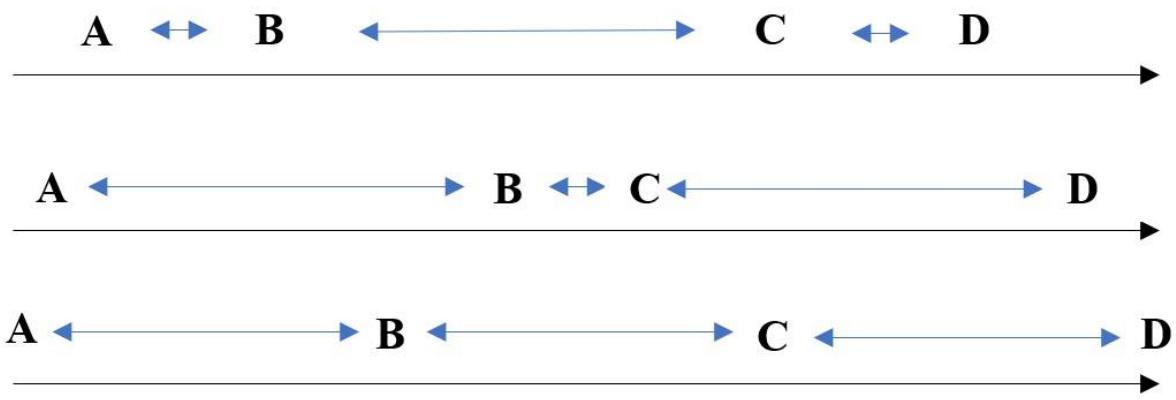


Figure 2 Three different sequences with different time intervals between events. I call this kind of model "time as duration".

In this figure, the black arrows represent fixed ΔT (specific durations/lapses of time). Note that the letters are not equally spaced in the three sequences. In the first at the top, there is a big space between B and C, and small ones between A and B, and C and D. By contrast, in the second, middle sequence, there is a small space between B and C and big ones A and B, and C and D. Finally, in the third sequence at the bottom, the letters are equally spaced. The spaces between letters correspond to the relative lapses of time between events. Thus, this model describes the linear order of the events being considered as well as their relative temporal relations.

When applied to our case study, this second model of time as duration describes the temporal intervals between different ORF selection events (e.g., ORF1 → → ORF2 → ORF1 → → → → ORF1 → ORF2). The variable number of arrows represents temporal lapses of

different durations between two consecutive start-codon selection events (i.e. the initiation of translation by a ribosome at the level of different ORFs), and each sequence represents some possible alternative sequences for this sort of event. This model describes a second way in which stochasticity can intervene by affecting timing in translation. It can also be the source of *alternative* spacing between different ORF selection events. Indeed, in our hypothetical situation we could have the following alternative scenarios:

[1] ORF1 → → ORF2 → ORF1 → → → → ORF1 → ORF2 → …

or

[2] ORF1 → ORF2 → → ORF1 → ORF1 → → → ORF2 → …

or

[3] ORF1 → ORF2 → → ORF1 → ORF1 → → → → → ORF2 → …

(and so on)

As before, having sequence 1 or 2 or 3 (or any other), with their specific durations between ORF selection events, is a matter of chance. Again, this second linear model allows us to identify the elements that "could be otherwise" (Wong 2020) to which stochasticity refers. Indeed, the relative temporal intervals between possible ORF selection events are the elements that "could be otherwise".[256]

Up to this point, I have proposed a conceptual reflection on the possible ways of framing stochasticity in translation, focusing on the parameter of time. More specifically, we showed that, by considering two linear models of time and applying them to our case study, stochasticity is shown to be a potential source of *alternative* orders of start-codon selection events over time and/or of *alternative* temporal spacings between them. This is indeed a precious theoretical gain. Why? Because it allows us to point precisely to what is chancy in the process of translation – to what "could be otherwise". Moreover, the result of our analysis up to now is a first step towards reconsidering the role that stochasticity (chance) can play at the molecular level, in particular in the process of gene expression.

[256] The two models can obviously be integrated. The various sequences may vary in both order and temporal spacing between the different ORFs.

3.3. Two specifications

This short section addresses two specifications related to the fact that, when applying the models of time as order and time as duration to our case study, I implicitly assumed that each sequence of start-codon selection events is produced from a single RNA.

First, it might be though that, according to my analysis, only a single ribosome scans an RNA sequence at any one time, thus producing a succession of start-codon selection events. However, more than one ribosome can be attached to a given RNA sequence at any one time and, as a consequence, more than one sequence of ORF/start-codon selection events can be simultaneously produced. In such a scenario, the various sequences of stochastic events would have to be combined and analyzed together in order to have any idea of their overall impact on the end result on translation.

Second, I do not ignore that there can be multiple copies of the same type of RNA in a cell, which can then be translated simultaneously. This makes the scenario more complex insofar as it enhances the possibility of overlaps between sequences of start-codon selection events. Moreover, each sequence can be the result of different dynamics. Multiple activations[257] of two ORFs (e.g., ORF1and ORF2) in a single RNA or the sum of solitary successive activations of different ORFs in multiple RNAs of the same type can both be at the origin of a given sequence of activation events in translation. For instance, see [actual 1] sequence below.

[actual 1] ORF1 → → ORF2 → ORF1 → → → → ORF1 → ORF2 → …

More precisely, the sequence [actual 1] can be the result of two possible scenarios (see Figure 3). The first shows a single RNAs from which multiple events of translation activation (represented by the arrows) take place. The second scenario shows multiple RNAs of the same type, each of which activates translation (again represented by the arrows) of a single ORF.

[257] By "multiple activations" we refer to the possibility whereby the ORFs of a single RNA are activated more than once. In this case, ribosomes assemble and run on the ORF_1 and ORF_2 more than once.

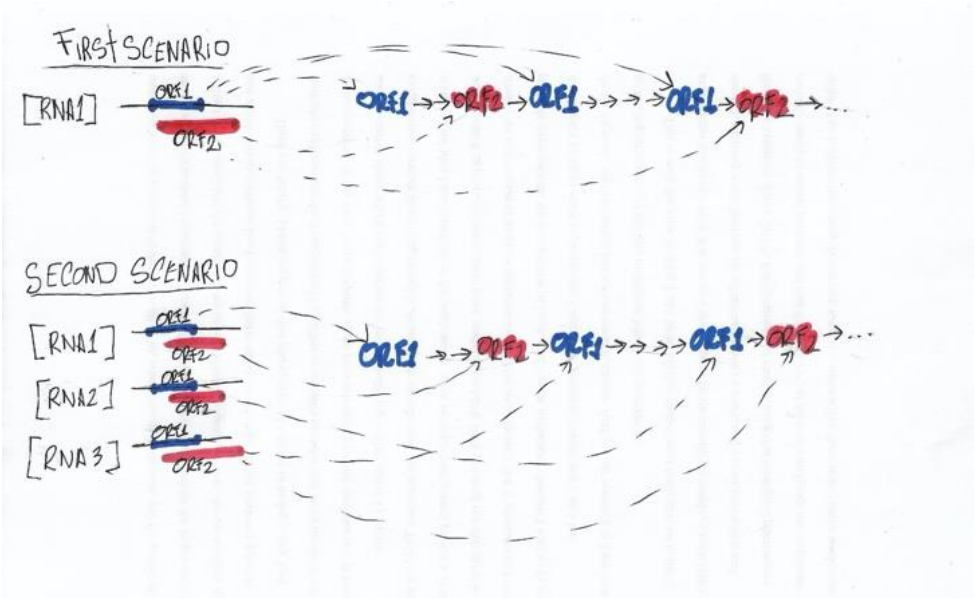


Figure 3 Two hypothetical scenarios in which [actual 1] could be realized. At the top: multiple activations of two ORFs (i.e. ORF1 and ORF2) in a single RNA. At the bottom: a single activation of ORFs in multiple RNAs of the same type.

What matters here is to recognize that the actual sequence of different ORF activation events – which constitutes [actual 1] and is characterized by a specific linear order and specific time intervals – is the result of successive stochastic events of start-codon selection by ribosomes, which can be realized by either of these dynamics (the first involving just one RNA, the second more than one (in this example, three)).

In general, the fewer the number of proteins produced, more important is stochasticity's effect. This is because when there are fewer proteins, their presence, absence, order, and duration carry greater importance for cellular life. In statistics this is called the low number effect: when fewer elements are involved, the difference between outcomes that are possible because of variations of those elements grows in magnitude. Hacking (2001) explains this statistical property with a simple example. Males and females are born in roughly equal numbers. If we have two hospitals, one in which many babies are born and one in which few babies are born, the question Hacking asks (or rather that he poses to the reader) is, in which hospital will we find unusual weeks in which babies, female or male, exceed 55% of births? The answer is the small one, because large deviations are more probable, and then common, in small populations (p. 250). In our case study, the greater the number of overlapping sequences of start-codon selection events, the smaller the possible impact of stochasticity on the end result of translation. Why? Because the potential impact, in terms of proteins produced, of each sequence stochastically produced can be compensated by the other sequences (which are also stochastically produced). This is to say that the low number effect no longer applies.

4. The biological relevance of my philosophical distinction

The question that I want to now ask is: why should we see the distinction between these two ways in which stochasticity intervenes in the timing of alternative start codon selection events as biologically relevant? In order to answer this question, we need to take another step into the scenario's complexity. To ask more specifically whether this distinction is relevant for the study of cell functioning is to ask whether the proteins produced by the activations of ORFs – with their order of appearance as well as the time interval between them – can have an impact on cell metabolism. Note that, when I discussed the low number effect at the end of the previous section, I referred to the number, the sequence, and the order *of proteins produced,* and not to the activation of ORFs. Thus, before answering the question above concerning the biological relevance of my analysis, the purpose of this section is to give an explicit argument for this change of focus.

4.1 Order of translation starts vs. order of full protein completion

In the models I have so far proposed, I focused on the activation of ORF events, namely on events in which ORFs are bonded by ribosomes and the synthesis of polypeptide starts. We talked about two ORFs – ORF1 and ORF2. Now, we have to add a further level of complexity because the order of ORF *activation* does not necessarily reflect the order of the *appearance* of proteins. By the appearance of proteins, we mean the moment in which the ribosome finishes synthesizing a polypeptide, and in which this latter is "ready" to achieve its role. Let's imagine we have the activation of ORF_x and ORF_y and that ORF_x encodes for a very long proteins and that ORF_y, by contrast, encodes for a very short polypeptide. If the time it takes to synthesize protein X is bigger than sum of the time it takes for ORF_y to activate and protein Y to synthesize, then the order of ORF activation does not reflect the order in which the proteins appear.[258]

[258] In this scenario we assume a constant rate of protein synthesis.

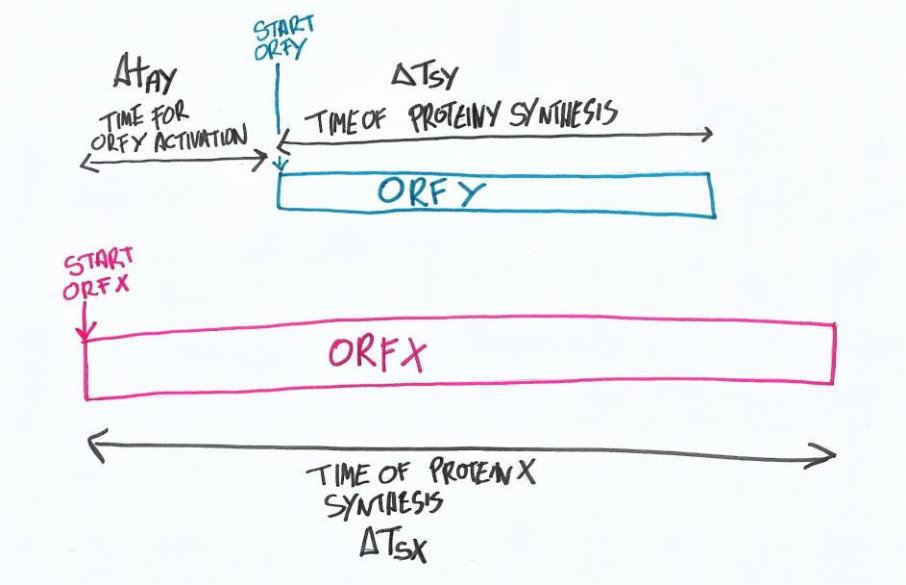


Figure 4 A representation of different time periods during which the activation and synthesis of the proteins Y and X could take place.

In Figure 4 we can see that if $\Delta T_{AY} + \Delta T_{SY} < \Delta T_{SX}$ then the order of activation does not reflect the order in which the proteins appear:

Order of activation

$ORF_X \rightarrow ORF_Y$

Order of appearance of proteins

Proteins Y → Proteins X

Could this possible asymmetry between these two kinds of orders (i.e. activation and appearance) be a problem for the analysis of stochasticity that has, up to now, been proposed? I think the answer is no and that, on the contrary, this could be seen as a further insight into the complexity of stochasticity in translation. Besides the two different ways in which stochasticity can intervene in the parameter of time, we can also say that we have an additional source of stochasticity in the order of appearance of proteins that cannot necessarily be reduced to the order of activation events. In the case shown in Figure 4, we can reasonably assume that the order of protein appearance is predictable from the order of ORF activation since we know the time of activation and synthesis of both of the two proteins. But in biological reality these times are not always fixed and predictable. Such stochasticity is therefore not reducible to our momentary ignorance with respect to the value of the parameters discussed above. It is, rather,

an original peculiarity with which living systems enrich their strategies of variability in protein production.

4.2 Some plausible biological scenarios

Let us return to the question regarding the biological relevance of the distinction between the two ways in which stochasticity (can) intervene in the timing of translation. I suggest considering invented but realistic, which is to say plausible, biological scenarios of the temporal dynamics of protein production in order to show more precisely how stochasticity can have a decisive impact on these dynamics, and thus on their outcome.

The fist plausible scenario that I propose is as follows: consider two different possible orders of appearance of two antagonist proteins, protein 1 and protein 2, one of which inhibits the activity of the other. Depending on the order that takes place, these proteins will have a different impact on cellular metabolism (Figure 5):

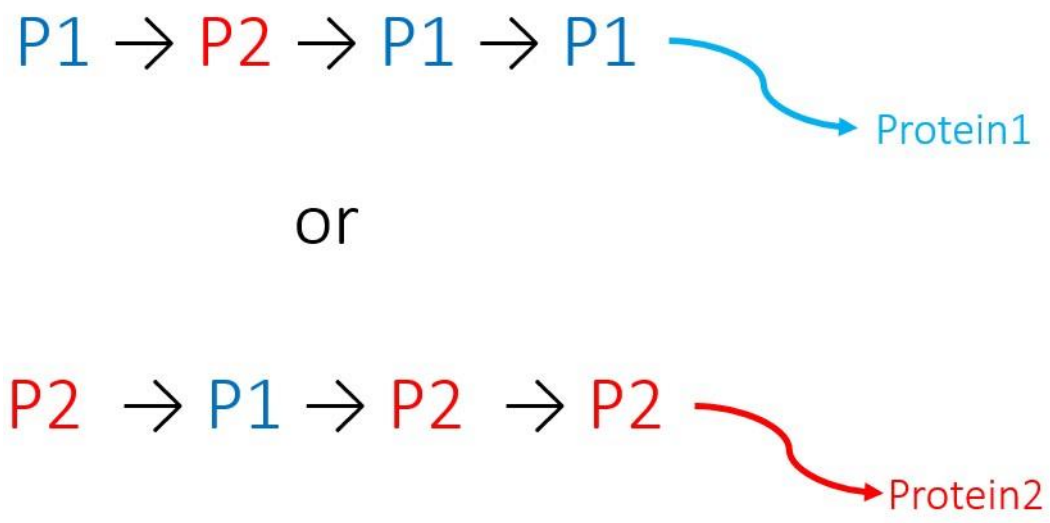


Figure 5 Representation of a possible biological scenario in which a change in the order of the appearance of proteins could affect cell metabolism.

In this case, the specific order of the appearance of proteins could be crucial in determining cell metabolism and behavior. In the upper part of Figure 5, we can see that the first specific order allows mainly protein 1 to perform its role. In the lower part of Figure 5, by contrast, we can see that the second order allows mainly protein 2 to perform its role. Stochasticity in this case intervenes by determining which of the two orders actually takes place, and consequently the protein's appearance. Stochasticity affecting timing conceived as the order of protein

appearance may therefore be relevant to the extent that it can affect the metabolic activity of the cell.

The second plausible scenario that I propose is one in which there are differences in terms of time intervals between the appearance of two antagonist proteins. In figure 6, we can see that the lapse of time between the appearance of protein 1 and the appearance of protein 2 is crucial in determining the effects that they might possibly have on cell metabolism and behaviour. In the upper part of Figure 6, protein 1 appears first. Since the time interval between its appearance and the appearance of protein 2 is long enough to allow protein 1 to act, we could imagine that protein 1 could impact metabolism/behaviour prior to the synthesis and appearance of antagonist protein 2. Seen the other way around, in the lower part of Figure 6, the time interval between the appearance of the two proteins is quite short. Then, we might deduce that protein 1 does not have enough time to achieve its activity before the synthesis and appearance of antagonist protein 2.

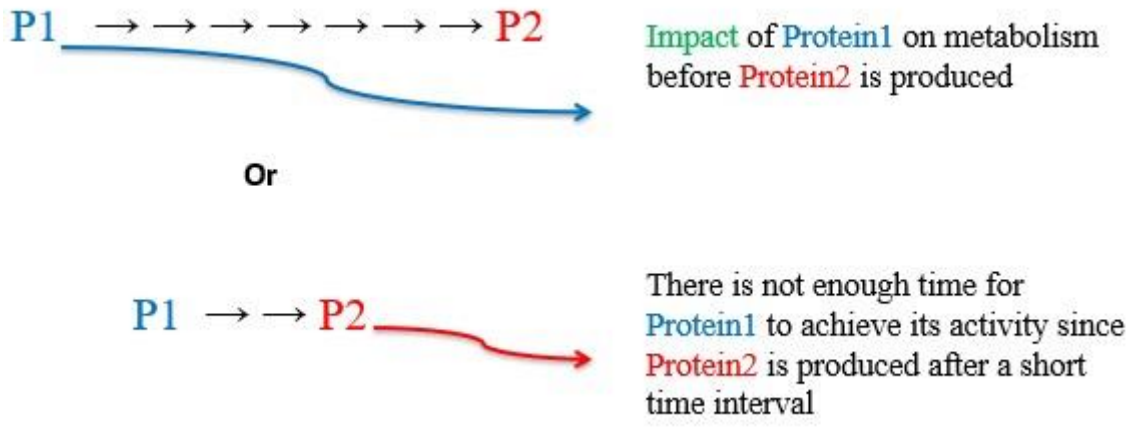


Figure 6 Representation of a plausible biological scenario in which the change in duration between the appearance of proteins could affect cell metabolism.

Stochasticity in this case intervenes by determining the change in duration between the appearance of proteins 1 and 2. This difference in duration can consequently determine which of the two antagonistic proteins will act first on cellular metabolism. This second plausible scenario also shows that stochasticity can have important effects on cellular metabolism.

To sum up, these two qualitative examples (i.e. Figures 5 and 6) suggest that our reflection on the different ways in which stochasticity can intervene in the timing of translation could be relevant to the study of cell behavior. These scenarios are *hypothetical* and do not stem from empirical evidence, though they are nonetheless plausible. In addition, we saw that in order to provide them, I had to shift attention from the order of *activation* of ORFs to the order

of protein *appearance*. Far from being a problem, I think instead that it adds precision to my argument. This relation does not always receive the correct attention by scientists and I think, by contrast, that it is crucial for understanding the complexity behind stochasticity in translation.

5. Philosophical implications of alternative-start codon selection analysis

In this last section, I propose comparing my stochastic explanation (SE) instantiated through the alternative start-codon selection with Lewis's well-known metaphysical conception of chance. By proposing a very different account, I am able to better underline the virtues of my work. Specifically, dealing with a precise account of chance understood in a strongly metaphysical register allows me to emphasize how shifting the focus away from investigating the ontological, uncaused nature of chance and towards the epistemological functionality that it exercises in the explanation is of epistemological benefit. I do not wish to exclude the possibility that the ontological study of chance as uncaused cannot be useful for understanding biological processes more deeply (see studies on quantum biology, for example). However, this comparison clearly highlights why we should consider chance firstly at an epistemological level (cf. introduction of the dissertation, section 1.4).

Lewis writes that we cannot explain why one thing happens rather than another if that thing happens by chance (p. 230). Indeed, he asks rhetorically, "is it not the very essence of chance that one thing happens rather than another for no reason whatsoever" (*ibid.*). He adds that if we had two identical causal chains, one actualized and the other unactualized, and we had two different outcomes from each chain, let's say X and Y, then we would not be able to explain why the actual causal history produces result X rather than another identical causal history producing Y. It seems therefore that Lewis conceives of chance as something metaphysical (although it is not very clear what it is that he refers to) which is set apart from (and/or beyond) causal histories. He does not deny that it is possible to explain chance phenomena (e.g. chance outcomes like X and Y), but denies that it can be done in a contrastive way. According to him, if chance is involved, I can explain why I have a certain effect with respect to an actualized causal chain, but I cannot explain why I have that effect *rather than another* (i.e. I cannot provide a contrastive explanation). In other words, if we are dealing with chance, the only type of explanation possible is the one that answers to a plain "why" question (p. 231). I propose overturning Lewis's position and instead emphasizing the relevance of the analysis of stochasticity in translation in alternative start codon selection events proposed in

this final chapter, and, more generally, of my account of stochastic explanation (SE). Let's first of all look at the differences between my account and Lewis's.

Lewis's account of chance is metaphysical and focused on outcomes. As noted above, he writes that "we cannot explain why a chance process yields one outcome rather than another" (Lewis 1983, p. 230). He thus talks about chancy outcomes and not chancy processes. Specifically, he invites us to think of two identical causal histories that yield two different outcomes (X) and (Y), one actualized (1) and one non-actualized (2). He states that we cannot explain why a given result (e.g. X) happens rather than another (e.g. Y) when one of the two identical causal chains is actualized (e.g., causal history 1). Therefore, he conceives of chance in metaphysical terms as pure indetermination. In other words, he conceives of chance as uncaused. By contrast, my account of chance is epistemological and focused on processes. I start from an *explanandum/explanans* scheme in which chance (i.e. CPC) characterizing the *explanans* allows the different ways in which the *explanandum* can be produced to be held and synthesized. In this context, it is thanks to, and not in spite of, chance that we can provide a stochastic explanations (SE).

With the differences between Lewis's metaphysical account and my own epistemological, process-specific one in mind, let's now show in clearer terms the value of my account of SE. In order to do that, the first step is to show that stochastic explanation (SE) can apply to the models proposed for alternative start codon selection.

5.1 Models of time for the alternative-start codon selection provide SE

Consider a model of alternative start-codon selection events that represents both order and duration. Stochasticity in this case can be the source of both the *alternative* orders and the alternative durations in sequences of start codon selection events, which could randomly be one of the following:

(e.g.)

(1) ORF1 → → ORF2 → ORF1 → → → → ORF1 → ORF2

or

(2) ORF2 → → → → → → ORF1 → ORF2 → → → ORF2

or

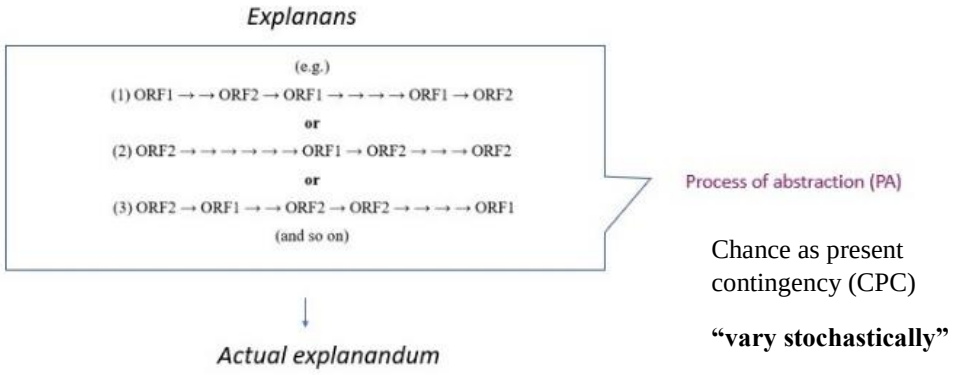
(3) ORF2 → ORF1 → → ORF2 → ORF2 → → → → ORF1

(and so on)

We can see that the order of ORFs changes from sequence to sequence as well as the time lapse between successive ORFs selection events. This model thus nicely synthesizes our two models of time as order and time as duration. Let us now try to express the kind of explanation this model might provide.

Explanandum: How can we explain variation in the kind of proteins produced over time from the same single RNA transcript in a situation of alternative start-codon selection events?
Explanans: We can explain this variation using the models above. The events of ORF activation (and of protein production) can ***vary stochastically*** *with respect to both temporal order and duration (the model proposed above qualitatively exemplifies three ways in which the explanandum can be realized).*

In this case, we can identify the notion of chance as present contingency (CPC) as the fact that, given a single RNA, the order and duration of the ORF activation sequences can be otherwise in a disjunctive way.[259] This is expressed in the *explanans* with the phrase "vary stochastically" and we can visualize these "disjunctive ways" in the sequences (1), (2), and (3). The following schema represents the general schema of stochastic explanation (SE) applied to the case of alternative start-codon selection events:



(Variation in types of proteins produced for the same single RNA transcript in time)

Figure 7 A schema of stochastic explanation (SE) showing disjunctive ways in which the alternative start-codon selection can take place.

[259] We must notice that in a scenario in which we have more RNAs of the same type, we could have also a situation of cumulative (and not only disjunctive) representations of different alternative start codon-selection event sequences.

On the left of Figure 7, we can see that in the *explanans* we have the disjunctive ways in which alternative start-codon selection events can take place over time. In the *explanans* these possibilities are held and synthetized via the notion of chance as present contingency (CPC) expressed as "vary stochastically" (Figure 7, on the right). In other words, if in my general account of SE, CPC allows for the different ways in which the *explanandum* could be realized to be held and synthesized, in the case of alternative start-codon selection events, these "different ways" are represented by the disjunctive set of possible ORF activation sequences which differ with respect to timing as order and as duration.

Now that I have succinctly applied my idea of SE to the alternative start-codon selection events models, let's see how this account can be used to emphasize still better the value of my SE account by comparing it with Lewis's ideas of chance and explanation.

5.2 Overturning Lewis's account; showing the contribution of my account of SE

At the beginning of section 5, I suggested that my account is upside down compared to that of Lewis. Let us now detail in what sense this is so. The overturning that I propose is as follows:

Lewis: "[we can never] explain why a chance process yields one outcome rather than another" (p. 230).

Me: "1 – In the context of an SE *explanans*, it is not our interest to know why one process takes place rather than another because it is exactly the fact that a process could be otherwise – possibility that is held and synthetized in the notion of chance as present contingency (CPC) – that enables a given *explanandum* to be explained;

2 – In addition, the fact that an SE *explanans* contains the notion of CPC does not prevent us from explaining why the *explanandum* is X rather than Y".

Let's try to unpack the two points above with the following Figure 8:

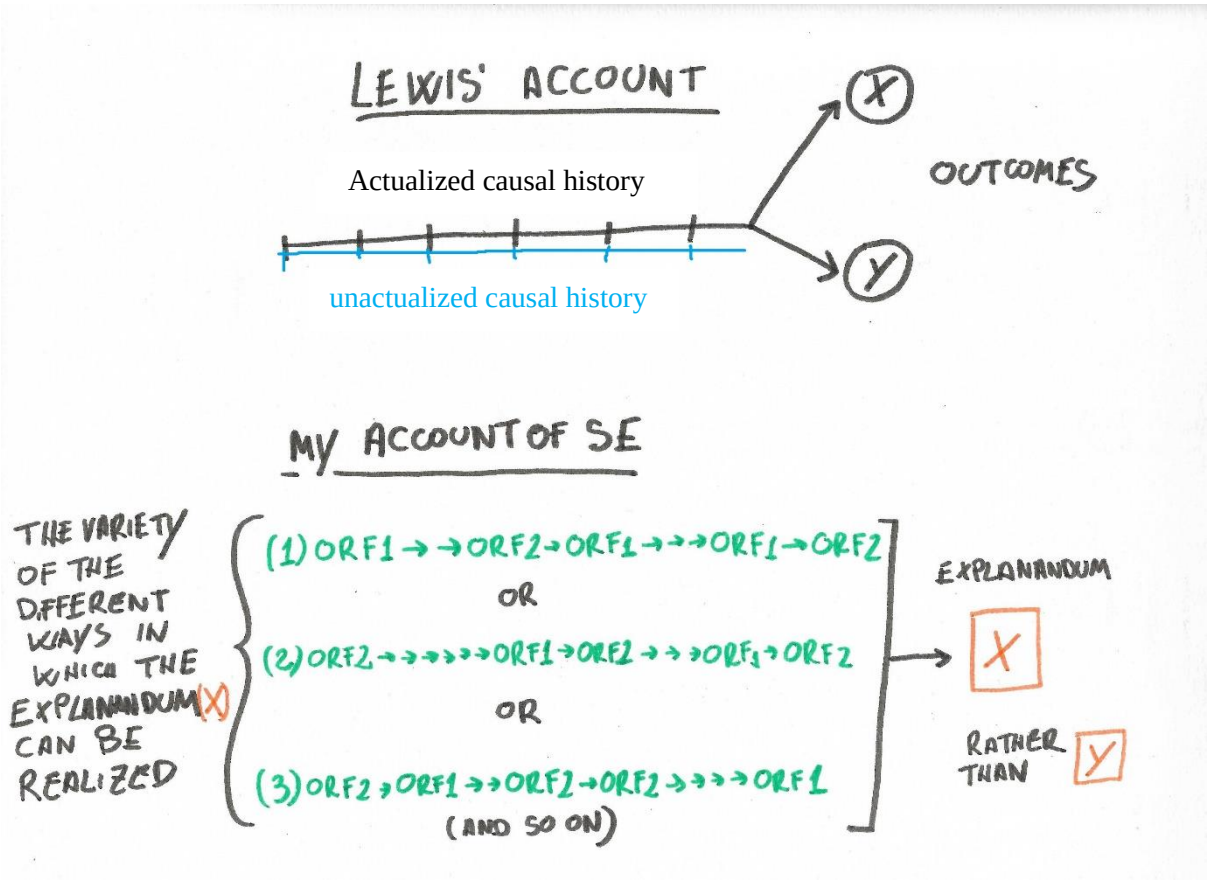


Figure 8 A comparation of the account by Lewis with my idea of stochastic explanation in the case of alternative start-codon selection events.

Unpacking point 1

Following Lewis, chance consists of not being able to explain why one given outcome takes place rather than another. For Lewis, this is a limitation since it does not allow contrastive explanations. Indeed, for the author chance cannot have an explanatory power in sense of forming contrastive explanations. In Figure 8 above, we can see two superposed identical causal histories (in which only one is actualized) and the two outcomes, X and Y. Chance as pure indeterminacy is present at the level of these outcomes, which is why, according to Lewis, we cannot explain why we have X rather than Y. By contrast, in my account, not being able to explain why I have one process rather than another is not a limitation because my picture is broader. Indeed, I do not focus on two (or more) outcomes of a single (or two identical) causal chain(s), but on the description of the possible processes involved in the explanation of an *explanandum*. Referring to the bottom of Figure 8, it is precisely the fact that the process could be otherwise (as in the case of the alternative start-codon selection events sequences; Figure 8 in green) in a cumulative or/and disjunctive way that means the *explanans* is able to explain the *explanandum* (Figure 8, in orange).

Unpacking point 2

Since he adheres to a metaphysical conception of chance as uncaused, Lewis affirms that we cannot say why we have outcome X instead of outcome Y (i.e. pure chance, cf. Millstein 2011). I affirm otherwise: chance does not prevent me from explaining why I have *explanandum* X instead of *explanandum* Y. In fact, an *explanans* with CPC can explain why I have *explanandum* X versus an *explanandum* Y. Indeed, the many possible ways in which a sequence of ORF activations, in terms of temporal order and duration, can occur (i.e. the *explanans* containing CPC) allows me to explain why we observe variation in the types of proteins produced over time from a single RNA transcript (i.e., the *explanandum* X), rather than just the same type of protein (Y).

This overturning of Lewis's approach allows me to go even deeper than my proposal of stochastic explanation (SE) and further emphasize its value. The fact of having a description in which processes can be otherwise is not a limitation but rather the very explanatory power of SE.

Conclusion

For this final chapter of the dissertation I set the following goals: 1) To propose shifting attention from transcription (usually the process in which the study of chance is considered most important) to translation (a process that is often seen as a copy-paste of transcription, and thereby receives less attention in the analysis of its stochastic events); 2) To propose a new attitude towards studying chance, focusing on the various parameters in which it can intervene. I propose focusing on time for the reasons laid out in the introduction of the present chapter. Nonetheless, several other parameters can also be approached with the same attitude, for example, intracellular space, ATP metabolism, etc.; 3) By confronting my own account of SE with Lewis's I tried to make its novelties even more evident by comparing it to a strongly metaphysical account of chance. Chance is not to be conceived as a metaphysical property that prevents the explanation of why one event happened instead of another. It is, instead, precisely chance that allows us to go deeper with respect to the possible ways in which a process can occur, and with respect to the explanation, and to the possible ways in which the *explanandum* can be realized.

Conclusion of the dissertation

My contribution to the philosophical debate on chance in biology consists of widening the field of reflection towards areas of biological research beyond evolutionary biology. Taking the analysis of explanations found in molecular cellular biology as a starting point, my goal is to give an epistemological dignity to chance in this context. The intention of my account of "stochastic explanation" (SE) developed in this dissertation is for it to function as a philosophical tool which allows the central role of chance as an abstractor in cellular and molecular explanation to be recognized and accounted for. More specifically, the purpose of this account is:

1) To make explicit the epistemological role of chance in terms of abstraction, as implied in existing biological explanations. This is the meta-normative part of my work which consists of evaluative claims (i.e. I claim that some explanations are good because they use chance as an explicative element).
2) To propose that in some contexts, such as the study of noise in gene expression, it is a matter of reevaluating how the chance element is conceived, and its role. I have argued that this is fundamental to explaining certain phenomena in depth. This is the most strongly normative part of my work and consists of meta-normative prescriptive claims (i.e. I make claims about the approach biologists should adopt in developing certain explanations).

1. Mapping stochastic explanations (SE) in the cell

In the present dissertation, I have drawn attention to many examples of explanations of different cellular and molecular processes. What I would like to propose in this section is a visual map that can usefully gather them together into a comprehensive picture. More explicitly, in providing this map I have two aims: 1) I want to highlight how the role of chance is fundamental for explaining many cellular and molecular processes; 2) I want to underline that there remain numerous other explanations in molecular cellular biology in which the notion of stochasticity need to be unpacked and clarified.

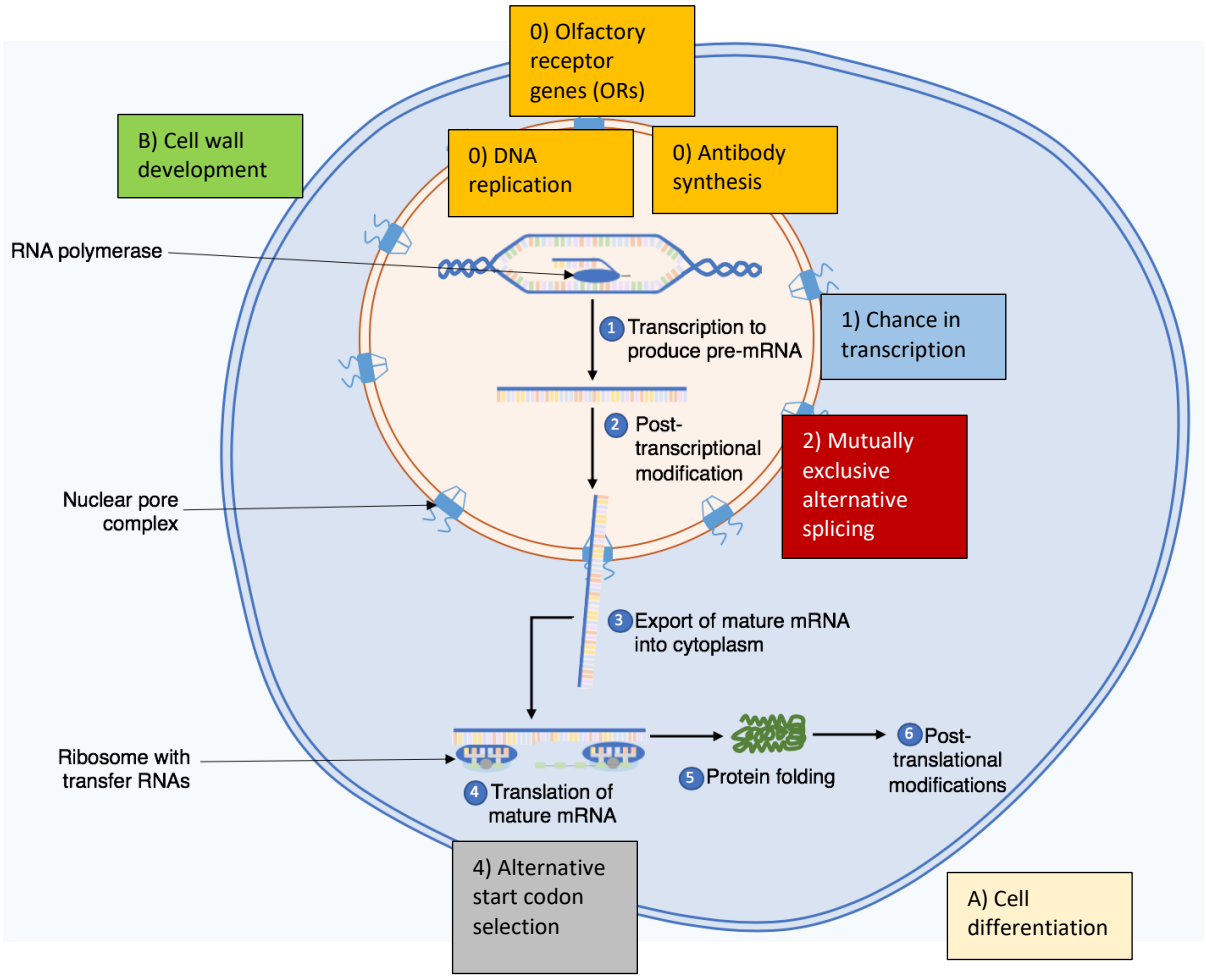


Figure 1 Representation of a cell annotated with the processes explained by stochastic explanations (SE)

Figure 1 represents an ideal cell, insofar as no real cell undertakes all the processes mentioned. Only certain cells do certain processes, others not. For example, while DNA duplication takes place in all cells, the expression of olfactory genes does not, as it is a process that only concerns the olfactory sensory neuron (OSN). The story is the same with cell wall development in plant cells. Moreover, though my dissertation has not been exclusively focused on researches on eukaryotes, the figure represents an ideal eukaryotic cell. Nonetheless, it allows us to map most of the examples analyzed in the present dissertation in a single drawing. Let's follow the progression outlined visually in Figure 1 in the text below:

- Clustered together here are explanations that refer to DNA: the expression of olfactory receptor genes, DNA replication and antibody synthesis. Indeed, in Chapter 3 and Chapter 4, I showed that stochasticity is considered to play a very important role at the heart of the cell (Kupiec 2009a), and in explaining the processes in which DNA is involved.

- Moving forward and following the steps of gene expression, we encounter transcription. Although the notion of chance in transcription has often been developed in terms of noise, what I have tried to show is that this prevents an appreciation of the role of chance in explaining this step of gene expression. In Chapter 5, I gave three arguments on how it is often better to use a stochastic explanation (SE) rather than a non-stochastic, more highly detailed one.
- Moving on, we encounter post-transcriptional modifications. There are various processes associated with these types of modifications. I focused on the explanation of mutually exclusive alternative splicing since its relationship to the notion of stochasticity is particularly controversial and, therefore, philosophically interesting. In Chapter 6, I tried to contribute to the clarification of this relationship by proposing a definition of chance as an "insufficiency of causal dependence" (ICD). This definition sought to go beyond the characterization of chance as a merely metaphysical notion or as being due to our cognitive limitation.
- One of the last steps of gene expression is translation. It is here that RNA is translated into amino acid chains. In Chapter 7, I analyzed a specific event in this process, named the alternative start-codon selection. Modeling stochastic phenomena by making the parameter of time explicit was useful in further highlighting the crucial role of chance in the explanation of this biological event, and, more generally, in the explanation of translation. In addition, chance's essential role becomes even more evident when we consider that this is one of the last steps in gene expression, meaning the effects of stochasticity are likely to have direct quantitative and qualitative effects on the proteins produced.
- Apart from gene expression, two additional boxes can be seen in Figure 1. One refers to cell differentiation (A) and the other to the development of the plant cell wall (B). On one hand, cell differentiation is a process that involves the interaction of multiple cells.[260] We have seen how a stochastic element is at play in explaining this phenomenon, specifically in explaining the synthesis and subsequent interaction between the two membrane proteins, Notch and Delta. On the other hand, since we are considering the cell wall, a plant cell's ultimate

[260] The interaction between multiple cells is not depicted in Figure 1. I have limited myself to putting the box of "cell differentiation" outside the cell (see letter A).

> outer envelope, the example of plant cell wall development is located at the very edge of the cell. Even in this case, explaining the spherical development of the cell wall requires the use of a stochastic element, in this case specifying that the microfibrils are randomly arranged.

This map shows clearly where the phenomena, explainable through stochastic explanations (SE), take place in a cellular context. But this map is by no means exhaustive. Many other cellular and molecular processes present in biological literature are characterized as chancy, stochastic, random, etc. The interesting thing is that up to now, no philosophical reflection has been proposed for them. For example point (3), the export of mature mRNA into the cytoplasm, is worthy of attention in future work. Indeed, contemporary biological articles speak at length and expressly of stochastic transport (Bressloff and Newby 2013) and it would be interesting to investigate in what sense they talk about stochasticity. Even protein folding (5) and post-translational modifications (6) are often characterized as stochastic (Hansmann and Okamoto 1999; Gō 1983). But again, theoretical reflections on the meaning and role of chance in this context is still lacking. Going beyond Figure 1, ion channels on the cell membrane,[261] vesicular transport, and the assembly of protein systems (such as ribosome assembly just prior to translation) are other examples of phenomena often described as stochastic but for which no epistemological elucidation concerning their chancy nature has yet been developed. Could we say that all these processes can also be explained by stochastic explanation (SE)? This is a question that I leave open for future work on the topic.

2. On the applicability and scope of stochastic explanation (SE)

As should be clear from the title of my dissertation, my investigation, and specifically my account of stochastic explanation (SE), have been developed from a careful analysis of explanatory practices in cellular molecular biology. Regarding scientific explanation, in Chapter 2 I strongly rejected the mechanistic account, pointing out that explanation in terms of mechanisms is an unsatisfactory framework for accommodating the explanation of processes considered stochastic. However, the question of whether my account of stochastic explanation (SE) is compatible with other philosophical accounts of scientific explanation and/or whether this type of explanation is applicable to other biological domains is still an entirely open

[261] Actually, a philosophical literature on ion channels already exists. It focused on understanding whether molecular processes can be framed as genuinely indeterministic processes or not (e.g., see Gessell 2017; Weber 2000, 2005, 2008).

question. As to the first point, I maintain the mechanistic account of explanation – in particular the idea of stochastic mechanism – was necessary, since it provided the "lever" by which I made a case for the need to propose novel approaches to the study of stochasticity. However, this does not mean that in future work we cannot reflect further on the possible usefulness of stochastic explanation (SE) for the mechanistic account.

The basis of future investigations of this issue can already be found in the present dissertation. Indeed, some of the biological explanations that I proposed framing with my account of SE could also be understood by some philosophers as *mechanistic*. For example, if we consider Machamer *et al*'s (2000) definition of mechanism as composed by "entities and activity…" (p. 3), the explanation of the expansion of the cell wall analyzed in Chapter 3 could be seen as a mechanistic explanation where the "entities" are the microfibrils and the "activities" are the random rearrangements of these entities in space. Following this hypothesis, it seems that this explanation can be conceived as stochastic and mechanistic *at the same time*. However, no philosophical work has been proposed to substantiate this hypothesis. In this vein, one of the possible directions my future work could take would be to attempt to provide to mechanistic explanation the stochastic aspect that it has always struggled to accommodate (cf. Chapter 2). In this sense, it would also be worthwhile analyzing the famous example proposed by Machamer *et al* (2000) of neurotransmitter release, investigating the ways in which it can be conceived as stochastic as well as whether its *explanation* can be conceived as stochastic too. This case study, which sits at the heart of the mechanistic framework, could become fertile ground for elaborating on the compatibility of mechanistic and stochastic explanation (SE).

The same suggestion I have just proposed for mechanistic explanation could also apply in the case of the other philosophical accounts of explanation, derived directly from biological practice, that I have mentioned and analyzed throughout this dissertation. However, I have not investigated what kind of philosophical interpretation might be associated with them. For example, were we to show that some of the explanations could be considered as causal, statistical, mechanistic, etc., it would be interesting to investigate their relationship with my own interpretation of them as stochastic explanations. In Chapter 3 (section 1.2.2.1) I defended my own agnostic position with respect to whether a stochastic explanation (SE) is compatible with a causal explanation or not. In that context, the stochastic explanation did not draw its explanatory power from an explanation being causal (or not). Rather, its power derived from the notion of chance as present contingency (CPC) which is present in its *explanans*. But this did not imply that the explanations I was working with could not also be conceived as causal. Let's assume that this was the case for some. In this context, having worked with a causal

explanation and labelled it as SE prompts novel questions with respect to how the causal explanation is compatible with the stochastic. The same works for statistical explanations: could Lange's really statistical (RS) explanation and Ariew's statistical autonomous explanation (SAE) both be conceived as both as statistical and stochastic? Finally, could Sober's explanation be conceived as both equilibrium and stochastic?

The take home message is that future work might attempt to state that stochastic explanation (SE) could be conceived as a meta-explanation that can be applied to existing explanations in order to highlight chance role within them. This meta-stochastic explanation would provide a useful philosophical tool that would invite both philosophers and biologists to be more careful when it comes to chance, randomness and stochasticity.

The last point I would like to make with respect to the purpose of stochastic explanation (SE) is the following. Although the refrain of the entire work has been to specify how chance plays an essential role in explanation, I have not gone so far as to explicitly say that stochasticity is an indispensable element for the agent in accounting for a certain *explanandum*. Let us now unpack this idea by taking up the reflection made on Lewis's work on chance and explanation in the last part of Chapter 7.

One of the points argued in Chapter 7 was that, while we cannot say why chance processes yield one result over another, for Lewis, conversely the fact that an SE *explanans* contains the notion of chance as present contingency (CPC) ***does not prevent us*** from explaining why the *explanandum* is X rather than Y (see Chapter 7, section 5.2, point 2). This statement can be put more strongly if, instead of writing "does not prevent us", we put instead "is indispensable for". In this way the claim would become:

"The fact that a SE *explanans* contains the notion of chance as present contingency (CPC) **is indispensable for** explaining why the *explanandum* is X rather than Y"

This thesis is stronger because it implies that without the stochastic element, we could no longer explain *explanandum* X. Or, in other words, to explain X, it is necessary to rely upon stochasticity. In this case, it is implied that stochasticity is the *only* way to explain X. But is this really the case? We could use a very detailed *explanans*, in which we specify all the relevant causal factors, as an argument to say that CPC, at the end of the day, is not indispensable for explaining X. I would like to make three objections. First, looking for a more detailed explanation goes with a commitment to determinism: by specifying all the possible ways in which the *explanandum* can be realized, we seem to assume that finding all the causes enables

us to explain all the consequences. From this perspective, chance has no room. I reject this idea of determinism because it is grounded on a strong metaphysical position that is difficult to support for a bottom-up philosophy of biology project. As already specified in Chapter 1, this kind of project has to be as close as possible to scientific practice, which means avoiding any kind of *a priori* metaphysics.[262] The second resides in the fact that, even if we are actually able to develop a more detailed explanation in which we specify all the relevant causal factors (which would not anyway be clear), it would not be a good explanation because its sensitivity would increase and its cognitive salience decrease (cf. Chapter 3, section 3.4). Elaborating a very detailed *explanans* means that if any element changes in the phenomenon under examination, the explanation loses its effectiveness as the *explanans* is no longer able to account for the *explanandum*. Moreover, such a large amount of detail drastically decreases the cognitive salience of the agent because the information content becomes too complex for comprehension by a single (or community of) human being(s). Third, even if an explanation with this degree of detail were possible, it would lose fundamental elements such as systematicity (showing how the elements of the process behave as a whole) and multiple realizability (showing how the same *explanandum* can be achieved through alternative processes).

Taking into account these three points might suggest that I could go further and actually claim that chance as present contingency (CPC) is indispensable in accounting for *explanandum* X. But while I do not stand against the proposition to label as indispensable chance as present contingency (CPC), I am nonetheless convinced that further in depth philosophical work, beyond the scope of this dissertation, is required to defend this strong account.

3. Towards a naturalized metaphysics of chance in cellular and molecular biology

My work on stochastic explanation (SE) provides a great springboard for addressing ontological questions on chance. Contemporary ontological reflections *on chance* in biology mainly focus on the debate around the interpretation of probabilities. When biological theories use probability, what exactly are those probabilities? Are probabilities ontological, describing the ontological properties of the world, or are they epistemic, resulting merely from our own cognitive limitations? As our knowledge of the world improves, can we expect probabilities to be reabsorbed into a more comprehensive framework of understanding, or will they remain as irreducible descriptions of the behavior of some natural systems (Gilles 2000)?

[262] Whether science itself can have certain metaphysical commitments is a hotly debated issue (cf. Chakravartty 2013), but for the purpose of the argument, it is not necessary to address the question in the present context.

The epistemological account of chance proposed in the present dissertation opens the door to future research on novel questions regarding ontology and chance. My account of stochastic explanation (SE) says that chance plays an essential role in some biological explanations because it provides abstraction (cf. Chapter 3). What ontological considerations come into play in my epistemological study of chance? And, following that, what kind of ontology of chance can be developed from the results of my epistemological analysis? I suggest adopting a naturalistic stance for developing a metaphysics of chance, that is to say, a metaphysics that foregrounds the importance of scientific practice for metaphysical inquiry, or, in Chakvravartty's words, a metaphysics "inspired and constrained by the output of our best science" (2013, p. 33). However, we must specify what "naturalistic" means in this case.

Intuitively, it is often said that a naturalized metaphysics is a kind of metaphysics that is based on, or/and constrained by, scientific practice. But the meaning of "based on" and "constrained by" is still a matter of discussion within the metaphysics of science. By contrast, when I talk about a naturalized metaphysics, I have clearly in mind my proposition for the relationship between science and metaphysics. A naturalized metaphysics of chance must build on scientific explanations that can be conceived as stochastic explanations (SE). Indeed, taking SEs as a starting point, my question is, what is the discriminating factor between a process that can be explained by an SE and one that cannot?

One of the novelties of this line of research is that it proceeds from an epistemological account of chance in order to seek out and elaborate on a compatible naturalized metaphysics. This is somewhat unusual. Philosophical work often starts with an analysis of the ontological status of chance, and only then deals with epistemological problems and peculiarities. Millstein (2011), for example, defines the various meanings of chance in evolutionary biology by referring to the peculiar features of (chancy) processes this concept can refer to. She talks about chance as indiscriminate sampling, chance as the encounter of independent causal chains, chance as a process that does not have a predetermined end, and chance as a process that is independent of the generally adaptive direction of natural selection. She goes into epistemological analysis only in the second part of her work, wondering what kind of causes are permitted, ignored or/and forbidden in the ontological definitions of chance just provided.

By contrast, my suggestion for future work is to completely overturn the analysis normally pursued when studying chance in biology. Indeed, it will not be my intention to begin from an ontological analysis of what chance is in biological processes and only later move to epistemological questions that might arise from that first study. My proposition will start from the epistemological account of chance defended in this dissertation, stochastic explanations,

that are good – even better than others, sometimes – not in spite of the presence of chance, but thanks to it. Starting from this solid epistemological basis, my aim will be to develop a naturalistic ontology of phenomena related to chance by addressing the following questions: Why can some phenomena can be explained by SE and others cannot? Could this epistemological discriminator be the trace of an ontological one? Could phenomena explainable by SE have some kind of property that phenomena not explainable by SE do not? If so, what kind of properties related to chance are we talking about? In this future work, I would hope for a balance between the ontological and epistemological dimensions – balance that has been deliberately held at bay in the elaboration of this present work.

Bibliography

Literature on philosophy of science and philosophy of biology: chance, probability, scientific explanation

Andersen H., (2012), *The Case for Regularity in Mechanistic Causal Explanation*, Synthese, 189 (3): pp. 415–432.

Ankeny R., Chang H., Boumans M., & Boon M, (2011*), Introduction: philosophy of science in practice*, European journal for philosophy of science, 1(3), pp. 303-307.

Ariew A., Rice C., & Rohwer Y., (2015), *Autonomous-statistical explanations and natural selection*, British Journal for the Philosophy of Science, 66(3), pp. 635-658.

Ariew A., Rohwer Y., and Rice C., (2017), *Galton, reversion and the quincunx: The rise of statistical explanation*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 66, pp. 63-72.

Baetu T., (2015), *The completeness of mechanistic explanations*, Philosophy of Science, 82(5), pp. 775-786.

Id., (2019), *Mechanisms in molecular biology*, Cambridge University Press.

Bechtel W., and Abrahamsen A., (2005), *Explanation: A mechanist alternative*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36(2), pp. 421-441.

Bechtel W. and Richardson R., (1993), *Discovering complexity*, Princeton University Press.

Id., (2010), *Discovering complexity: decomposition and localization as strategies in scientific research*, Cambridge, Mass, MIT press.

Boniolo G., and Campaner R., (2020), *Life Sciences for Philosophers and Philosophy for Life Scientists: What Should We Teach?*, Biological Theory, 15(1), pp. 1-11.

Bogen J., (2005), *Regularities and Causality; Generalizations and Causal Explanations*', Studies in History and Philosophy of Biological and Biomedical Sciences, 36, pp. 397–420.

Boone W., and Piccinini G., (2016), *Mechanistic abstraction*, Philosophy of Science, 83(5), pp. 686-697.

Braillard P. and Malaterre C., (2015), *Explanation in Biology: An Introduction*, in *Explanation in Biology*. *An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*, History, Philosophy and Theory of the Life Sciences, Springer.

Brigandt I., (2013), *Explanation in Biology: Reduction, Pluralism, and Explanatory Aims*, Science & Education January, Volume 22, Issue 1, pp. 69–91.

Brigandt I., Green S., and O'Malley M., (2016), *Systems biology and mechanistic explanation*.

Bruce V. and Young A., (1986), *Understanding face recognition*, British Journal of Psychology 77, pp. 305-327.

Callebaut W., (1993), *Taking the naturalistic turn, or how real philosophy of scince is done*, Chicago: University of Chicago Press.

Id. (2005), *Again, what the philosophy of biology is not*, Acta Biotheoretica, 53, pp. 93-122.

Cartwright N., (2004), *Causation: One word, many things*, *Philosophy of Science, 71*, pp. 805-819.

Chakravartty A., (2017), *Scientific Realism*, Stanford Encyclopedia of Philosophy.

Cath Y., (2016), *Reflective equilibrium*, Cappelen, Gendler and Hawthorne, pp. 213-230.

Cerredo V. (2020), *Variational Probabilities and Developmental Propensities. A Philosophical Study of Chance in Evolutionary Variation*, PhD dissertation, University of Madrid.

Craver C., (2006), *When Mechanistic Models Explain*, Synthese, 153, pp. 355–76.

Id., (2007), *Explaining the brain. Mechanisms and the mosaic unity of neuroscience*, Oxford: Oxford University Press.

Id., (2008), *Physical Law and Mechanistic Explanation in the Hodgkin and Huxley Model of the Action Potential*, Philosophy of Science, 75, pp. 1022–33.

Id., (2014), *The Ontic Account of Scientific Explanation*, in (eds.) Kaiser M., Scholz O., Plenge D., and Hüttemann A., *Explanation in the Special Sciences: The Case of Biology and History*, Dordrecht: Springer, pp. 27–52.

Craver C., and Kaplan D., (2020), *Are more details better? On the norms of completeness for mechanistic explanations*, The British Journal for the Philosophy of Science, 71(1), pp. 287-319.

Darden L., and Craver C., (2002), *Strategies in the interfield discovery of the mechanism of protein synthesis*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, *33*(1), pp. 1-28.

Denny M., and Gaines S., (2002). *Chance in biology: using probability to explore nature*. Princeton University Press.

DesAutels L., (2014), *Mechanism and Chance: Toward an Account of Stochastic Mechanism for the Life Sciences*, (Doctoral dissertation).

Id., (2015), *Toward a propensity interpretation of stochastic mechanism for the life sciences*, Synthese, *192*(9), pp. 2921-2953.

Douglas H., (2009), *Reintroducing prediction to explanation*, Philosophy of Science, 76(4), pp. 444-463.

Diacoins P. and Skyrms B., (2017), *Ten Great Ideas about Chance*, Princeton University Press.

Drouet I. and Merlin F., (2015)**,** *The Propensity Interpretation of Fitness and the Propensity Interpretation of Probability*, *Erkenntnis*, 80 (3): pp. 457-468. https://doi.org/10.1007/s10670-014-9681-2

Dudley J., (2012), *Aristotle's Concept of Chance*, NY: State University of New York Press.

Earman J, (1986), *A Primer on Determinism*, Dordrecht, Reidel Publishing Company.

Eigenbrode S. *et al* (2007), *Employing philosophical dialogue in collaborative science*, BioScience 57.1, pp. 55-64.

Ellis-Davies G., (2013), *A chemist and biologist talk to each other about caged neurotransmitters*, Beilstein journal of organic chemistry, 9.1, pp. 64-73.

Feller W., (1968), *An Introduction to Probability Theory and Its Applications*, New York: John Wiley & Sons.

Fine A (1984), *The natural ontological attitude*, In Scientific Realism (ed. J. Leplin), University of California Press, Berkeley, CA., pp. 83–107.

Franklin-Hall L., (2016), *New mechanistic explanation and the need for explanatory constraints*, In Scientific composition and metaphysical ground, Palgrave Macmillan, London, pp. 41-74.

Fodor J, (1976), *The language of thought*, Hassocks.

Garfinkel A., (1981), *Forms of Explanation*, Yale University Press, New Haven, CT.

Green, S. and Jones N., (2016), *Constraint-based reasoning for search and explanation: Strategies for understanding variation and patterns in biology*. Dialectica, 70(3), pp. 343-374.

Galavotti M. (2005), *A Philosophical Introduction to Probability*, Stanford, CSLI Publications.

Gayon J., (2009), *Philosophy of biology: an historico-critical characterization*, In: Brenner A, Gayon J (eds), French studies in the philosophy of science: contemporary research in France. Springer, Dordrecht, pp. 201–212.

Id., (2005), *Chance, explanation, and causation in evolutionary theory*, History and philosophy of the life sciences, pp. 395-405.

Gebharter A., (2013), *Solving the flagpole problem*, Journal for general philosophy of science, *44*(1), pp. 63-67.

Gervais R., (2013), *Explanation in the cognitive sciences and biology. Mechanisms, laws, and their explanatory virtues*, Ph.D. Dissertation, Ghent University.

Gilles D., (2000), *Philosophical Theories of Probability*, Routledge London and New York.

Glennan S., (1996), *Mechanisms and the Nature of Causation*, Erkenntnis 44 (1): pp. 49-71.

Id., (1997), *Probable causes and the distinction between subjective and objective chance*, Noûs, 31(4), pp. 496-519.

Id., (2002), *Rethinking Mechanistic Explanation*, Philosophy of Science 69 (3), pp. S342-

S353.

Id., (2005), *Modeling mechanisms*, Studies in the History and Philosophy of Science Part C. 56(2), pp. 443-464.

Id., (2010), *Ephemeral Mechanisms and Historical Explanation*, Erkenntnis, 72, (2): pp. 251-266.

Id., (2017), *The new mechanical philosophy*, Oxford University Press.

Godfrey-Smith P., (2009), *Abstractions, idealizations, and evolutionary biology*, In Mapping the future of biology, Springer, Dordrecht, pp. 47-56.

Green, S., and Jones, N., (2016), *Constraint-based reasoning for search and explanation: Strategies for understanding variation and patterns in biology*, Dialectica, 70(3), pp. 343-374.

Hacking I., (2001), *An introduction to probability and inductive logic*, Cambridge university press.

Id., (2006), *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*, Cambridge University Press.

Handfield T., (2012), *A philosophical Guide to Chance: Physical Probability*, Cambridge University Press.

Hanson N., (1959), *On the symmetry between explanation and prediction*, The Philosophical Review, 68(3), pp. 349-358.

Henning B., and Scarfe A., (Eds.), (2013), *Beyond mechanism: Putting life back into biology*. Rowman & Littlefield.

Hempel C., (1965), *Aspects of scientifi c explanation. And other essays in the philosophy of science*, New York: Free.

Hempel C., and Oppenheim P., (1948), *Studies in the logic of explanation*, Philosophy of Science,

15 (2), pp. 135–175.

Holmes T., (2021), *Unifying statistically autonomous and mathematical explanations*, Biology & Philosophy, 36(3), pp. 1-22.

Hull D., (1969), *What philosophy of biology is not*, Journal of the History of Biology, 2(1), pp. 241-268.

Humphreys P., (1989), *The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences*. Princeton, NJ: Princeton University Press.

Huneman, P., (2010), *Topological explanations and robustness in biological sciences*, Synthese, 177(2), pp. 213-245.

Illari P., (2013), *Mechanistic explanation: Integrating the ontic and epistemic*, Erkenntnis, 78(2), pp. 237-255.

Illari P., and Williamson J., (2011), *Mechanisms Are Real and Local*', in Illari P., Russo F., and Williamson J. (eds), Causality in the Sciences, Oxford: Oxford, University Press, pp. 818–844.

Id., (2012), *What is a mechanism? Thinking about mechanisms across the sciences*, European Journal for Philosophy of Science, 2(1), pp. 119-135.

Jackson F. and Pettit P., (1992), *In Defense of Explanatory Ecumenism*, Economics and Philosophy, N. 8, pp. 1-21, Printed in the United States of America.

Jones M., (2005), *Idealization and abstraction: A framework*, In Idealization XII: Correcting the model, Brill Rodopi, pp. 173-217.

Id., (2013), *Don't blame the idealizations*, Journal for general philosophy of science, 44(1), pp. 85-100.

Kaplan D., (2011), *Explanation and Description in Computational Neuroscience*, Synthese 183 (3), pp. 339–73.

Kaplan D., and Craver C., (2011), *The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective*', Philosophy of Science, 78, pp. 601–27.

Kaplan J., and Turkheimer E., (2021), *Galton's Quincunx: Probabilistic causation in developmental behavior genetics*. Studies in History and Philosophy of Science, 88, pp. 60-69.

Kellert S., Longino H., and Waters K. (eds.), (2006), *Scientific pluralism*, Vol. 19, U of Minnesota Press.

Khalidi M., (1993), *Carving nature at the joints*, Philosophy of Science, *60*(1), pp. 100-113.

Kim J., (1988), *What is "naturalized epistemology?"*. Philosophical perspectives, 2, pp. 381-405.

Krickel B., (2018), *A regularist approach to mechanistic type-level explanation*, The British Journal for the Philosophy of Science, 69(4), pp. 1123-1153.

Fagot-Largeault A., (2009), *The Legend of Philosophy's Striptease (Trends in Philosophy of Science)*, In French Studies In The Philosophy Of Science Springer, Dordrecht, pp. 25-49.

Laplane L., (2016), *Cancer stem cells*. Harvard University Press.

Laplane L. *et al* (2019), *Why science needs philosophy*, Proceedings of the National Academy of Sciences 116.10, pp. 3948-3952.

Lee R. (1979), *A critical analysis of the thesis of the symmetry between explanation and prediction: including a case study of evolutionary theory*, Doctoral dissertation, University of British Columbia.

Levy A, and Bechtel W., (2013), *Abstraction and the organization of mechanisms*, Philosophy of science, 80.2, 241-261.

Levy A., (2014), *What was Hodgkin and Huxley's achievement?*, The British Journal for the Philosophy of Science, 65(3), pp. 469-492.

Id., (2018), *Idealization and abstraction: refining the distinction*. Synthese, pp. 1-18.

Levins R., (1966), *The strategy of model building in population biology*, American scientist, 54(4), pp. 421-431.

Lewis D., (1986), *Causal Explanation*, in Philosophical Papers, vol. 2, Oxford, Oxford University Press, pp. 214–240.

2009: 43–63.

Machamer P., (2004), *Activities and Causation: The Metaphysics and Epistemology of Mechanisms*, International Studies in the Philosophy of Science 18 (1), pp. 27-39.

Machamer P., Darden L., and Craver C., (2000), *Thinking about mechanisms*, Philosophy of science, 67(1), pp. 1-25.

Malaterre C., Pulizzotto D., and Lareau F., (2020), *Revisiting three decades of Biology and Philosophy: a computational topic-modeling perspective*, Biology & Philosophy, 35(1), pp. 1-25.

Matthewson J., (2020), *Detail and generality in mechanistic explanation*, Studies in History and Philosophy of Science Part A, 80, pp. 28-36.

Matthen M., (2009), *Drift and "statistically abstractive explanation"*, Philosophy of Science, 76(4), pp. 464-487.

Matthiessen D., (2017), *Mechanistic explanation in systems biology: Cellular networks*, The British Journal for the Philosophy of Science, 68(1), pp. 1-25.

Mellor D., (2005), *Probability: A Philosophical Introduction*, Routledge London and New York.

Mayr E., (1961), *Cause and effect in biology: kinds of causes, predictability, and teleology are viewed by a practicing biologist*, Science, 134(3489), pp. 1501-1506.

Merlin F., (2015a), *Developmental Noise: Explaining the Specific Heterogeneity of Individual Organisms*, in *Explanation in Biology*. *An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*, History, Philosophy and Theory of the Life Sciences, Springer.

Miller T. *at al* (2008), *Epistemological Pluralism: Reorganizing Interdisciplinary Research*, Ecology and Society, 13(2): 46.

Mitchell S.; (1997), *Pragmatic Laws*, Philosophy of Science 64 (4), pp. 468-479.

Id., (2000), *Dimensions of Scientific Law*, Philosophy of Science 67 (2), pp. 242-265.

Id., (2002), *Biological Contingency and Laws*, Erkenntnis 57 (3), pp. 329-350.

Id., (2003), Biological Complexity and Integrative Pluralism. Cambridge University Press, New York.

Morange M., (2015), *Is There an Explanation for … the Diversity of Explanations in Biological Studies?*, in *Explanation in Biology*. *An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*, History, Philosophy and Theory of the Life Sciences, Springer.

Morrison M., (2009), *Understanding in Physics and Biology*, In De Regt, Leonelli, and Eigner (eds.), *Scientific Understanding philosophical Perspectives*, pp 123 - 145.

Newton-Smith W., (2000), *Explanation*, in Newton-Smith W. (ed.), *A Companion to the Philosophy of Science*, Oxford, Blackwell, pp. 127–133.

Nicholson D., (2012), *The concept of mechanism in biology*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43(1), pp. 152-163.

Nicholson D., and Gawne R., (2014), *Rethinking Woodger's legacy in the philosophy of biology*, Journal of the History of Biology, 47(2), pp. 243-292.

Id., (2015), *Neither logical empiricism nor vitalism, but organicism: what the philosophy of biology was*, History and philosophy of the life sciences, 37(4), pp. 345-381.

O'Rourke M. and Crowley S. (2013), *Philosophical intervention and cross-disciplinary science: the story of Toolboox Project,* Synthese, 190.11, pp. 1937-1954.

Pence C. (2021), *The Rise of Chance in Evolutionary Theory: A Pompous Parade of Arithmetic*, Academic Press / Elsevier (London).

Polger W. and Shapiro L., (2016), *The multiple realization book*. Oxford University Press.

Popper K., (2002[1932]), *The Logic of Scientific Discovery*, Routledge Classics.

Id., (1959). *The propensity interpretation of probability*, The British journal for the philosophy of science, 10(37), 25-42.

Pradeu T., (2011). *The limits of the self: immunology and biological identity*. Oxford University Press.

Id., (2017), *Thirty years of Biology & Philosophy: philosophy of which biology?*, Biology & Philosophy, *32*(2), pp. 149-167.

Id., (2018), Philosophy of biology, in *The Philosophy of Science. A Companion*, pp. 430-463.

Putnam H., (1973), *Reductionism and the Nature of Psychology*, Cognition 2 (1), pp. 131–46.

Id., (1975), *Philosophy and Our Mental Life*, In Philosophical Papers, vol. 2, Mind, Language and Reality, Cambridge: Cambridge University Press, pp. 291–303.

Quine W., (1969), *Epistemology Naturalized*, Columbia University Press, 3, pp. 69-90.

Railton P., (1980), *Explaining Explanation: A Realist Account of Scientific Explanation and Understanding*, PhD dissertation, Princeton University.

Id., (1981), *Probability, Explanation, and Information*, Synthese 48 (2), pp. 233–256.

Rescher N., (1958), *On prediction and explanation*, The british journal for the philosophy of science, 8(32), pp. 281-290.

Rice C., (2015), *Moving beyond causes: Optimality models and scientific explanation*, Noûs, 49(3), pp. 589-615.

Richardson, A. W., & Reichenbach, H. (1938). E*xperience and Prediction: An Analysis of the Foundations and the Structure of Knowledge.*

Ruse M., (1989), *David Hull through two decades*, In What the Philosophy of Biology Is. Springer, Dordrecht, pp. 1-15.

Skillings J., (2015), *Mechanistic explanation of biological processes*, Philosophy of Science, 82(5), pp. 1139-1151.

Skipper R., and Millstein R., (2005), *Thinking about evolutionary mechanisms: natural selection*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 36(2), pp. 327-347.

Théry F., (2015), *Explaining in Contemporary Molecular Biology: Beyond Mechanisms*, in *Explanation in Biology*. *An Enquiry into the Diversity of Explanatory Patterns in the Life Sciences*, History, Philosophy and Theory of the Life Sciences, Springer.

Id., (2016). *La face cachée des cellules: quand le monde des ARN bouscule la biologie*. Editions Matériologiques.

Salmon C., (1971), *Statistical explanation*, In Salmon C., (Ed.), *Statistical explanation and Statistical relevance,* Pittsburgh: University of Pittsburgh Press, pp. 29–87.

Id., (1984), *Scientific explanation and the causal structure of the world,* Princeton: Princeton University Press.

Id., (1989), *Four Decades of Scientific Explanation*, Minneapolis: University of Minnesota Press

Id., (1999), *The spirit of logical empiricism: Carl G. Hempel's role in twentieth-century philosophy of science*, Philosophy of Science, 66(3), pp. 333-350.

Scheffler I., (1957), *Explanation, prediction, and abstraction*, The British Journal for the Philosophy of Science, 7(28), pp. 293-309.

Schiemann G., (2019), *Old and New Mechanistic Ontologies*, In Mechanistic Explanations in Physics and Beyond Springer, Cham, pp. 33-46.

Sellars W., (1963), *Philosophy and the scientific image of man*, Science, perception and reality, 2, pp. 35-78.

Soler L., Zwart S., Lynch M., and Isreal-Jost V., (Eds.), (2014), *Science after the practice turn in the philosophy, history, and social studies of science*. New York: Routledge

Strevens M., (2000), *Do Large Probabilities Explain Better?*, Philosophy of Science, The University of Chicago Press on behalf of the Philosophy of Science Association Vol. 67, No. 3, pp. 366-390.

Id., (2008), *Depth. An account of scientific explanations*, Cambridge: Harvard University Press.

Id., (2019), *Explanation, Abstraction, and Difference-Making*, Philosophy and Phenomenological Research, 99(3), pp. 726-731.

Suárez, M. (2020). *Philosophy of Probability and Statistical Modelling*. Cambridge University Press.

van Thiel G. and van Delden J., (2010), *Reflective Equilibrium as a Normative Empirical Model*, Ethical Perspectives 17, pp. 183–202.

Vorms M., (2018), *Theories and models*, The philosophy of science: A companion, pp. 171-224.

Waters C., (2014), *Shifting Attention from Theory to Practice in Philosophy of Biology*, New Directions in the Philosophy of Science, New Directions in the Philosophy of Science, 2014, pp. 121 - 139.

Weber E. and Van Bouwel J., (2007), *Assessing the explanatory power of causal explanations*, in *Rethinking explanation*, edited by Persson J. and Ylikoski P., Boston Studies in the Philosophy of Science 252, Springer, pp. 109-118.

Weslake B., (2010), *Explanatory depth*, Philosophy of Science 77.2, pp. 273-294.

Wimsatt W., (1972), *Complexity and organization*, In K. F. Schaffner, & R. S. Cohen (Eds.), PSA 1972: Proceedings of the Philosophy of Scgoience Association (pp. 67–86). Dordrecht: Reidel.

Wouters A, (2007), *Design explanation: determining the constraints on what can be alive*, Erkenntnis 67.1, pp. 65-80.

Ylikoski P. and Kuorikoski J., (2010), *Dissecting explanatory power*, Philosophical studies 148.2, pp. 201-219.

Ylikoski P., (2001), *Understanding interests and causal explanation*, PhD dissertation, Department of Moral and Social Philosophy, University of Helsinki.

Id., (2007), *The Idea of Contrastive Explanandum*, in Rethinking Explanation, edited by Persson J. and Ylikoski P., Boston Studies in the Philosophy of Science 252, Springer, pp. 27-42.

Molecular and cellular biology

Asiaee, A., Abrams, Z. B., Nakayiza, S., Sampath, D., & Coombes, K. R. (2020), *Explaining Gene Expression Using Twenty-One MicroRNAs*, Journal of Computational Biology, 27(7), 1157-1170.

Bapteste E. and Dupré J., (2012), *Towards a Processual Microbial Ontology*, Biology & philosophy, vol. 28, pp. 379–404.

Bickle j., (2003), *Philosophy and Neuroscience. A ruthlessly reductive account.* Dordrecht: Kluwer Academic Publishers:

Brangwynne C., Koenderink H., MacKintosh C., and Weitz D., (2009), *Intracellular transport by active diffusion*, Trends in cell biology, 19(9), pp. 423-427.

Boersma S. *et al* (2019), *Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding*, Cell, Volume 178, Issue 2, pp. 458-472.

Buck L., and Axel R., (1991), *A novel multigene family may encode odorant receptors: a molecular basis for odor recognition*, Cell 65, pp. 175–187.

Bressloff P. and Newby J., (2013), *Stochastic models of intracellular transport*, Reviews of Modern Physics, 85(1), 135.

Chargaff E., Lipshitz R., and Green C., (1952), *Composition of the desoxypentose nucleic acids of four genera of sea-urchin*, Journal of Biological Chemistry, 195(1), pp. 155-160.

Crick F., (1966), *Of molecules and men*, Seattle: University of Washington Press.

Cummins R., and Roth M., (2009). *Traits have not evolved to function the way they do because of a past advantage*. *Contemporary debates in philosophy of biology*, 72-85.

Delbruck M., (1971), *Aristotle-totle-totle*, reprinted from Of Microbes and Life, Monod J. And Borek E., eds. Columbia University Press.

Dobzhansky T., (1973), *Nothing in biology makes sense except in the light of evolution*, The american biology teacher, *35*(3), pp. 125-129.

Doolittle W., (2018), *We simply cannot go on being so vague about 'function'*, *Genome biology*, *19*(1), pp. 1-3.

Dupré (2015), *A Process Ontology for Biology. Functions are just Fast Processes and Structures are (relatively) Slow Processes*, Physiology News, Issue 100, pp. 33-34.

Elson D., and Chargaff E., (1952), *On the desoxyribonucleic acid content of sea urchin gametes*, Experientia, 8(4), pp. 143-145.

Endy D. and Brent R., (2001), *Modelling cellular behavior*, Nature, Vol. 409, pp. 391–395.

Falk R, (1986), What is a Gene?, Studies In History and Philosophy of Science Part A 17, pp. 396–406.

Id., (2000), *The Gene: A Concept in Tension*, in Beurton P., Falk R. and Rheinberger H. (eds.), The Concept of the Gene in Development and Evolution, Cambridge, Cambridge University Press, pp. 317-348.

Id., (2010) *What is a Gene? – Revisited*, Studies in History and Philosophy of Biological and Biomedical Sciences, 41, pp. 396–406.

Firth A., and Brierley I., (2012), *Non-canonical translation in RNA viruse*, The Journal of general virology, 93(Pt 7), 1385.

Fuss S., and Ray A., (2009), *Mechanisms of odorant receptor gene choice in Drosophila and vertebrates*, Molecular and Cellular Neuroscience, 41(2), pp. 101-112.

Gehring W., (1998), *Master Control Genes in Development and Evolution: The Homeobox Story*, the terry Lectures Series, Yale University Press

Griffiths P., Neumann-Held E.M., (1999), *The Many Faces of the Gene*, BioScience, Vol. 49, No. 8, pp. 656-662.

Griffiths P., Stotz K., (2007), *Gene*, Cambridge Companion to Philosophy of Biology. M. Ruse and D. Hull. Cambridge, Cambridge University Press, pp. 85- 102.

Gō, N, (1983), *Protein folding as a stochastic process*, Journal of Statistical Physics, 30(2), pp. 413-423.

Gupta P., Fillmore C., Jiang G., Shapira, S., Tao K., Kuperwasser C., and Lander E., (2011). *Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells*. Cell, 146(4), pp. 633-644.

Guttinger, S. (2021), *Covid-19 and the need for more history and philosophy of RNA*, History and philosophy of the life sciences, 43(2), 1-6

Hall B., (2003), *Unlocking the Black Box between Genotype and Phenotype: Cell Condensations as Morphogenetic (modular) Units,* Biology and Philosophy, 18, pp. 219–247.

Hansmann U., and Okamoto Y., (1999), *New Monte Carlo algorithms for protein folding. Current opinion in structural biology*, 9(2), pp. 177-183.

Hiesinger P., and Hassan A., (2018), *The evolution of variability and robustness in neural development*, Trends in Neurosciences, 41(9), pp. 577-586.

Hodgkin A., and Huxley A., (1952), *A quantitative description of membrane current and its application to conduction and excitation in nerve*, The Journal of physiology, *117*(4), pp. 500-544.

Honegger K., and de Bivort B., (2018), *Stochasticity, individuality and behavior*, Current Biology, 28(1), pp. R8-R12.

Ingolia N., (2014), *Ribosome profiling: new views of translation, from single codons to genome scale*, Nature reviews genetics, 15(3), pp. 205-213.

Jacob F. and Monod J., (1961), *Genetic regulatory mechanisms in the synthesis of proteins*, Journal of Molecular Biology, 3, pp. 318–356.

Jamieson A. and Radick G., (2017), *Genetic Determinism in the Genetics Curriculum. An Exploratory Study of the Effects of Mendelian and Weldonian Emphases*, Science and Education, 26 (10). pp. 1261-1290.

Jerne N., (1955), *The natural-selection theory of antibody formation*, Proceedings of the National Academy of Sciences of the United States of America 41.11, pp. 849-857.

Kaiser M., (2011), *The Limits of Reductionism in the Life Sciences*, History and Philosophy of the Life Sciences, Vol. 33, No. 4, Causation and Disease, pp. 453-476.

Id., (2012), *Why is Time to Move Beyond Nagelian Reduction*, in Dieks D., Gonzalez W., Hartmann S., Stöltzner M., Weber M. (eds), *Probabilities, Laws, and Structures. The Philosophy of Science in a European Perspective*, vol 3. Springer, Dordrecht.

Id., (2015), *Reductive Explanation in the Biological Sciences*, History, Philosophy and Theory of the Life Sciences, Springer.

Id., (2018): *Individuating Part-whole Relations in the Biological World*, in: Bueno O., Chen R., Fagan M. (eds.): Individuation, Process and Scientific Practices. Oxford: Oxford University Press.

Id., (2019), *Normativity in the philosophy of Science*, Metaphilosophy, 50 (1-2), pp. 36-62.

Kaiser M., and Krickel B., (2017), *The metaphysics of constitutive mechanistic phenomena*, The British Journal for the Philosophy of Science, 68(3), pp. 745-779.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., Hudspeth, A. J. (2013), *Principles of Neural Science*, (5th ed.) McGraw-Hill Inc.

Keeling D., Garza P., Nartey C., and Carvunis A., (2019), *Philosophy of Biology: The meanings of 'function' in biology and the problematic case of de novo gene emergence*, Elife, *8*, e47014.

Keller F., (2000a), *Decoding the Genetic Program, in The Concept of gene in Development and Evolution*, Cambridge University Press, pp. 159-177.

Id., (2000b), *Models of and models for: Theory and practice in contemporary biology*, Philosophy of science, 67, pp. S72-S86.

Keller F. and Harel D., (2007) Beyond the gene, PLoS ONE (2), 11, e1231, pp. 1-11.

Kitcher P. and Salmon W., (1987), *Van Fraassen on Explanation*, The Journal of Philosophy, Vol. 84, No. 6, pp. 315-330.

Kitcher P., (2003). *In Mendel's mirror: Philosophical reflections on biology*. Oxford University Press.

Kupiec J., (1997), *A Darwinian theory for the origin of cellular differentiation*, Molecular and General Genetics MGG, 255(2), pp. 201-208.

Id., (2009a), *Le hasard au coeur de la cellule*, Paris, Syllepse.

Id., (2009b), *The Origin of Individual*, World Scientific.

Id., (2014a), *Cell differentiation is a stochastic process subjected to natural selection*, In Minelli A., Pradeu T. (eds.), *Towards a Theory of Development*, pp. 155-173.

Id., (2014b), *Toward a probabilistic biology*, Math. Struct. in Comp. Science, vol. 24, iss. 3, e240307, Cambridge University Press, pp. 1-3.

Id., (2019), *Et si le vivant était anarchique : La génétique est-elle une gigantesque arnaque ?*, Éditions Les Liens qui libèrent.

Id., ([1983]2020), *A probabilistic theory for cell differentiation, embryonic mortality and DNA c-value paradox*, Organisms. Journal of Biological Sciences, 4(1), pp. 80-85.

Kupiec J. and Sonigo P., (2009), Né dio né genoma, Milano, Elèuthera.

Lewens T., (2004), *Organisms and artifacts: Design in nature and elsewhere*, The MIT press.

Lewontin R., (2002), *The Triple Helix: Gene, Organism, and Environment*, Harvard University Press.

Loison, L. (2022). *The environment: An ambiguous concept in Waddington's biology*, Studies in History and Philosophy of Science, 91, pp. 181-190.

Merlin F., (2009), *Le hasard et les sources de la variation biologique: analyse critique d'une notion multiple*, Ph.D. Thesis, Institut d'Histoire et de Philosophie des Sciences et des Techniques, University of Paris 1, Paris.

Id., (2010), *Evolutionary Chance Mutation: A Defense of the Modern Synthesis' Consensus View*, Philosophy & Theory in Biology, Volume 2(e103).

Id., (2013), *Mutations et aléas. Le hasard dans la théorie de l'évolution*, Paris: Hermann.

Id., (2015b), The *Multi-faceted Idea of Chance in Darwin's Writings*, Science & Education 25, pp. 1159–1164 (2016). https://doi.org/10.1007/s11191-016-9869-8.

Id., (2016a), *Weak Randomness at the Origin of Biological Variation: The Case of Genetic Mutations*, in Ramsey and Pence (ed.), Chance in Evolution, The University Chicago Press, Chicago & London, pp. 176–195.

Miné-Hattab J. and Taddei A., (2019), *Physical principles and functional consequences of nuclear compartmentalization in budding yeast*, Current opinion in Cell Biology, 59, pp. 105-113.

Merlin F. and Loison L. (2021), *Cybernétique enzymatique. Un inédit de Jacques Monod* (édition critique).

Moore P., (2012), *How should we think about the ribosome?*, Annual review of biophysics, 41, pp. 1-19.

Mombaerts P., (2004), *Odorant receptor gene choice in olfactory sensory neurons: the one receptor–one neuron hypothesis revisited*. Current Opinion in Neurobiology. 14, pp. 31–36.

Morange M., (2000), A History of Molecular Biology, Harvard University Press.

Id., (2006), Post-genomics, between Reduction and Emergence, Synthese, 151, pp. 355-360.

Id., (2009b), Articulating Different Modes of Explanation: The Present Boundary in Biological Research, in Barberousse, Morange and Pradeu (Eds.), *Mapping the Future of Biology. Evolving Concepts and Theories*, Boston Studies in the Philosophy of Science, Springer.

Id., (2019), *A Time to Model and a Time to Experiment*, Journal of Computational Biology, 26(7), pp. 629-636.

Morrison S., Perez S., Qiao Z., Verdi J., Hicks C., Weinmaster G., *et al* (2000), *Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells*, Cell 101, pp. 499-510. Available from: https://doi.org/10.1016/S0092-8674(00)80860-0.

Moss L., (1992), *A Kernel of Truth? On the Reality of the Genetic Program*, Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol.1, Contributed Papers, pp. 335-348.

Moreno A. and Mossio M., (2015), *Biological Autonomy. A philosophical and Theoretical Enquiry*, History, Philosophy and Theory of the Life Sciences, Springer.

Moxon E., Rainey B., Nowak A., and Lenski R., (1994), *Adaptive evolution of highly mutable loci*

in pathogenic bacteria. Current Biology 4. pp 24-33.

Moxon E., Bayliss C., and Hood D., (2006), *Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation*. Annual Review of Genetics 40, pp. 307-333.

Nicholson D., (forthcoming), *What is life revisited.*

Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford University Press.

Peluffo A., (2015), *The "genetic program": Behind the genesis of an influential metaphor*, Genetics, 200(3), pp. 685-696.

Pauli A., Valen E., and Schier A., (2015), *Identifying (non-) coding RNAs and small peptides: Challenges and opportunities*, Bioessays, 37(1), pp. 103-112.

Perlman M., (2010), *Traits have evolved to function the way they do because of a past advantage*" *Contemporary debates in philosophy of biology* (2010): pp. 53-71.

Peterson E., (2016), *The Life Organic: The Theoretical Biology Club and the Roots of Epigenetics*, University of Pittsburgh Press.

Portin P., (1993), *The Concept of the Gene: Short History and Present Status*, The Quarterly Review of Biology, 68 (2), pp. 173-223.

Id., (2009), *The Elusive Concept of the Gene*, Hereditas, 146, pp. 112–117.

Radick G., (2011), *Physics in the Galtonian sciences of heredity*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, *42*(2), pp. 129-138.

Ramsey and Pence (ed.) (2016), *Chance in Evolution*, The University Chicago Press, Chicago & London.

Rheinberger H. and Muller-wille S., *The Gene: From Genetics to Postgenomics*, University of Chicago Press.

Schaffner K. (1967), *Approaches to reduction*, Philosophy of Science, *34* , pp. 137–147.

Id., (1969), *The Watson-Crick model and reductionism*, British Journal for the Philosophy of Science, 20, 325–348.

Id., (1974), *The peripherality of reductionism in the development of molecular biology*, Journal of the History of Biology, 7 (1), pp. 111–139.

Schrödinger E., (2013[1947]), *What is life*, Cambridge University Press.

Spemann H., (1938), *Embryonic Development and Induction*, Yale University Press, New Haven.

Stamos D. N., (2001), *Quantum indeterminism and Evolutionary Biology*, Philosophy of Science, Vol.68, No. 2, pp. 164-184.

Id., (2010), *Quantum indeterminism, mutation, natural selection, and the meaning of life*, Quantum Biochemistry: Electronic Structure and Biological Activity; Wiley-VCH, Weinheim, pp 837-871.

Symmons O., and Raj A., (2016), *What's luck got to do with it: single cells, multiple fates, and biological nondeterminism*, Molecular cell, *62*(5), pp. 788-802.

Szathmáry E., Jordán F., and Pál C., (2001), *Can genes explain biological complexity?*, Science, 292(5520), pp. 1315-1316.

Van Fraassen B., (1980), *The Scientific Image*, The Clarendon Library of Logic and Philosophy.

Noise in gene expression

Avner P. and Heard E., (2001), *X-Chromosome Inactivation: Couning, Choice and Initiation*, Nature Reviews, genetics, Vol. 2, pp. 59-67.

Balázsi G. *et al*, (2011), *Cellular Decision Making and Biological Noise: From Microbes to Mammals*, Cell 144, Elsevier Inc., pp. 910-925.

Beslon G., (2012), *Chance at the Heart of the cell*, Editorial / Progress in Biophysics and Molecular Biology 110, pp. 1–4.

Birnbaum and Kussell (2011), *Measuring cell identity in noisy biological systems*, Nucleic Acids Research, 2011, Vol. 39, No. 21, pp. 9093–9107.

Blake L., and Wu B., (2019), *One message, many translations: Heterogeneity revealed with multicolor imaging*. Molecular cell, 75(1), pp. 3-4.

Borges J., (1962), *Fictions*, Grove Press.

Casali M. and Merlin F., (2020), *Rethinking the role of chance in the explanation of cell differentiation*, in Levine H., Jolly M., Kulkarni P., Nanjundiah V. (eds.), *Phenotypic Switching: Implications in Biology and Medicine*, Academic Press, Elsevier, pp. 23-51.

Cohen L., (2005), *The history of noise [on the 100th anniversary of its birth]*, Signal Processing Magazine, *IEEE* 22, pp. 20-45.

Coffman J., (2014), *On the Meaning of Chance in Biology*, Biosemiotics, 7(3), 377–388, p. 1-12, doi: 10.1007/s12304-014-9206-z.

Corson, F., and Siggia E., (2012), *Geometry, epistasis, and developmental patterning*, Proceedings of the National Academy of Sciences, 109(15), pp. 5568-5575.

Duchesneau F., (2012), *Determinism and probability in the development of the cell theory*, Progress in Biophysics and Molecular Biology 110, pp. 34-40.

Efroni I., Ip P., … Birnbaum D., (2015), *Quantification of cell identity from single-cell gene expression profiles*, Genome Biology, 16:9, pp. 1-12, DOI 10.1186/s13059-015-0580-x.

Eldar A. and Elowitz M., (2010), *Functional roles for noise in genetic circuits*, Review, Nature, Vol. 467, pp. 167-173, doi:10.1038/nature09326.

Ellis R., (2001), *Macromolecular Crowding: Obvious but Underappreciated*, Review, Trends in Biochemical Sciences Vol.26 No.10, pp. 597- 604.

Elowitz M. *et al*, (2002), *Stochastic Gene Expression in a Single Cell*, Reports, Science, volume 297, pp. 1183 – 1186.

Feinberg A., and Irizarry R., (2010), *Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease*, Proceedings of the National Academy of Sciences of the United States of America, 107 Suppl 1(Suppl 1), pp. 1757–1764. doi:10.1073/pnas.0906183107.

Franceschelli S., (2012), *Some remarks on the compatibility between determinism and unpredictability*, Process in Biophysics and Molecular Biology 110, pp. 61-68.

Gillespie D., (2007), *Stochastic Simulation of Chemical Kinetics*, Annual Review of Physical Chemistry 58, pp. 35-55.

Golubev, (2012), *Transition probability in cell proliferation, stochasticity in cell differentiation, and the restriction point of the cell cycle in one package*, Progress in Biophysics and Molecular Biology 110, pp. 87-96.

Huang S., (2009), *Non-genetic heterogeneity of cells in development: more than just noise,* Development 136, pp. 3853-3862, doi:10.1242/dev.035139.

Han R., Huang G., Wang Y., Xu Y., Hu Y., Jiang W., ... and Zheng D., (2016), *Increased gene expression noise in human cancers is correlated with low p53 and immune activities as well as late stage cancer*, Oncotarget, 7(44), 72011.

Harton M. and Batchelor E., (2017), *Determining the limitations and benefits of noise in gene regulation and signal transduction through single cell, microscopy-based analysis*, Journal of Molecular Biology, 429, pp. 1143 – 1154, doi:10.1016/j.jmb.2017.03.007.

Hattori D., Demir E., …, Dickson B., (2007), *Dscam diversity is essential for neuronal wiring and self-recognition*, Nature, 13, 449(7159), pp. 223–227.

Heams T., (2014), *Randomness in Biology*, Mathematical Structures in Computer Science, vol. 24, iss. 3, e240308, Cambridge University Press, pp. 1 – 24, doi:10.1017/S096012951200076X.

Heard E. (2005), *Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome*, Current Opinion in Genetics & Development, Elsevier Ltd, pp. 482-489.

Huang S., (2012), *Tumor progression: Chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution*, Progress in Biophysics and Molecular Biology 110, pp. 69-86.

Jaeger S. and Fernandez B. (2012), *Epigenetic Aspects of Lymphocyte Antigen Receptor Gene Rearrangement or 'When Stochasticity Completes Randomness'*, Inserm Immunology, Blackwell Publishing Ltd, Immunology, 139, pp. 141–150.

Jin Y. and Li H. (2018), *Revisiting Dscam diversity: lessons from clustered protocadherins*, Cellular and Molecular Life Sciences, pp. 1-14.

Jin Y., Dong H., Shi Y., Bian L., (2017), *Mutually exclusive alternative splicing of pre-mRNAs*, Wiley Interdisciplinary Reviews: RNA 9.3, e1468, pp. 1-14

Johnston R. and Desplan C. (2010). *Stochastic mechanisms of cell fate specification that yield random or robust outcome.* Annual review of cell and developmental biology, 26, pp. 689-719.

Kærn *et al* (2005), *Stochasticity in gene expression: from thoeries to phenotypes*, Nature Reviews, Genetics, Volume 6, pp. 451- 464.

Kint C., Verstraeten N., Fauvart M., and Michiels J., (2012). *New-found fundamentals of bacterial persistence*, Trends in microbiology, 20(12), pp. 577-585.

Kitcher P., (1984), *1953 and all that. A Tale of Two Sciences*, The Philosophical Review, XCIII, No. 3, pp. 335 – 373.

Klingenberg C., (2005), *Developmental Constraints, Modules, and Evolvability*, In Hallgrimsson & Hall (eds.), Variation. A Central Concept in Biology, pp. 219-247.

Latter B., (1964), *Selection for a Threshold Character in Drosophila*, I, Analysis of the Phenotypic Variance of the Underlying Scale, Genetics Research 5, pp. 198-210.

Landau D., Clement K., Ziller, M., Boyle P., Fan J., Gu H., ... and Kotliar, D., (2014), *Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia*, Cancer cell, 26(6), pp. 813-825.

Lenormand T. *et al* (2009), *Stochasticity in Evolution*, Review, Trends in Ecology and Evolution Vol.24 No.3, pp. 157-165.

Levine H., Jolly M., Kulkarni P., and Nanjundiah V., (Eds.), (2020). *Phenotypic Switching: Implications in Biology and Medicine*. Academic Press.

Lestas I., Vinnicombe G., Paulsson J., (2010), *Fundamental limits on the suppression of molecular fluctuations*. Nature 467, pp. 174-178.

Losick R. and Desplan C. (2008), *Stochasticity and Cell Fate*, Science, 320(5872): 65–68. doi:10.1126/science.1147888

Liu J. et al, (2006), *Intrinsic Disorder in Transcription Factors*, Biochemistry, 45(22), 6873–6888, pp. 1-38, doi:10.1021/bi0602718.

Maamar H., Raj A., and Dubnau D., (2007), *Noise in gene expression determines cell fate in Bacillus subtilis*, Science, 317, pp. 526–529.

Maheshri, N. & O'Shea, E. (2007), *Living with noisy genes: How cells function reliably with inherent variability in gene expression*, Annual Review of Biophysics and Biomolecular Structures, 36, pp. 413–434.

McAdam, H. & Arkin, A. (1997). *Stochastic mechanisms in gene expression*. Proceedings of the National Academy of Sciences USA 94. pp. 814-819.

Melamud E. and Moult J., (2009), *Stochastic noise in splicing machinery*, Nucleic Acids Research, Vol. 37, No. 14, pp. 4873–4886.

Meyer M. and Roeder A., (2014), *Stochasticity in plant cellular growth and patterning*, Frontiersin Plant Science, Plant Evolution and Development, Volume 5, Article 420, pp. 1 – 14.

Mettetal J. and van Oudenaarden A., (2007), *Necessary Noise*, Science, Volume 317, pp. 463-464.

Morizot B., (2012), *Chance: From metaphysical principle to explanatory concept. The idea of uncertainty in a natural history of knowledge*, Progress in Biophysics and Molecular Biology, 110, pp. 54-60.

Morris S., (2019), *The evolving concept of cell identity in the single cell era*, The Company of Biologists Ltd, Development, 146, dev169748, pp. 1-5.

Paldi A., (2003), *Stochastic gene expression during cell differentiation: order from disorder?*, Cellular and Molecular Life Sciences, CMLS, 60, pp. 1775–1178.

Id., (2012), *What makes the cell differentiate?*, Progress in biophysics and molecular biology, 110(1), pp. 41-43.

Id., (2018), *Conceptual Challenges of the Systemic Approach in Understanding Cell Differentiation*, System Biology, Springer.

Id., (2020). *Random walk across the epigenetic landscape*. in (eds.) Levine H., Jolly M., Kulkarni P., Nanjundiah V., Phenotypic Switching. Implications in Biology and Medicine, Elsevier; pp. 53-76.

Pilpel Y., (2011), *Noise in Biological Systems: Pros, Cons, and Mechanisms of Control*, in Castrillo J.,and Oliver S., (eds.), Yeast Systems Biology, Methods in Molecular Biology 759, pp. 407-425.

Pujadas E. and Feinberg A., (2012), *Regulated Noise in the Epigenetic Landscape of Development and Disease*, Cell 148, Elsevier Inc, pp. 1123-1131.

Raj A. and van Oudenaarden A., (2008), *Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences*, Cell 135, Elsevier Inc., pp. 216-226.

Raser J. and O'Shea E., (2005), *Noise in Gene Expression: Origins, Consequences, and Control*, Science, 23, 309(5743): 2010–2013, pp. 1-11.

Reeve E. and Robertson F., (1953), *Analysis of Environmental Variability in quantitative inheritance*, Nature, 171, pp. 874-875.

Reiter M., Kirchner B., …, Pfaffl W., (2011), *Quantification noise in single cell experiments*, Nucleic Acids Research, 2011, Vol. 39, No. 18, e124, pp. 1-11.

Rosenberg A., (2001), *Reductionism in a Historical Science*, Philosophy of Science, 68, pp. 135 – 163.

Ruphy S., (2016), *Scientific Pluralism Reconsidered: A New Approach to the (Dis)Unity of Science*, University of Pittsburgh Press.

Samoilov M., Price G., Arkin A. (2006). *From fluctuations to phenotypes: The physiology of noise*, Science STKE, 366(re 17), pp. 1–9.

Sanchez A. and Golding I., (2013), *Genetic Determinants and Cellular Constraints in Noisy Gene Expression*, Science, 342(6163): 1188–1193, pp. 1-14, doi:10.1126/science.1242975.

Santos R., Fuertes A., …, Cohen-Cory S., (2018), *DSCAM differentially modulates pre- and postsynaptic structural and functional central connectivity during visual system wiring*, Neural Development, 13:22, pp. 1-19.

Schmucker D., Clemens C., …, Zipursky L., (2000) *Drosophila Dscam Is an Axon Guidance Receptor Exhibiting Extraordinary Molecular Diversity*, Cell, Vol. 101, pp. 671–684.

Schmucker D., (2007), *Molecular diversity of Dscam: recognition of molecular identity in neuronal wiring*, Nature Reviews, Neuroscience, Vol. 8, pp. 915-920.

Shannon C., (1948), *A mathematical theory of communication*, The Bell system technical journal, 27(3), 379-423.

Sober E.,

Sober E., (1983), *Equilibrium explanation*, Philosophical Studies, 43(2), 201-210.

Id., (1999), *The Multiple Realizability Argument Against Reductionism*, Philosophy of Science, 66, pp. 542 – 564.

Sprinzak *et al* (2010), *Cis Interactions between Notch and Delta Generate Mutually Exclusive Signals States*, Nature, 465(7294), pp. 86-90.

Swanton C. and Beck S., (2014), *Epigenetic noise fuels cancer evolution*, Cancer cell, 26(6), pp. 775-776.

Thoday J., (1956), *Balance, Heterozygosity, and developmental Stability*, Cold Spring Harbor Symposia on Quantitative Biology 21, pp. 318-326.

Turner T., Schnell S., and Burrage K., (2004), *Stochastic approaches for modeling in vivo reactions*, Computational Biology and Chemistry, 28, pp. 165-178.

Umerez J., Mossio M, (2016), *Constraint*. In Dubitzky W. *et al* Encyclopedia of Systems Biology, Springer, pp. 490-493, 2013, 978-1-4419-9863-7.

Various (2017), *What Is Your Conceptual Definition of ''Cell Type' in the Context of a Mature Organism?*, Cell Systems, pp. 255-259.

Velasco D., (2012), *Objective and subjective probability in gene expression*, Progress in Biophysics and Molecular Biology 110, pp. 5-10.

Viñuelas J., (2012), *Towards experimental manipulation of stochasticity in gene expression*, Progress in Biophysics and Molecular Biology 110, pp. 44-53.

Werndl C., (2009). *Are deterministic descriptions and indeterministic descriptions observationally equivalent?*. Studies in history and philosophy of science part B: studies in history and philosophy of modern physics, 40(3), pp. 232-242.

Id., (2012), *Probability, Indeterminism and Biological Processes*, in D. Dieks D. *et al* (eds), Probabilities, Laws and Structures. The Philosophy of Science in a European Perspective, Volume 3, Springer.

Id., (2016) *Determinism and indeterminism*. In: The oxford handbook of philosophy of science. Oxford University Press, Oxford.

Woodward J., (2003), *Scientific Explanation*, Stanford Encyclopedia of Philosophy.

Id., (2005), *Making things happen: A theory of causal explanation*, Oxford university press.

Id., (2010), *Causation in Biology: Stability, Specificity, and the Choice of Level of Explanation*, Biology and Philosophy, 25, pp. 287-318.

Id., (2013), *II—Mechanistic explanation: Its scope and limits*. In Aristotelian Society Supplementary, Vol. 87, No. 1. Oxford, UK: Blackwell Publishing Ltd, pp. 39-65.

Wolfe C., (2012), *Chance between holism and reductionism: Tensions in the conceptualization of Life*, Progress in Biophysics and Molecular Biology, 110, pp. 113-120.
Wright S., (1952), *The Genetics of Quantitative Variability,* In ECR Reeve & CH Waddington (eds.), Quantitative Inheritance, London, Stationery Office, pp. 5-41.

Epigenetics

Baedke J., (2013), *The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences*, Studies in History and Philosophy of Biological and Biomedical Sciences 44, pp. 756–773.

Id., (2018), *Above the Gene, Beyond Biology: Toward a Philosophy of Epigenetics*, University of Pittsburgh Press.

Gilbert S. (1991), *Epigenetic Landscaping: Waddington's Use of Cell Fate Bifurcation Diagrams*, Biology and Philosophy, 6, pp. 135-154.

Morange M., (2002), *The Relations between Genetics and Epigenetics: A Historical Point of View*, Annals of the New York Academy of Sciences, 981, pp. 50-60.

Id., (2009), *What history tells us XVII. Conrad Waddington and The nature of life*, Journal of Biosciences, 34(2), pp. 195–198

Pujadas E. and Feinberg A., (2012), *Regulated Noise in the Epigenetic Landscape of Development and Disease*, Cell 148, Elsevier Inc., pp. 1123-1131.

Van de Vijver G., (2009), *No Genetics without Epigenetics? No Biology without Systems Biology? On the Meaning of a Relational Viewpoint in a Complex Account of Living Systems*, Natural Genetic Engineering and Natural Genome Editing: Annals of the New York Academy of Sciences. 1178, pp. 305–317.

Verhoeven K. *et al* (2016), *Epigenetics in ecology and evolution: what we know and what we need to know*, Molecular Ecology, 25, pp. 1631–1638.

Waddington C., (1957), *The Strategy of Genes*, Routledge Library Editions: 20th Century Science.

Contingency

Desjardins, E. (2011), *Reflections on path dependence and irreversibility: lessons from evolutionary biology*, Philosophy of Science, 78(5), 724-738.

Gould S., (1989), *Wonderful Life: The Burgess Shale and the Nature of History*, New York: W. W. Norton & Company.

Huneman P., (2017), *Macroevolution and Microevolution: Issues of Time Scale in Evolutionary Biology*, (con Bouton C.) (dir.), *Time of Nature and the Nature of Time. Dordrecht: Springer*, pp. 315 – 357.

Lewis C., (2018), *The domain relativity of evolutionary contingency*, Biology & Philosophy, 33(3), pp. 1-18.

Losos L., (2017), *Improbable Destinies: Fate, Chance, and the Future of Evolution*, Riverhead Books.

Turner D., (2011a), *Paleontology: A Philosophical Introduction*, Cambridge Introductions to Philosophy and Biology, Cambridge University Press.

Id., (2011b), *Gould's replay revisited*, Biology & Philosophy 26.1, pp. 65-79.

Wong T. (2019), *A Study of Evolutionary Contingency. Characterization, mechanisms, and evidence*, PhD dissertation, Fitzwilliam College University of Cambridge.

Id., (2020), *Sources of evolutionary contingency: chance variation and genetic drift*, Biology & Philosophy, pp. 35, 36.

Indeterminism/determinism, natural selection and genetic drift: debates on chance at the natural population levels

Beatty J., (1984), *Chance and Natural Selection*, Philosophy of Science, 51, p. 183-211.

Id., (1992) *Random Drift*, in E. F. Keller and E. A. Lloyd (eds.), Keywords in Evolutionary Biology, Harvard University Press, Cambridge, MA, pp. 273–281.

Id., (2006a), *Chance Variation: Darwin on Orchids*, Philosophy of Science, 73, pp. 629-641.

Id., (2006b), *Replaying life's tape*, The Journal of philosophy, 103(7), pp. 336–362.

Brandon R.N., (2006), *The principle of drift: Biology's first law*, The Journal of Philosophy, 103(7), pp. 319-335.

Brandon R. N., Carson S., (1996), *The Indeterministic Character of Evolutionary Theory: No "No Hidden Variables Proof" but No Room for Determinism Either*, Philosophy of Science, Vol. 63, No. 3, pp. 315-337.

Brandon R., and McShea D., (2020), *The Missing Two-Thirds of Evolutionary Theory*, Cambridge University Press.

Bravi B., Longo G., (2015), *The Unconventionality of Nature: Biology, from Noise to Functional Randomness*, Conference Paper, ResearchGate, DOI:10.1007/978-3-319-21819-9_1.

Buiatti M., Longo g., (2013), *Randomness and Multi-level Interaction in Biology*, in Theory of Biosciences, vol. 132, n.3, pp. 139-158.

Calude C., and Longo G., (2016). *Classical, quantum and biological randomness as relative unpredictability*. Natural Computing, 15(2), pp. 263-278.

Darwin C., ([1859] 1964), *On the Origin of Species: A Facsimile of the First Edition*, Harvard University Press, Cambridge, Massachusetts.

Dodson E., and Dodson P., (1985), *Evolution: Process and product*, Weber and Schmit, Boston.

Eble, G., (1999), *On the dual nature of chance in evolutionary biology and paleobiology*, Paleobiology, 25, pp. 75-87.

Gessell B., (2017), *Indeterminism in the brain*, Biology & Philosophy, *32*(6), pp. 1205-1223.

Graves L., Horan B. L., Rosenberg A., (1999), *Is Indeterminism the Source of the Statistical Character of Evolutionary Theory?*, Philosophy of Science, Vol. 66, No. 1, pp. 140-157.

Hodge J. S., (2016), *Chance and Chances in Darwin's Early Theorizing and in Darwinian Theory Today*, in Ramsey and Pence (ed.), Chance in Evolution, The University Chicago Press, Chicago & London, pp. 41–75.

Horan B. L., (1994), *The Statistical Character of Evolutionary Theory*, Philosophy of Science, 61, pp. 76-95.

Johnson C., (2015), *Darwin's Dice*, Oxford University Press.

Longo G., (2018), *How future depends on past and rare events in systems of life*., Foundations of Science 23.3, pp. 443-474.

Longo G., Montévil M., (2012a), *From physics to biology by extending criticality and symmetry breakings*, in Systems Biology and Cancer, Progress in Biophysics and Molecular Biology, vol.106 (2), pp. 340–347.

Id., (2012b), *Randomness Increases Order in Biological Evolution*, in: Dinneen M., Khoussainov B., Nies A., (eds.) Computation, Physics and Beyond, Lecture Notes in Computer Science, vol. 7160, pp. 289 - 308. Springer Berlin / Heidelberg.

McShea D., and Brandon R., (2010), *Biology's first law: the tendency for diversity and complexity to increase in evolutionary systems*, University of Chicago Press.

Merlin F., (2015), *Monod's concept of chance: its diversity and relevance today*, Comptes Rendus de Biologie de l'Académie des Sciences, 338, pp. 406-412.

Merlin F., (2016b), *The Multi-faceted Idea of Chance in Darwin's Writings*, Springer.

Merlin F., Malaterre C., (2015), *The (In)Determinism of Biological Evolution: Where Does the Stochastic Character of Evolutionary Theory Come From?*, in Heams T., Huneman P.,

Lecointre G., Silberstein M. (ed.), *Handbook of Evolutionary Thinking in the Science*, Springer Science+Business, Media Dordrecht, pp. 349-366.

Millstein, R. L. (1996), *Random Drift and the Omniscient Viewpoint*, Philosophy of Science, 63, pp. S10-S18.

Id., (1997). *The Chances of Evolution: An Analysis of the Roles of Chance in Microevolution and Macroevolution*, Ph.D. Dissertation. University of Minnesota.

Id., (2000a), *Interpretations of Probability in Evolutionary Theory*, Philosophy of Science, 70, pp. 1317–1328.

Id., (2000b), *Is Evolutionary Process Deterministic or Indeterministic? An Argument for Agnosticism*, presented at the Biennial Meeting of the Philosophy of Science Association Vancouver, Canada, (http://philsci-archive.pitt.edu/4544/1/Millstein2000Agnosticism.pdf).

Id., (2000c), *Chance and Macroevolution*, Philosophy of Science, 67, pp. 603-624.

Id., (2002), *Are Random Drift and Natural Selection Conceptually Distinct?*, Biology and Philosophy,17, pp. 33-53.

Id., (2003), *How Not to Argue for the Indeterminism of Evolution: A Look at Two Recent Attempts to Settle the Issue*, In Determinism in Physics and Biology, edited by Hüttemann A., Paderborn, Mentis.

Id., (2006), *Discussion of "Four Case Studies on Chance in Evolution" Philosophical Themes and Questions*, Philosophy of Science, 73, pp. 678-687.

Id., (2011), *Chances and Causes in Evolutionary Biology: How Many Chances Become One Chance*, In Causality in the Sciences, eds. McKay Illari P., Russo F., and Williamson J., pp. 425-444, Oxford, Oxford University Press.

Id., (2016) *Genetic drift*. In: Zalta EN (ed) The stanford encyclopedia of philosophy. https://plato.stanford.edu/archives/fall2 017/entries/genetic-drift/

Mills S.K., Beatty J.H., (1979), *The Propensity Interpretation of Fitness*, Philosophy of Science, 46/2, pp. 263-86.

Monod J., (1971), *Il caso e la necessità*, Milano, Mondadori.

Montévil M., Mossio M., Pocheville A., Longo G., (2016), *Theoretical principles for biology: Variation*, Progress in Biophysics and Molecular Biology, Available online, ISSN 0079-6107, http://dx.doi.org/10.1016/j.pbiomolbio.2016.08.005.

Mossio M., Montévil, M., and Longo G., (2016), *Theoretical principles for biology: Organization*, Progress in Biophysics and Molecular Biology, 122(1), pp. 24-35.

Pavé A. (2007), *Necessity of chance: biological roulettes and biodiversity*, Comptes rendus biologies, 330, pp. 189–198.

Rosenberg, A., (1994), *Instrumental Biology or the Disunity of Science*. Chicago: University of Chicago Press.

Id., (2006), *Darwinian reductionism. Or, how to stop worrying and love molecular biology*, Cambridge: University of Chicago Press.

Rosenberg A., Bouchard F., (2015), *Fitness, The Stanford Encyclopedia of Philosophy*, (https://plato.stanford.edu/entries/fitness/).

Ross W., (1936), *Aristotle's Physics*, Publisher: Oxford University Press

Roughgarden J., (1979), *Theory of Population Genetics and Evolutionary Theory: An Introduction*. New York: Macmillan

Shanahan T., (1989), *Beatty on Chance and Natural Selection*, Philosophy Faculty Works, Paper 10, (http://digitalcommons.lmu.edu/phil_fac/10).

Id., (1991), *Chance as an Explanatory Factor in Evolutionary Biology*, History and Philosophy of the Life Sciences, 13, pp. 249-269.

Id., (1992), *Selection, Drift, and the Aims of Evolutionary Theory*, in P. Griffiths (ed.), Trees of Life: Essays in Philosophy of Biology, Kluwer, Dordrecht, pp. 131–161.

Id., (2003), *The evolutionary indeterminism thesis*, Bioscience, 53, pp. 163–169.

Soto A., Longo G., Miquel P., Montévil M., Mossio M., Perret N., Pocheville A., Sonnenschein C., (2016), *Toward a theory of organisms: Three founding principles in search of a useful integration*, in Progress in Biophysics and Molecular Biology.

Weber M., (2000), *Determinism, Realism, and Probability in Evolutionary*, Philosophy of Science, Vol. 68, No. 3, pp. S213-S224.

Id., (2005*), Indeterminism in Neurobiology*, Philosophy of Science, Vol. 72, pp. 663-674.

Id., (2008), *Causes without mechanisms: Experimental regularities, physical laws, and neuroscientific explanation*, Philosophy of Science, 75(5), pp. 995-1007.

Natural kinds

Campbell K., O'Rourke M., and Slater M., (Eds.). (2011). *Carving nature at its joints: Natural kinds in metaphysics and science*. MIT Press.

Sabbarton-Leary N., Bortolotti L., Broome M., (2015), *Natural and Pare-Natural Kinds in Psychiatry*. In Zachar P. (ed.), Stoyanov D. (ed.), Aragona M. (ed.), Jablensky A. (ed.), Alternative Perspectives on Psychiatric Validation DSM, ICD, RDoC, and Beyond, Oxford, UK: Oxford University Press, pp. 76 – 93.

Sorensen R., (2011), *Para-natural Kinds*, In J. K. Campbell, M. O'Rourke, and M. H. Slater

(eds), *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science*, Cambridge, MA: MIT Press.

Handbooks

Alberts B., Johnson A., Lewis J., Raff M., Roberts, K., and Walter P., (2008), *Molecular biology*
of the cell (5th ed.), New York: Garland Science.

Chakravartty A., (2013), *On the prospects of Naturalized Metaphysics*, in Ross D., Ladyman J., and Kincaid H., (Eds.), *Scientific metaphysics*. Oxford University Press.

Id., (2017), *Scientific ontology: Integrating naturalized metaphysics and voluntarist epistemology*. Oxford University Press.

Bachelard G., (1934), *La formation de l'esprit scientifique*, (4 e édition), Paris: Vrin.

Evert R. and Elchhorn S., (2013), *Raven. Biology of Plants*, W. H. Freeman and Company.

Galton F., (1894). *Natural inheritance*, Macmillan and Company.

Godfrey-Smith P., (2003), *Theory and Reality*, The University of Chicago Press / Chicago and London.

Griffiths P., (2007), *Philosophy of biology*, in Sarkar S. and Pfeifer J. (Eds.), The philosophy of Science, An encyclopedia, New York: Routledge, pp. 68–75).

Hoquet T. and Merlin F., (2014), *Précis de philosophie de la biologie*, Vuibert.

Thom, R. (1992). *Prédire n'est pas expliquer*. Revue Philosophique de la France Et de l, 182(2).

Livingstone Smith D., (2017), *How biology Shapes Philosophy*, Cambridge University Press.

Mak T. and Saunders M., (2008), *Primer to the Immune Response*, first ed. Elsevier.

Nagel E., (1961), *The Structure of Science. Problems in the logic of scientific explanation.* London: Routledge.

Sterelny K. and Griffiths P., (1999), *Sex and Death. An Introduction to Philosophy of Biology*, University Chicago Press.

Résumé conséquent de la thèse en français

Introduction

Cette thèse en philosophie de la biologie porte sur le rôle du hasard dans les explications moléculaires et cellulaires, avec quelques excursus provenant également de la biologie du développement. Dans ce qui suit, j'introduirai certains des éléments nécessaires à la compréhension du type de travail dans lequel le lecteur s'aventurera : le contexte philosophique et biologique de ma recherche, le sujet principal de la thèse, le problème spécifique que j'aborde, la nature des questions et des réponses que je propose, ainsi que les particularités et les limites de mon analyse. Dans la première partie de l'introduction, je ne donne pas de définition précise des notions que j'utiliserai plus fréquemment au cours de mon travail, c'est-à-dire le hasard, la stochasticité et l'aléatoire. La raison en est de garder l'esprit du lecteur aussi libre que possible de toute définition spécifique pendant que je présente des problèmes et des scénarios qui peuvent être compris comme des prolégomènes à mon travail, c'est-à-dire avant de poser la question portant directement sur le hasard, la stochasticité et le caractère aléatoire. Ce n'est que dans la section 1.6 que je commencerai à circonscrire les significations que ces notions auront pour le développement de ma thèse. Pour l'instant, je peux dire que le hasard peut intuitivement être conçu comme un concept polysémique. Millstein (2010) démontre très bien ce point, en soulignant à travers quatre exemples comment le hasard a déjà une sémantique riche dans le langage courant :

« Je donnerais à ce cheval une chance sur deux (i.e. degré de croyance) de gagner.
J'espère avoir la chance (i.e. l'opportunité) de vous voir.
J'ai trouvé ce super restaurant par hasard (i.e. accident).
Je suis désolé, mais je ne peux pas prendre le risque » (p. 426).

De même, nous verrons que même en biologie, les significations attribuées au « hasard » sont très disparates. Cependant, son pluralisme sémantique ne sera pas le sujet principal de mon travail. Le cœur de mon travail sera plutôt le hasard comme élément de nos descriptions et de nos explications du monde biologique. En d'autres termes, il sera surtout question du hasard comme propriété épistémologique plutôt que comme caractéristique intrinsèque des systèmes biologiques existants.

1. Premier tournant : de l'évolution aux biomolécules et aux cellules

Depuis que Darwin a formulé pour la première fois la théorie de l'évolution par sélection naturelle ([1859] 1964), le hasard en biologie est considéré comme intimement lié à l'origine de la variation héritable. Dans la lignée de la pensée de Darwin, la Synthèse moderne, développée dans les années 1940 et 1950, a invoqué le hasard pour caractériser la variation génétique via les mutations aléatoires, ainsi que les processus évolutifs de la dérive génétique. La philosophie de la biologie, en tant que domaine de recherche autonome au sein de la philosophie des sciences, est apparue environ dix ans plus tard, c'est-à-dire dans les années 1960 et 1970. Lorsqu'elle a commencé à s'intéresser au hasard en biologie, vers les années 1980 et 1990, elle a hérité de cet intérêt pour la réflexion sur l'évolution. En effet, les chercheurs se sont concentrés sur les questions philosophiques soulevées par le hasard en biologie évolutive et, plus spécifiquement, en théorie de l'évolution (TE). Dès lors, un ensemble de réflexions philosophiques et historiques sur l'étude du hasard dans l'évolution en a découlé. Par exemple, dans les années 1990 et 2000, un débat clé a éclaté entre des philosophes se demandant si la TE est une théorie probabiliste (Horan 1994 ; Millstein 2000a). Peu après, le débat s'est déplacé vers des questions métaphysiques plus spécifiques : si le processus évolutif lui-même est déterministe ou indéterministe (Brandon et Carson 1996 ; Graves, Horan, Rosenberg, 1999), et si la sélection naturelle et la dérive génétique aléatoire peuvent être distinguées conceptuellement et/ou empiriquement (Beatty 1984 ; Millstein 2002). A partir de là, de nombreuses autres réflexions sur le hasard ont été proposées à travers l'analyse de concepts clés importants en biologie évolutive. Il s'agit notamment de la mutation aléatoire (Millstein 1997 ; Merlin 2010 ; 2016a), de la fitness (Drouet et Merlin 2015 ; Rosenberg et Bouchard 2015), de la probabilité et de la propension (Cerredo 2020 ; DesAutels 2014), du hasard évolutif (Eble 1999 ; Shanahan 1989, 1991, 1992, 2003) et de la contingence (Beatty 1984, 1992, 2006a, 2006b ; Cerredo 2020 ; Gould 1989 ; Losos 2017 ; Wong 2019).

Même si cette liste n'est pas exhaustive, elle est néanmoins représentative en montrant que les débats philosophiques sur le hasard en biologie se sont principalement concentrés sur la théorie de l'évolution, et plus généralement sur le cadre de la biologie évolutive. Mais qu'en est-il du hasard sous l'évolution ? Voyons ce que j'entends par « sous ». Ce travail se concentre sur des domaines dans lesquels le hasard n'est pas nécessairement réductible à ses significations évolutionnistes, ou du moins, sa signification est plus riche que ses définitions d'un point de vue évolutionniste (par exemple, en tant que « source de variation non dirigée » ou en tant que « processus d'échantillonnage indiscriminé », cf. Millstein 2011). Tout d'abord, par « sous », je ne fais pas référence au rejet de l'évolution, mais à un mode d'enquête qui examine ce que

le hasard pourrait signifier à une échelle spatio-temporelle plus fine. La biologie évolutive se concentre généralement, mais pas exclusivement, sur le niveau des populations naturelles en évolution. Ce projet vise à explorer le niveau physiologique individuel où se déroulent les processus moléculaires et cellulaires, qui se trouvent sous le principal niveau d'étude de l'évolution. Deuxièmement, « sous », dans mon cadre, indique aussi métaphoriquement la volonté de faire une étude qui encadre le hasard avec des outils épistémologiques autres que ceux de l'évolution classique. Regarder « sous » l'évolution signifie analyser le hasard d'un point de vue épistémologique qui, même en partant du cadre évolutionniste, essaie de proposer quelque chose de nouveau.

Je ne peux qu'être d'accord sur l'importance de poursuivre les travaux philosophiques qui étudient et explorent les différentes notions de hasard dans l'évolution. Cependant, je pense également que l'étude du hasard ne doit pas se limiter à son analyse strictement évolutionniste car cela pourrait conduire à une sous-estimation de la richesse des questions philosophiques, pas toujours directement liées à l'évolution, que le hasard peut soulever dans d'autres domaines biologiques. Sarkar et Plutynski ont déjà dénoncé cette tendance de la philosophie de la biologie à sous-estimer les champs de recherche biologique qui ne sont pas directement concernés par l'évolution :

« Traditionnellement, l'évolution a été le centre de la plupart des attentions philosophiques. Alors qu'il reste sûrement vrai que 'rien en biologie n'a de sens si ce n'est à la lumière de l'évolution' (Dobzhansky 1973), cette tradition au sein de la philosophie de la biologie est myope dans la mesure où elle ignore une grande partie – sinon la plupart – des travaux en biologie contemporaine » Sarkar et Plutynski (2008, p. xviii).

Le fait que les philosophes contemporains ne trouvent généralement pas l'aspect moléculaire intéressant a également été soutenu par Pradeu (2017). Il a analysé tous les articles publiés par le journal *Biology and Philosophy* de 2003 à 2015 et tous les articles scientifiques publiés par les *Proceedings of the National Academy of Science of the USA* (PNAS) pendant la même période. En comparant ces données, il a constaté un décalage important entre les domaines biologiques explorés par les philosophes de la biologie publiant dans *Biology and Philosophy*, et ceux couverts par les biologistes publiant dans *PNAS* (Pradeu 2017, pp. 150-151). Plus précisément, il a constaté que 62 % des articles philosophiques sont consacrés à l'évolution (i.e. *Biology and Philosophy*) mais seulement 5 % des articles biologiques (i.e. revue *PNAS*). Il a

également constaté que la biologie moléculaire et cellulaire est aux mieux mentionnés de façon marginale, et au pire complètement absente, dans la littérature de la philosophie de la biologie.

Ces données pourraient suggérer que les domaines biologiques, autres que ceux de l'évolution, pourraient en fait être moins intéressants d'un point de vue philosophique. Il y a moins d'ouvrages parce que les sujets autres que ceux de l'évolution ne sont pas vraiment intéressants d'un point de vue philosophique. Mais ce raisonnement ne tient pas si l'on se concentre sur des travaux récents qui montrent exactement l'inverse, par exemple la réflexion philosophique de Merlin (2016a) sur la stochasticité des mutations génétiques au niveau moléculaire, les travaux de Baedke (2018) sur l'épigénétique, la problématisation du concept de soi en immunologie par Pradeu (2011), l'analyse des ARN non codants par Théry (2016), la refonte des propriétés des cellules souches par Laplane (2016), etc. Si cette liste est loin d'être complète, elle permet tout de même de montrer que les sujets non évolutifs sont un terrain fertile pour la réflexion philosophique.

Partant, nous pouvons maintenant nous demander s'il ne serait pas intéressant d'élaborer une analyse philosophique sur le hasard sous la biologie évolutive. En regardant autour de nous, il est clair que ce type d'analyse est très rare dans les réflexions philosophiques, et laisse inexplorée une importante lacune conceptuelle que les biologistes eux-mêmes soulignent souvent. Par exemple, Heams (2012) parle d'une « biologie schizophrénique » (p. 9). Si, en biologie évolutive, le hasard est généralement conçu comme une source de nouvelles variations, en biologie moléculaire, il est en revanche souvent conçu comme une « erreur » ou un « bruit », à savoir une source de nuisance pour les systèmes (biologiques). Mais cette situation paradoxale ne concerne pas exclusivement la comparaison entre la biologie évolutive et la biologie moléculaire. En immunologie, par exemple, le hasard joue un rôle central pour expliquer la synthèse des anticorps. Mais les entités dont traite ce domaine sont des biomolécules, les mêmes entités que celles abordées par la biologie moléculaire ! Plus généralement, comment est-il possible que le hasard soit conçu si différemment, avec des connotations contradictoires (négatives et positives), dans les différents domaines de la recherche biologique ? Comment expliquer cette hétérogénéité conceptuelle ?

Dans ce contexte, il pourrait être utile de distinguer deux points qui pourraient être poursuivis en parallèle, à savoir 1) la signification du hasard et 2) son rôle :

1. Il existe peu d'analyses philosophiques de la signification du hasard en biologie moléculaire et cellulaire. Que peut-on dire à ce sujet ? Peut-on le réduire uniquement aux

significations évolutionnistes que Millstein (2011) nous a proposées ? Ou, au contraire, le hasard a-t-il des significations spécifiques en biologie cellulaire et moléculaire ?

2. A l'exception de l'immunologie, le hasard en biologie moléculaire et cellulaire est souvent conçu comme une erreur ou un bruit et donc ne fait pas partie de l'explication des systèmes biologiques (l'inverse est vrai en biologie évolutive). Cela soulève la question suivante : le hasard peut-il avoir un rôle explicatif en biologie moléculaire et cellulaire ? Et si oui, dans quelle mesure et de quelle manière le hasard peut-il être explicatif ?

Ces deux points résument bien l'objectif de l'ensemble de la thèse. Les chapitres 1, 2, 3 et 4 s'attacheront principalement à répondre à la deuxième question, concernant le rôle explicatif du hasard, ce qui correspond au cœur de ma recherche. La première question concernant la signification du hasard fera l'objet du chapitre 6. Dans celui-ci, je proposerai une définition du hasard à travers une étude de cas spécifique. Les chapitres 5 et 7 répondront aux deux questions.

1.2 Deuxième tournant : des théories à l'explication

Une importante littérature philosophique sur le hasard entre les années 1990 et 2000 s'est attachée à répondre aux raisons pour lesquelles la théorie de l'évolution est probabiliste. Deux positions principales ont été discutées : la TE est probabiliste puisque le processus évolutif est intrinsèquement indéterministe ; dit autrement, puisque nous ignorons les détails du processus qui, s'ils étaient connus, nous permettraient de rendre compte de phénomènes déterministes, la TE est probabiliste. Afin de comprendre pourquoi ces réflexions philosophiques se concentrent sur la théorie de l'évolution, plutôt que sur les modèles et/ou les explications de l'évolution, prenons un peu de recul et examinons le contexte dans lequel la philosophie de la biologie a émergé au cours des années 1960-70 en tant que domaine de recherche autonome.

La philosophie des sciences s'est souvent intéressée à la physique, en particulier durant la première moitié du XXe siècle, lorsque la physique était considérée comme la science paradigmatique. Cet intérêt a culminé dans le développement de la philosophie de la physique qui réfléchit typiquement sur les théories puisqu'il y en a beaucoup dans ce domaine d'étude. La mécanique quantique et la relativité générale et restreinte sont deux piliers exemplaires de la réflexion philosophique sur les théories physiques.

En biologie, les théories sont moins nombreuses. En fait, il y en a trois (Hoquet et Merlin 2014) : la théorie de l'évolution (TE), qui est au centre de l'attention depuis que Darwin l'a proposée en 1859 ; la théorie cellulaire ; et le cadre théorique à partir duquel la biologie

moléculaire a été fondée (c'est-à-dire le dogme central, si on la considère comme une « théorie »). Nonobstant, Nicholson et Gawne (2015) ont montré que l'empirisme logique, un courant philosophique né dans la première moitié du XXe siècle au sein du Cercle de Vienne, était une tentative « de rendre les *théories* biologiques plus semblables à leurs homologues physiques en les intégrant de force dans des systèmes déductifs rigoureux » (Nicholson et Gawne 2015, p. 347 ; je souligne ; voir aussi Callebaut 1993). Plus précisément, faisant écho à Wolters (1999, p. 195), Nicholson et Gawne précisent que cette « manière forcée » de faire de la philosophie de la biologie que poursuivaient les empiristes logiques consistait à appliquer trois thèses qui sont en réalité malencontreuses lorsqu'elles sont appliquées à la biologie plutôt qu'à la physique : « l'antimétaphysique » (la pensée que toute métaphysique est à éviter dans la réflexion philosophique), la « réduction » (la croyance que la biologie est réductible à l'explication physico-chimique) et « la physique comme science modèle » (la conception que la physique est la science modèle que la biologie devrait essayer d'imiter) (Nicholson et Gawne 2015, p. 350). Cette dernière thèse est celle qui nous concerne le plus dans la présente section. Sans entrer dans le fascinant débat historique concernant les racines de la philosophie de la biologie (Nicholson et Gawne 2014, 2015 ; Pradeu 2018), ce que je veux souligner ici, c'est que 1) la tradition en philosophie des sciences a pu influencer la focalisation de la philosophie de la biologie sur la théorie – la troisième thèse empiriste logique en est un argument, et 2) la théorie de l'évolution est « LA » théorie en biologie et de nombreux travaux philosophiques se sont concentrés sur les différentes questions qu'elle soulève.

En ce sens, il y a eu ensuite, dans les années 1960, une rupture méthodologique proposée par Hull (1969) qui a été l'instigateur d'un abandon de cette façon de faire de la philosophie des sciences quand elle a trait à la biologie. Afin de développer une réflexion philosophique pertinente pour la science, il a suggéré de regarder ce que font les biologistes. Ce texte très cité est celui dans lequel Hull affirme qu'il y a beaucoup à faire d'un point de vue philosophique en ce qui concerne les questions faisant partie intégrante de la biologie – mais, au moment de la rédaction de ce texte, en 1969, rien n'avait encore été fait. Pradeu (2018) affirme avec enthousiasme que le désir de Hull a en fait été réalisé dans les décennies suivantes : « Les souhaits de Hull ont été exaucés, mieux qu'il ne pouvait l'espérer » (p. 454). Aujourd'hui, la philosophie de la biologie est une branche bien formée et riche de travaux « bottom-up », c'est-à-dire de travaux qui regardent de près la pratique scientifique pour en déduire des questions philosophiques intéressantes et potentiellement pertinentes pour la science (cf. chapitre 1). Mais ce qui m'intrigue à ce stade, c'est une autre citation moins célèbre qui s'insurge directement

contre l'obsession des empiristes logiques d'axiomatiser toutes les sciences, y compris la biologie :

« Trop souvent, les applications de la logique mathématique aux problèmes de la biologie donnent l'impression que des idées plus ou moins banales ont été exprimées avec une exactitude fastidieuse alors qu'elles auraient pu être transmises plus facilement et plus directement en quelques phrases d'un anglais simple » (Hull, 1969, p. 173).

Ici, Hull est clair : il critique la tentative des empiristes logiques de faire entrer la biologie dans des axiomes formels et mathématiques. Il s'agit d'une percée majeure qui a conduit la philosophie de la biologie à s'ouvrir à des méthodologies différentes, plus attentives à la pratique biologique et non canalisées par la perspective étroite de travailler uniquement sur des théories (cf. chapitre 1).

De ce que nous avons dit jusqu'ici, pourrions-nous affirmer que la pratique quotidienne de la biologie ne consiste pas à formuler/construire des théories, mais plutôt des modèles et des explications ? (cf. Waters 2014). Cette affirmation serait trop forte, mais on peut en dire quelque chose de moins direct. Si l'on prend n'importe quel manuel de biologie, on constate que ce qui compte le plus en biologie, c'est l'explication des processus : un manuel de biologie moléculaire vise à expliquer, par exemple, la duplication de l'ADN, le processus menant à l'expression des gènes, les modifications post-traductionnelles, etc. Un manuel de biologie cellulaire vise à expliquer, par exemple, les mitoses, la méiose, la différenciation cellulaire, etc. Dans le même ordre d'idées, l'un des principaux objectifs des articles sur la biologie est d'expliquer les dernières observations empiriques en biologie (même si, bien entendu, on trouve aussi des articles descriptifs). Des titres tels que « Explaining Gene Expression Using Twenty-One MicroRNAs » (Asiaee et al 2020) ou « Can Genes Explain Biological Complexity » (Szathmáry et al 2001) témoignent de cette tendance.

Mais une objection se présente : la biologie fournit plusieurs théories à la place ! Comme nous l'avons déjà mentionné, il en existe au moins trois : la théorie de l'évolution, la théorie cellulaire et le dogme central dans la biologie moléculaire. De plus, dire que la biologie se concentre sur les explications et les modèles et non sur les théories pourrait stimuler la curiosité du lecteur quant à ce que j'entends par ces trois catégories. Mais prendre le chemin de la définition d'un modèle et d'une théorie prendrait trop de temps, et n'est pas indispensable pour cette thèse (même si, par ailleurs, je passerai beaucoup de temps à définir l'explication stochastique). Ce qui m'intéresse, c'est seulement de préciser que, dans la mesure où l'on parle

de hasard en biologie moléculaire et cellulaire, il est *de facto* présent principalement dans les explications et les modèles, alors que par « explication » je fais simplement référence à toute sorte de réponse à une question du « pourquoi », et par « modèle », à toute sorte de schéma et/ou de représentation qualitative/quantitative qui vise à décrire et à expliquer les processus biologiques. En notant que le hasard est la plupart du temps mentionné dans les explications et les modèles, je me libère du poids de dire qu'en biologie en général il n'y a que des explications et des modèles. Dans le présent travail, je me concentre principalement sur les explications en posant l'hypothèse que souvent les modèles sont construits pour fournir des explications. Par exemple, nous verrons qu'un modèle d'expansion de la paroi cellulaire est proposé pour expliquer la forme sphérique de certaines cellules végétales. Mais afin d'éviter de toujours mentionner les explications et les modèles, je parlerai désormais principalement des explications. Les discussions sur les modèles, si elles sont pertinentes, suivront éventuellement.

Enfin et surtout, l'accent mis sur les explications (plutôt que sur les théories) semble particulièrement pertinent et fructueux par rapport au type de philosophie de la biologie qui caractérise ma recherche. Love (2008) écrit : « [la] relation entre la philosophie et la science peut être bénéfique, mais elle est intrinsèquement précaire car elle exige de maintenir une *tension* entre la proximité conceptuelle de la pratique scientifique et la distance interprétative des réflexions philosophiques » (pp. 66-67 ; je souligne). Pour au moins deux raisons, l'explication présente un sujet parfait pour trouver un bon équilibre à cette tension entre philosophie et biologie. La première est que l'explication me permet de travailler à proximité de la pratique scientifique, puisque l'explication est l'une des catégories les plus développées et utilisées dans la pratique biologique discutant du hasard. Cela peut sembler trivial mais ne l'est pas, car cela trace une position précise par rapport au type de philosophie que je propose dans le présent travail. Comme on le verra plus clairement au chapitre 1, un projet philosophique centré sur la science doit prendre au sérieux ce qui se passe en biologie et chercher à être aussi proche que possible de sa pratique. Cette position, qui est déjà codifiée dans l'univers de la philosophie des sciences, est appelée « philosophie des sciences en pratique » (PSP ; Ankeny 2011). Cette façon de faire de la philosophie est un positionnement clair contre une philosophie « top-down » qui est moins sensible à la biologie (ou à la science) contemporaine, et qui tire plutôt sa plus grande force d'arguments philosophiques *a priori* (cf. Sterelny et Griffiths 1999, p. xi ; Turner 2011a, p. 12). La deuxième raison est que, en me concentrant sur l'explication, je dispose d'un très large fond de référence philosophique. Depuis plus de 50 ans, l'explication scientifique a fait l'objet de critiques et de discussions philosophiques minutieuses : du premier modèle de Hempel et Oppenheim (1948) jusqu'aux

toutes dernières propositions de Strevens (2008) et Woodward (2003), l'explication scientifique a été soigneusement analysée avec un ensemble hétérogène de comptes rendus philosophiques. Ce que j'essaie de suggérer est que le thème de l'explication dans l'étude du hasard pourrait être considéré comme un excellent terrain de rencontre entre la biologie (qui travaille tout le temps avec des explications) et la philosophie (qui étudie les explications depuis plus de 50 ans).

Après avoir précisé que mon sujet est le hasard en biologie moléculaire et cellulaire dans une perspective non évolutive, et que son cadre est l'explication scientifique, je voudrais brosser un tableau plus complet du type de travail épistémologique que je vais mener, à partir de ces deux éléments. La section suivante se concentre justement sur ce point.

1.3 La nature épistémologique de mon travail

Dans les travaux cellulaires et moléculaires, le hasard est souvent conçu comme une erreur ou un bruit, et reste la plupart du temps à l'écart, considéré comme un obstacle épistémologique, c'est-à-dire un problème pour l'avancement de la recherche. Plus précisément, la question très générale souvent abordée dans la littérature biologique est la suivante : quelle est la signification et le rôle du hasard dans les phénomènes biologiques ? Les travaux qui en découlent oscillent comme un pendule entre limitation cognitive et réductionnisme ontologique. Dans la littérature biologique, nous pouvons trouver deux significations principales à l'étiquette « hasard » lorsqu'elle est utilisée pour caractériser un phénomène biologique. La première fait référence au fait qu'il est tout simplement trop difficile de déballer toutes les dynamiques spécifiques impliquées dans le phénomène examiné. En d'autres termes, lorsqu'il devient problématique de comprendre ce qui se passe dans un phénomène ou comment il produit certains résultats, on dit souvent que « ce processus est le fruit du hasard ». Dans cette signification, le hasard fait référence à notre limitation cognitive de la connaissance, de l'explication et de la compréhension de certains phénomènes. D'autre part, en biologie cellulaire et moléculaire, le hasard est souvent décrit comme un processus physique (par exemple, l'agitation thermique des molécules) qui influence les comportements des phénomènes biologiques. Cette deuxième signification du hasard est engagée dans une sorte de réductionnisme ontologique, selon lequel le hasard au niveau biologique n'est rien d'autre que l'agitation thermique des molécules au niveau physique. Je voudrais échapper à ce mouvement de balancier. Dans le présent travail, je proposerai une idée du hasard comme élément essentiel de l'explication biologique plutôt que de l'utiliser pour souligner nos limites

cognitives ou pour le réduire aux phénomènes physiques sous-jacents aux processus biologiques (cf. chapitres 5 et 6).

Le défi philosophique de ce travail est de dépasser cette conception du hasard comme « ennemi » de la connaissance en biologie cellulaire et moléculaire. Je ne suis pas le premier à essayer de rendre cette sorte de « dignité » au hasard. Dans l'interface entre la physique et la biologie, entre autres, Bravi et Longo (2015) et Buiatti et Longo (2013) tentent de redonner un sens au hasard en termes de fonctionnalité qui peut s'étendre à tous les niveaux de l'organisation biologique. D'un point de vue purement biologique, Kupiec (2009a, 2009b) propose d'étendre le darwinisme jusqu'à l'intérieur des cellules, en conférant au hasard un rôle central dans la fonctionnalité de la plupart des phénomènes biologiques. L'un des livres qu'il a édités s'intitule même *Le hasard au cœur de la cellule*. Tous ces travaux représentent des contributions importantes au débat et j'aurai l'occasion de m'engager sur certains d'entre eux au chapitre 4, où j'examinerai certains comptes déjà proposés et discutés dans la littérature philosophique et biologique.

Mais ce que je voudrais préciser à ce point de l'introduction, c'est que ma contribution est plus modeste et circonscrite par rapport aux travaux qui viennent d'être mentionnés. Je n'ai pas l'intention de poursuivre une argumentation à la fois ontologique et épistémologique sur le hasard en biologie moléculaire et cellulaire (cf. section 1.4). Ma contribution se concentre sur les explications biologiques dans ces domaines disciplinaires, et donc sur le pan épistémologique, 1) en argumentant que le hasard peut avoir un rôle central dans l'explication de certains processus moléculaires et cellulaires au moins, et 2) en montrant pourquoi le hasard peut avoir un tel rôle fondamental. Même si le premier point est déjà abordé par des chercheurs tels que ceux mentionnés ci-dessus, ils oscillent néanmoins souvent entre des arguments ontologiques et épistémologiques, à savoir entre le rôle fonctionnel et explicatif du hasard (cf. section 1). En revanche, mon travail vise à rester principalement à un niveau épistémologique, en abordant la question du rôle du hasard dans l'explication, c'est-à-dire son pouvoir explicatif (voir après). Le second point, qui concerne les raisons pour lesquelles le hasard peut avoir un rôle fondamental dans l'explication de certains processus biologiques (en particulier au niveau moléculaire et cellulaire), est rarement abordé. Pour ce faire, il faut être ouvert à une sorte d'exploration philosophique qui cherche à aller plus loin en remettant en cause les réponses qui semblent évidentes *de prime* abord. En effet, on pourrait rendre cette deuxième question triviale en répondant que le hasard est essentiel parce que l'explication dans laquelle il est présent fonctionne, tout simplement. Bien que je sois d'accord avec cette réponse car elle présuppose que le hasard a un rôle dans certaines explications biologiques (ce qui est néanmoins une partie

importante de ce travail), la réponse n'est pas philosophiquement satisfaisante. En effet, la question demeure : pourquoi le hasard est-il essentiel dans certaines explications biologiques ? Il existe ici un terrain philosophique plus subtil qui n'a pas encore été pleinement discuté. Que possède le hasard pour le rendre si efficace dans l'explication ? En d'autres termes, d'où vient le pouvoir explicatif du hasard ?

L'exercice philosophique que je propose sur cette question me rappelle le jeu classique du « Pourquoi ? » qui a souvent lieu entre parents et enfants. « Pourquoi l'arbre est-il vert ? » « Parce qu'il y a des feuilles. » « Pourquoi les feuilles sont vertes ? » « Parce qu'il y a de la chlorophylle. » « Pourquoi la chlorophylle est-elle verte ? » « Parce que c'est une molécule qui reflète la lumière verte et absorbe la lumière rouge et bleue. » « Pourquoi les molécules reflètent-elles la lumière verte et absorbent-elles plutôt la lumière rouge et bleue ? » « Parce que les électrons à l'intérieur des molécules sont excités et passent à des niveaux d'énergie plus élevés », et ainsi de suite. En développant ce projet, je me suis senti comme l'enfant qui ne cesse de demander pourquoi. Je pense que c'est une bonne stratégie pour explorer les implications philosophiques cachées dans les manières dont les biologistes traitent le monde naturel.

Une autre précision que je dois donner concernant la nature épistémologique de mon travail est que je ne propose pas une nouvelle assertion de l'explication scientifique à ajouter à la liste déjà longue proposée dans la littérature de philosophie des sciences (voir Braillard et Malaterre 2015 ; Brigandt 2013). Ce que je veux proposer est, d'une part, une réflexion qui vise à rendre le lecteur capable de reconnaître la valeur du hasard dans certaines explications et de réévaluer le jugement épistémologique à son sujet. D'autre part, je voudrais suggérer que, au moins dans certaines circonstances, il pourrait être utile de recadrer certaines explications et d'en proposer de nouvelles, en utilisant la notion de hasard sans crainte.

Dans cette section, j'ai été explicite quant au type de travail que je poursuivrai, à savoir une enquête épistémologique qui centre son attention sur le rôle du hasard dans les explications scientifiques. Bien que j'aie précisé que je ne poursuivrai pas de travaux sur l'ontologie du hasard, pourquoi ai-je pris cette décision ? Pourquoi est-ce que je ne me concentre pas aussi sur l'ontologie ? Dans la section suivante, je justifie pourquoi, dans mon travail, je veux garder les dimensions épistémologiques et ontologiques aussi séparées que possible.

1.4 La séparation ontologique-épistémologique

Mon analyse n'est pas une discussion métaphysique/ontologique du hasard. Je ne veux pas me demander dans quelle mesure le hasard existe et quelles formes il prend dans le monde

biologique au niveau moléculaire et cellulaire. Je ne souhaite pas non plus examiner les définitions ontologiques que nous pouvons donner au hasard afin de donner un sens aux processus biologiques dans lesquels il opère, ni proposer un autre positionnement par rapport au débat sur la nature déterministe vs. indéterministe des phénomènes biologiques. Bien sûr, les systèmes biologiques sont des systèmes physiques qui sont inévitablement influencés par le hasard au niveau physique en termes d'agitation thermique et de phénomènes quantiques. De plus, il est maintenant bien connu que les systèmes biologiques gèrent, et parfois même exploitent, le hasard physique en fonction des circonstances et du phénomène de référence. Montrer comment le hasard peut être exploité par les systèmes biologiques est une question empirique et non philosophique. Néanmoins, comme le souligne Kaiser (2015), « [i]l y a plus à dire sur l'ontologie que 'Nous sommes tous des réductionnistes ontologiques. Affaire classée' » (pp. 92 - 93). En d'autres termes, dans notre contexte : il y a plus à dire sur l'ontologie du hasard en biologie moléculaire et cellulaire que le simple fait qu'il soit dû à l'agitation thermique et aux effets quantiques. Je suis tout à fait d'accord avec cette pensée. Ainsi, j'espère que les travaux futurs pourront apporter des contributions ontologiques plus riches sur le hasard en biologie.

Mais en fait, en regardant autour de soi, on se rend compte que le hasard en tant qu'agitation thermique et phénomènes quantiques n'est pas le seul moyen par lequel une discussion ontologique du hasard a été proposée. Il existe, en tout cas, un autre ensemble de publications dans la littérature philosophique qui met de côté l'enquête sur la nature du hasard pour se demander plutôt comment interpréter les probabilités qui sont en fait utilisées pour le décrire et le mesurer. J'identifie ce domaine d'étude comme une réflexion métaphysique/ontologique parce qu'il interroge la nature même de la probabilité, en demandant « Que sont les probabilités ? » Si le hasard existe, que décrivent et mesurent les probabilités ? À l'inverse, si le hasard n'est pas là, peut-on dire que nous utilisons les probabilités uniquement parce que nous sommes des êtres humains limités, incapables de « tailler la nature à sa jointure » (voir Campbell et al. 2011 ; Khalidi 1993) ? De nombreuses réponses ont été données quant à ce que sont les probabilités ; les interprétations classique, logique, subjective, fréquentiste, bayésienne et des propensions (voir Gilles 2000 ; Mellor 2005) sont les principaux exemples (mais pas les seuls) de cette richesse. Dans l'ensemble, pas une seule proposition n'a permis de clore le débat. Dans la mesure où le débat est toujours ouvert, pourquoi mon projet ne prend-il pas position ? Une première réponse pourrait être que la richesse des travaux déjà réalisés me décourage de tenter une quelconque prise de position – mais c'est une raison superficielle. La véritable raison, plus profonde, concerne l'objectif de ma recherche. Pour défendre le rôle

explicatif du hasard dans l'explication biologique, je trouve une stratégie argumentative plus convaincante – j'espère – en ne m'enfermant pas dans le débat sur l'interprétation des probabilités. Je vais montrer que la notion de hasard peut avoir un pouvoir explicatif pour certaines explications biologiques. Je me concentrerai donc sur la notion de hasard et non sur la manière de la décrire et de la mesurer (c'est-à-dire la probabilité). Mais on peut se demander si la probabilité n'aurait pas elle aussi un certain pouvoir explicatif. Comme cela apparaîtra clairement au cours de cette thèse, je me concentrerai principalement sur certaines explications que nous pouvons trouver dans les manuels de biologie et les articles scientifiques de la biologie moléculaire et cellulaire contemporaine. Ces explications ne font pas appel aux probabilités. En effet, elles sont développées dans le but d'expliquer qualitativement le développement d'un certain processus (et non pour mesurer leur caractère hasardeux). Par exemple, l'explication du développement sphérique des parois cellulaires mentionnées ci-dessus ne fait pas appel aux probabilités. Elle précise seulement que les microfibrilles (qui sont les principaux composants de la paroi cellulaire) sont orientées de manière aléatoire (cf. chapitre 3). Cette référence à une orientation « aléatoire » prouve bien sûr qu'une notion de hasard est engagée dans cette explication. C'est sur ce type de notion de hasard (sans probabilité) que je vais me concentrer dans les chapitres qui suivent.

Bien entendu, mon étude du hasard n'exclut pas la possibilité d'étudier le rôle de la probabilité dans certaines explications – ce qui serait sans doute d'un grand intérêt. Mais, pour le dire simplement, il s'agirait d'un autre type d'étude, différent de celui que je me suis fixé pour cette thèse. C'est une chose de s'interroger sur le rôle du hasard dans l'explication du développement de la paroi cellulaire et c'en est une autre de se demander quelle mesure de probabilité peut être attribuée à l'orientation possible des microfibrilles et comment cette probabilité peut être explicative, si elle l'est. Ces deux questions mettent en évidence deux programmes de recherche différents. Ce travail vise à se concentrer sur le premier.

Ainsi avons-nous établi ma volonté de développer une idée de hasard explicatif qui laisse de côté, dans une certaine mesure, la notion de probabilité. Comment, dès lors, ai-je l'intention de développer ce travail ? S'agit-il d'un travail exclusivement épistémologique ? Si oui, quels types de questions pourraient être développés ? J'ai déjà trouvé ma position par rapport au hasard explicatif, pourrais-je maintenant être plus précis et la relier à ma méthodologie pour développer un projet en philosophie de la biologie ?

Mon idée est de passer d'une question concernant le sens et le rôle du hasard dans les phénomènes biologiques à une question sur le sens et le rôle du hasard dans l'explication biologique en réfléchissant au schéma *explanans-explanandum*. Ce déplacement est crucial

pour comprendre ma recherche. Mais même à partir d'une présentation aussi rapide de mon projet, une autre remarque surgit d'emblée concernant l'ontologie. Dire que le hasard peut avoir un rôle central dans l'explication biologique signifie-t-il qu'il pourrait être central pour les phénomènes biologiques également ? En d'autres termes, la séparation épistémologique-ontologique est-elle tenable ou s'effondrera-t-elle inévitablement à un moment donné de l'argumentation ? Par exemple, si je dis que le hasard joue un rôle central dans la réplication de l'ADN, suis-je en train de dire que, là-bas (j'entends par là non seulement dans l'explication mais aussi dans les phénomènes), le hasard joue un rôle fonctionnel dans ce processus moléculaire ? En d'autres termes, à travers mon analyse épistémologique, est-ce que je propose aussi une nouvelle ontologie du hasard ? Il s'agit d'un aspect très délicat de ma recherche car de nombreux philosophes ne séparent pas les dimensions épistémologique et ontologique. Ils soutiennent qu'il est épuisant et inutile de le faire. Par exemple, Rosenberg écrit que la combinaison d'un réductionnisme ontologique et d'un antiréductionnisme explicatif conduit à un « équilibre instable » (Rosenberg 2006, p. 7 ; pour une discussion, voir également Brandon 1996). Cela m'amène alors à me demander si cette idée de séparer l'ontologie et l'épistémologie ne conduit pas inévitablement à un équilibre difficile à maintenir. Malgré mon intérêt pour l'analyse épistémologique, l'épistémologie et l'ontologie doivent-elles finalement aller de pair ? En fin de compte, Rosenberg a-t-il raison ?

Je n'ai pas de contre-argument convaincant et je suis en fait favorable à ce scepticisme général à l'égard de la séparation. Il pourrait être raisonnable de dire que la façon dont nous comprenons et rendons compte du monde va de pair avec la façon dont le monde est réellement. Je ne veux pas développer un propos sophistiqué sur le réalisme scientifique, mais je peux dire que je suis un réaliste en ce qui concerne la science. J'adopte une attitude positive en croyant que les meilleures théories, explications et modèles fournis par les biologistes contiennent des informations qui font progresser notre compréhension de ce à quoi ressemble le monde biologique. Par exemple, lorsque nous faisons référence aux causes, nous ne nous référons pas (seulement) au fait qu'elles peuvent être utiles pour expliquer quelque chose, mais que ces causes existent réellement et qu'en les expliquant, nous ajoutons des connaissances à la façon dont nous concevons le monde. Je suppose que nous sommes tous plus ou moins d'accord pour dire que si nous fournissons une quelconque explication causale, c'est parce que nous croyons que les causes, ou du moins les histoires causales, sont dans le monde – que Hume repose en paix. Je pense que, pour l'essentiel, les explications sont ce qu'elles sont parce que, dans une certaine mesure, elles tentent de donner un sens à la façon dont le monde est réellement. Ce que je viens d'esquisser ici est bien sûr une forme naïve de réalisme. Mais si nous prenons ce

réalisme pour acquis, cela suffit à soulever une objection importante à ma position : d'un point de vue réaliste (même naïf), pourquoi suis-je encore convaincu de proposer une séparation entre épistémologie et ontologie ?

Un argument expliquant pourquoi je me concentre sur l'aspect épistémologique vs. ontologique concernant le hasard est une sorte de priorité temporelle dans le développement de la recherche philosophique. Nous devons d'abord comprendre comment nous décrivons et expliquons le monde afin de pouvoir dire à quoi ressemble le monde (car nos descriptions et explications sont notre seul moyen d'y accéder). Il est donc raisonnable d'affirmer qu'il est nécessaire, dans un premier temps, d'étudier et d'enquêter sur la manière dont les biologistes travaillent, structurent leur travail, pensent et organisent leur pratique quotidienne, développent un certain problème, etc. et seulement ensuite d'aborder le problème de la nature de ce qu'ils étudient, c'est-à-dire l'ontologie du hasard. C'est pourquoi, lorsque j'ai commencé à réfléchir à l'élaboration d'une analyse du hasard en biologie moléculaire et cellulaire, les questions qui se sont posées étaient (surtout) épistémologiques. Mais quelles sont les articulations de ces questions épistémologiques ? La section suivante vise à proposer cette articulation en divisant ma thèse en une partie critique et une thèse positive. Ces questions sont évidemment des développements des deux questions fondamentales que j'ai proposées dans la section 1. En biologie cellulaire et moléculaire, le concept de hasard a-t-il une signification spécifique ? Et, le hasard peut-il avoir un rôle explicatif en biologie moléculaire et cellulaire ?

1.5 Articuler les questions épistémologiques sur le hasard : la partie critique et la thèse positive de mon travail

Ainsi ma contribution se compose-t-elle d'une partie critique et d'une thèse positive. La partie critique est centrée sur les questions suivantes. Pourquoi le hasard est-il conçu de différentes manières dans différents domaines biologiques ? Ces différentes significations influencent-elles le rôle que joue le concept de hasard dans ces divers contextes de recherche ? Une question philosophiquement intéressante est que l'hétérogénéité dans la conception du hasard se retrouve non seulement entre la biologie évolutionniste et les biologies non-évolutionnistes, mais peut également se retrouver au sein des biologies non-évolutionnistes, comme en témoigne l'exemple précité de la biologie moléculaire contre les immunologistes (cf. section 1). Je propose de réfléchir sur la biologie moléculaire et cellulaire parce que, ces dernières années, le rôle du hasard dans ce type d'explications a été de plus en plus étudié, enrichissant ainsi le terrain de la réflexion philosophique. Plus précisément, d'un point de vue épistémologique, j'analyserai les façons dont les biologistes conçoivent le hasard et la manière

dont ces conceptions influencent effectivement leurs recherches. D'un point de vue historique, en abordant spécifiquement la biologie moléculaire, je passerai en revue les motivations qui ont conduit les biologistes à parler du hasard comme d'un bruit, à savoir comme une source de nuisance pour les systèmes (biologiques), en essayant de comprendre pourquoi, aujourd'hui encore, la notion de bruit est utilisée de manière épistémologiquement négative.

En ce qui concerne ma thèse positive, je vais explorer les deux principales questions épistémologiques que j'ai déjà présentées ci-dessus. Le hasard a-t-il un rôle à jouer pour expliquer le bon fonctionnement de certains processus ? Et si oui, en vertu de quoi ? En d'autres termes, quels arguments philosophiques peuvent être formulés en faveur du hasard en tant qu'élément explicatif ? Mon objectif est de convaincre ce lecteur que, dans certaines explications, le hasard peut, en effet, avoir un rôle explicatif essentiel. Pour ce faire, je fournirai un argument philosophique qui repose sur l'idée que le hasard est un abstracteur qui permet à l'explicateur de faire abstraction de nombreux détails qui pourraient nuire à l'efficacité de l'explication. Ce que j'entends par « efficacité » d'une explication (j'utiliserai aussi l'expression de « bonne explication ») sera précisée au chapitre 3, où j'introduirai la notion d'abstraction. Les arguments que je proposerai comme justification sont purement philosophiques. En revanche, pour tester cette idée, j'utiliserai des exemples issus de la biologie moléculaire, cellulaire et du développement. Ensuite, la question de savoir si le compte rendu que je propose dans cette thèse pourrait être étendu à des domaines différents de ceux dans lesquels je l'ai testé (biologie du développement, cellulaire et moléculaire) est une question qui pourra être reportée dans des travaux futurs (cf. conclusion de la thèse).

Le fait de supposer que ces questions épistémologiques sont plus urgentes dans un projet sur le hasard en biologie moléculaire et cellulaire n'empêche pas qu'une analyse plus centrée sur l'ontologie pourrait suivre dans un travail futur. En effet, il me semble que mon travail épistémologique pourrait donner lieu à des questions ontologiques différentes de celles mentionnées jusqu'à présent (le hasard comme propriété des systèmes physiques, l'interprétation des probabilités). Tout au long de ma thèse, j'affirmerai que le pouvoir explicatif du hasard se trouve dans l'explication d'au moins certains phénomènes et pas d'autres. Ainsi apparaît-il que certains phénomènes possèdent des propriétés qui leur permettent d'être expliqués en termes de hasard, mais ce n'est pas le cas pour d'autres. Alors, si le hasard joue un rôle dans l'explication de certains phénomènes, que peut-on en dire ? Le hasard explique-t-il certains phénomènes biologiques parce que ces phénomènes ont des propriétés spécifiques ? Si oui, de quelle nature ? Et comment articuler la relation entre le rôle du hasard dans l'explication et son ontologie dans les phénomènes ? Ces questions ontologiques sont

intéressantes et originales. Cependant, elles dépassent le cadre de ce travail. Mais elles constituent certainement un bon matériau pour des recherches futures (cf. conclusion de la présente thèse).

1.6 Stochasticité, aléatoire et chance : deux caractérisations nominalistes et une philosophique

Je me suis longuement attardé sur certaines questions importantes concernant mon projet sur le hasard, mais j'ai peu parlé de l'objet de mon étude, qui est, pour être précis, le hasard en biologie cellulaire et moléculaire. Dans la dernière partie de cette introduction, je développe la définition de trois notions qui seront centrales pour le développement de mon travail, à savoir la stochasticité, le caractère aléatoire et le hasard. Je souhaite circonscrire les notions de stochasticité et d'aléatoire en m'appuyant directement sur la littérature biologique. En d'autres termes, je veux voir comment les biologistes utilisent ces deux notions, retracer leur signification, et les utiliser dans mon contexte de recherche. Par contraste, je fournirai une définition originale du hasard.

Mais avant de poursuivre, je voudrais formuler trois avertissements. Tout d'abord, le but ultime de cette thèse n'est pas de fournir une définition unifiée du hasard mais plutôt de mettre en évidence ce qui le rend explicatif par rapport à certains phénomènes moléculaires et cellulaires. Les notions que je proposerai contribueront à cet objectif principal. Deuxièmement, comme je l'ai mentionné dans la section précédente, je me concentre sur le hasard dans l'explication des processus biologiques (vs. leurs résultats biologiques). Troisièmement, la partie originale de mon travail ne se trouve pas seulement dans la réponse (à savoir pourquoi le hasard peut être explicatif) mais aussi dans la question elle-même. En effet, se demander si le hasard a un rôle dans les explications cellulaires et moléculaires est une question qui, à ma connaissance, n'a jamais été formulée aussi explicitement dans un contexte philosophique.

1.6.1. Les processus stochastiques

Même si ce n'est pas toujours le cas, lorsque nous pouvons trouver la notion de stochasticité dans la littérature biologique, elle se réfère généralement à une propriété des processus. Par exemple, nous pouvons trouver des déclarations telles que : « [n]ous constatons que la sélection du site de départ est *largement stochastique* » (Boersma et al. 2019, p. 459, je souligne). Ici, la « sélection du site de départ » est un processus par lequel le ribosome se fixe à un ARN pour initier la traduction et « stochastique » désigne la propriété attribuée à ce processus. Prenons un autre exemple :

« Le système immunitaire exploite la *stochasticité* inhérente aux petits nombres dans le ‘processus de recombinaison VDJ’ qui donne naissance à la diversité des récepteurs des cellules T » (Honegger et de Bivort, 2018, p. R10, je souligne).

La recombinaison du gène VDJ est un processus qui se produit au cours du développement des lymphocytes B. La recombinaison VDJ implique la partie des gènes qui, une fois que le lymphocyte est mature, coderont pour les anticorps. La recombinaison aléatoire VDJ permet la production d’une grande variété d’anticorps. Mais en dehors de ces détails techniques, ce qui est intéressant aussi dans ce cas, c’est que la « stochasticité » fait référence à la propriété d’un processus, c’est-à-dire la recombinaison VDJ.

Essayons maintenant de donner une définition plus formelle de ce que signifie la stochasticité par rapport aux processus (biologiques). Pour cette définition, j’emprunte le lexique donné au processus stochastique dans un ouvrage technique sur la théorie des probabilités :

« [Les] processus *stochastiques* [sont] des systèmes qui évoluent de manière probabiliste dans le temps ou, plus précisément, des systèmes dans lesquels existe une certaine variable aléatoire X(t) dépendant du temps [:]
P (x1t1 ; x2t2 ; x3t3...) » (Gadiner 2004, p. 42 ; souligné dans l’original)

Un processus est stochastique si les variables utilisées pour le décrire ne varient pas régulièrement dans le temps mais de façon aléatoire sur un espace de probabilité donné dans lequel sont distribuées ses multiples valeurs possibles. Ce type de variables est représenté dans la définition proposée par « X(t) », c’est-à-dire une variable « X » qui varie aléatoirement en fonction du temps « (t) ». « P(x1, t1 ; x2, t2 ; x3, t3 ; ...) » est une formalisation supplémentaire de cette idée : « P » représente la probabilité qu’à t1 « X » soit « x1 » et qu’à t2 « X » soit « x2 », et ainsi de suite.

Il y a une objection immédiate à cette définition. Dans la section 1.4, j’ai précisé qu’il n’était pas dans mon intention de rendre compte du hasard en termes de probabilité. En revanche, développer cette définition de la stochasticité en termes de probabilité pourrait conduire à la présomption que je cède à la contradiction. C’est une objection légitime. Cependant, il n’en est rien puisque cette notion de stochasticité ne sera pas utilisée pour développer ma proposition sur le hasard en biologie cellulaire et moléculaire. La fonction de cette définition est seulement de faire comprendre au lecteur ce que j’ai à l’esprit lorsque

j'utilise des notions telles que « processus stochastique », « phénomènes stochastiques », « expression génétique stochastique », et ainsi de suite. En bref, je peux dire que cette caractérisation de la stochasticité a une fonction nominaliste : elle sera utilisée pour identifier les processus que les biologistes décrivent comme stochastiques (voir Tableau 3).

1.6.2 Résultats aléatoires

Parlons maintenant de la notion d'aléatoire. Au regard de la littérature, nous constatons que le terme « aléatoire » est utilisé de différentes manières, et pour caractériser à la fois des processus et des résultats. À la suite de Merlin (2009), j'ai choisi d'utiliser « aléatoire » et « caractère aléatoire » pour désigner les résultats. Comme pour la « stochasticité », je donne à la notion d'« aléatoire » une fonction nominaliste : elle sera utilisée pour identifier les résultats qui sont décrits par les biologistes comme aléatoires, plus précisément, des résultats qui sont la plupart du temps imprévisibles, ou qui ne suivent pas des schémas réguliers, mais qui peuvent néanmoins (éventuellement) être décrits par une loi ou une distribution de probabilité (voir Tableau 1). Ainsi, comme pour les résultats du lancer d'un dé ou d'une pièce de monnaie, l'inactivation du chromosome X ou l'apparition de mutations sont dites aléatoires. Les deux notions de stochasticité et d'aléatoire ont été proposées pour caractériser respectivement les processus et les résultats à la lumière de leur description la plus répandue. Dans la section suivante, je suivrai un cheminement différent avec la notion de hasard.

1.7 Le hasard est une propriété de la description d'un processus

Permettez-moi de souligner à nouveau l'objectif de ma thèse. Avec un argument philosophique, je veux convaincre le lecteur que le hasard est essentiel pour certaines explications en biologie moléculaire et cellulaire. Pour ce faire, je proposerai ce que j'appelle « l'explication stochastique » (SE), une catégorie philosophique qui se réfère à tout type d'explication qui inclut un rôle explicatif pour le hasard. Mais la tâche d'expliquer en détail ce que j'entends par cette explication, à quoi elle s'applique, et quelles implications elle a par rapport au hasard sera développée dans les chapitres 2 et 3. La présente section vise uniquement à fournir une première définition générale de la notion de hasard qui sera au cœur du développement de l'explication stochastique. Elle se présente comme suit :

Le hasard est une propriété de la description d'un processus qui, étant donné un ensemble fixe de conditions initiales, pourrait être autre.

Examinons cette définition point par point.

1) « Le hasard est une propriété de la description d'un processus ». Le développement de cette partie de la définition est principalement motivé par le fait que, puisque mon travail est une analyse épistémologique, je ne veux pas le développer par des inférences ontologiques. Par conséquent, dans mon cadre, le hasard n'est pas une propriété du processus lui-même mais une propriété par rapport à la manière dont il est décrit. À cette fin, j'ai essayé de construire une définition qui ne se réfère pas directement au phénomène, mais à la manière dont il est décrit. Je postule que le hasard n'est pas une propriété du phénomène mais plutôt une propriété de la description d'un phénomène. « D'une description de » crée un espace épistémologique dans lequel un travail à l'abri de toute inférence forte par rapport à la métaphysique pourrait être construit.

2) « Étant donné un ensemble fixe de conditions initiales ». Cette première partie de la définition est essentielle car s'il n'y avait pas de conditions initiales fixes, le hasard n'existerait nulle part : tous les événements seraient causaux, dépendants de leurs conditions antérieures. Si la condition initiale X n'était pas le cas, l'événement Y ne s'ensuivrait pas. Au contraire, si la condition initiale avait été Z, l'événement aurait été Q. Cet argument est résumé dans le tableau 1.

Pas de conditions initiales fixes
Conditions initiales X → Événement Y
Conditions initiales Z → Événement Q
Le hasard n'est nulle part : Si les conditions initiales X étaient différentes (par exemple Z), l'événement successif Y serait également différent (par exemple Q).

Tableau 1

Le grand avantage de fixer les mêmes conditions initiales réside dans le fait que nous dépassons la notion de dépendance causale. De cette façon, nous établissons en fait le scénario pour construire une notion de cas en termes de descriptions de processus qui peuvent être différents à partir des mêmes conditions initiales fixées. De plus, nous devons préciser que l'ensemble des conditions initiales que nous prenons sont locales et non universelles. Voir le tableau 2 ci-dessous.

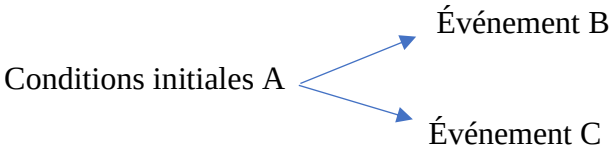
Conditions initiales fixes    Le hasard spécifie que, étant donné (un ensemble de) conditions initiales (A), plus d'un événement est possible.

Tableau 2

3) « [Le processus] peut être autrement ». Comme nous le rappelle Wong, l'évolution est contingente au moins parce que les résultats de l'évolution pourraient être différents (Wong 2020). Par exemple, l'hypothèse selon laquelle les résultats de l'évolution pourraient être différents est normalement associée aux contrefactuels « si nous avions A, nous aurions B ». Ce contrefactuel est basé sur le fait que l'histoire s'est déroulée d'une certaine manière, et nous nous demandons ce qui se serait passé si les choses avaient été différentes (même si elles ne l'ont pas été en réalité). Cette partie de ma définition du hasard reprend certains éléments du débat sur la contingence (qui s'est déroulé entre les années 1980 avec Gould jusqu'aux derniers travaux de Turner 2011a, 2011b) mais introduit quelques modifications. Tout d'abord, puisque je me concentre sur les processus répétables, je remplace « pourrait » par « peut ». Par exemple, je soutiendrai que l'explication de la réplication de l'ADN peut être décrite et expliquée par une étape fortuite importante, et jusqu'à présent négligée (voir chapitre 3, section 3.3). J'utilise « peut » puisque le processus considéré est la réplication de l'ADN, à savoir un processus qui se répète continuellement au cours de la vie cellulaire. Deuxièmement, je n'utilise pas le terme « résultat » dans ma définition, comme le fait Wong, mais « processus ». Je m'intéresse aux processus biologiques et à la manière dont nous pouvons les expliquer (cf. section 1.5). Je ne veux pas seulement expliquer pourquoi le résultat de la réplication de l'ADN peut être différent, par exemple en raison d'une « erreur » dans la duplication (dans les termes de Monod), mais je veux être capable d'expliquer comment le processus de réplication de l'ADN lui-même peut être différent et produire, en conséquence, des résultats différents. Par exemple, je m'intéresse à la manière dont nous pouvons expliquer que différents nucléotides peuvent être ajoutés à la séquence d'ADN nouvellement synthétisée par le biais de diverses interactions physico-chimiques (processus) et pas seulement au fait que ces nucléotides sont, en fin de compte, ajoutés (résultat). Résumons les définitions qui viennent d'être proposées dans le tableau 3 ci-dessous :

Termes	Définitions
Stochasticité (fonction nominaliste)	Terme général que j'utilise pour désigner les processus/événements. Je définis « stochastique » en utilisant la définition formelle suivante : « [Les] processus *stochastiques* [sont] des systèmes qui évoluent de manière probabiliste dans le temps ou, plus précisément, des systèmes dans lesquels il existe une certaine variable aléatoire X(t) dépendant du temps [...] » (Gadiner 2004, p. 42 ; souligné dans l'original).
Aléatoire (fonction nominaliste)	Termes généraux que j'utilise pour désigner les résultats de processus/événements : ils sont la plupart du temps imprévisibles ou ne suivent pas des schémas réguliers mais, néanmoins, peuvent (éventuellement) être décrits par une loi ou une distribution de probabilité.
Hasard (fonction philosophique)	Terme philosophique qui renvoie à la définition suivante : Le hasard est une propriété de la description d'un processus qui, étant donné un ensemble fixe de conditions initiales, peut être différent.

Tableau 3. Définitions des notions utiles à l'élaboration de mon affirmation épistémologique du hasard en biologie moléculaire et cellulaire.

Ces notions sont les bases à partir desquelles je vais commencer à développer mon idée d'explication stochastique (SE) en biologie cellulaire et moléculaire. Celle-ci sera développée en détail à partir du chapitre 2.

1.8 La structure symétrique du présent travail

Après avoir précisé le sujet et le but de ce travail, ce que sera l'attitude envers la notion de hasard que je vais proposer, et la signification spécifique de chaque terme et notion au cœur de l'argumentation de la thèse, je veux illustrer ici que mon travail a une structure symétrique.

Il est composé de sept chapitres, dont le quatrième est une sorte d' « interlude » qui permet au lecteur d'absorber ce qui a précédé et de se préparer à ce qui va suivre. Le chapitre 4 marque également une division en ce qui concerne le type de travail proposé. Les trois premiers chapitres ont une nature fortement philosophique, dans le sens où ils exposent les préliminaires métaphilosophiques à partir desquels je développe mon récit (Chapitre 1), discutent de la littérature philosophique pertinente pour mon travail (Chapitre 2), et présentent mon récit de l'explication stochastique en détail (Chapitre 3). Toutefois, cela ne signifie pas qu'ils ne contiendront aucun contenu biologique. Au contraire, le contenu biologique permettra de dégager de nouvelles questions philosophiques. En comparaison, les trois derniers chapitres sont davantage axés sur la biologie. Je décrirai et analyserai en détail des études de cas tirées de la biologie moléculaire, cellulaire et du développement, qui renforcent mon explication du hasard. Je fournirai une analyse de l'expression génétique qui concerne spécifiquement la transcription (chapitre 5), l'épissage alternatif (chapitre 6) et la traduction (chapitre 7). Cela ne signifie pas pour autant qu'une composante philosophique ne sera pas présente. A travers les exemples biologiques que nous considérerons, je montrerai qu'ils sont parfaitement compréhensibles comme des instances spécifiques de l'explication stochastique (ES).

Ainsi, cette thèse a une structure symétrique : les trois premiers chapitres philosophiques sont équilibrés par trois chapitres plus centrés sur la biologie. L'interlude agit comme un point d'appui et maintient les six chapitres en équilibre et en suspension (voir Figure 1).

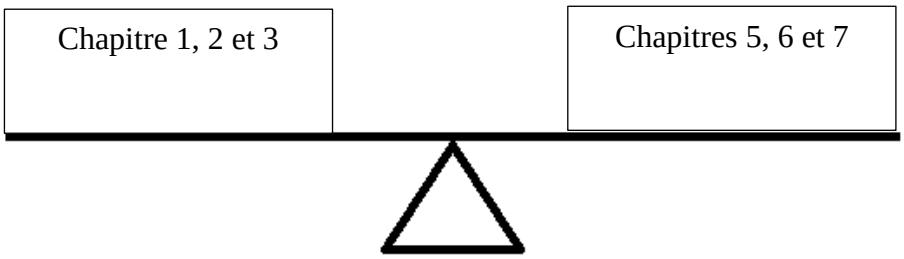


Figure 1. Une représentation qualitative de l'organisation de ma thèse.