

First draft – Do not quote without the authors' permission

Synthetic Modeling and the Functional Role of Noise

Tarja Knuuttila
University of Helsinki
Andrea Loettgers
California Institute of Technology

Abstract:

In synthetic biology the use of engineering metaphors to describe biological organisms and their behavior has become a common practice. The concept of noise provides one of the most compelling examples of such transfer. But this notion is also confusing: While in engineering noise is a destructive force perturbing artificial systems, in synthetic biology it has acquired an additional functional meaning. It has been found out that noise is an important factor in driving biological processes such as gene regulation, development, and evolution. How did noise acquire this dual meaning in the field of synthetic biology? In this paper we study the emergence of the functional meaning of noise in relation to synthetic modeling. We will pay particular attention to the interdisciplinary aspects of this process highlighting the way borrowed concepts, analogical reasoning and the use of cross-disciplinary computational templates were entwined in it.

1. Introduction

One of the most conspicuous features of contemporary modeling practice is the way modeling methods and formalisms move across disciplinary borders. Yet this is not a novel phenomenon. For example the pioneers of mathematical economics and biology in the 19th and early 20th century made extensive use of methods, templates and concepts borrowed from physics. This methodical and instrumental

transdisciplinarity has provided for some analysts the distinguishing mark of modeling, or computational science. Historian of science Giorgio Israel (1993) has claimed that the modeling paradigm maturing at the 20th century revolves around formal structures that are applied to various phenomena on the basis of some observed similarities. Somewhat similarly, Paul Humphreys (2002, 2004) has suggested that what he calls computational templates provide a convenient unit of analysis for computational science exactly because of their all-pervasiveness.¹

Though the notion of a model got its present-day meaning first during the 20th century, modeling as we understand it today started already earlier. Interestingly, what we call today mathematical models were called mathematical analogies in the 19th century (Bailer-Jones 2009, Israel 1993). This usage is telling as in modeling methods, formalisms and scientific concepts are often transferred from one domain to another by way of analogical reasoning. In the following we will study the intricacies of such analogical reasoning in synthetic biology, which is a relatively novel and highly interdisciplinary field located at the interface between engineering, physics, biology, chemistry and mathematics. The research practice of synthetic biology is a combination of methods, concepts, tools, and theories from these fields. Of these cross-disciplinary influences, however, those of engineering are most visible. In synthetic biology the use of engineering metaphors to describe biological organisms and their behavior has become a common practice. A host of engineering notions such as robustness, modularity, and redundancy have both served as basic theoretical concepts of the field as well as vehicles for public understanding of synthetic biology.

The concept of noise provides a particularly compelling and challenging example of such interdisciplinary transfer. This is due to the double meaning it has gained in the field of synthetic biology. In engineering noise is usually regarded as a nuisance: it is considered as a destructive force perturbing the functioning of engineered systems. Yet the situation is different when it comes to biology, where

¹ With computational templates Humphreys refers to genuinely cross-disciplinary computational devices, such as functions, sets of equations, and computational methods, which can be applied to different problems in various domains.

noise has also acquired an additional functional meaning. Biological systems appear to make good use of noise in diverse processes, including development (Paldi, 2003), differentiation (e.g. in genetic competence (Cagatay et al., 2009)), and evolution (Eldar & Elowitz, 2010). A new kind of a model, a synthetic model, contributed importantly to the development of the functional meaning of noise making it an intensively discussed and studied subject in synthetic biology. But side by side the functional meaning of noise also the more traditional understanding of noise as parasitic and harmful force still continued to exist in synthetic biology.

What is, then, the epistemic rationale of using the notion of noise in both of these opposite meanings – even within a same field? How did this situation come about? We will study the evolution of the functional meaning of noise paying attention to its material, disciplinary, historical, and conceptual conditions. As regards analogical reasoning, we will show that the two branches of synthetic biology, the engineering-oriented branch and the basic science-oriented branch made different uses of some basic engineering notions, such as electric circuit. Moreover, we will also examine the role played by the particular modeling praxis of the basic science oriented branch of synthetic biology that employs mathematical models, model organisms, and synthetic models in a combinatorial fashion. First, however, we will discuss briefly how philosophers and cognitive scientists have approached analogical reasoning.

2. Analogical reasoning and interdisciplinary exchange

In the philosophy of science analogical reasoning has often been discussed in the context of knowledge generation: in scientific discovery and theory development, and hypothesis formulation. The important role of analogies in the aforementioned activities has generally been admitted but their epistemic status has been a matter of disagreement. Whereas some philosophers have considered analogies as only heuristic tools, others have proposed that scientific theories and models could be conceived through the idea of analogy (e.g. Hesse 2001, Harre 1970). According to this view, which Hesse (2001) has dubbed the analogical conception of theories,

scientific models (or theories) are considered as analogues to the aspects of their real world targets. This conception has also been adopted by the discussion on mental modeling and (mental) model-based reasoning. Nancy Nersessian describes analogical modeling process as the evaluation of how well the “constraints of a model fit the salient constraints of target problem” (Nersessian 2002 , 138). As our interest is in understanding interdisciplinary exchange we focus rather on the analogies between different fields of inquiry. From this perspective analogical comparisons between two (or more) domains provide one prominent cognitive strategy used in modeling (cf. Bailer-Jones 2009). Mary Hesse’s work, especially Hesse (1966) provides a locus classicus also for this debate.² Her distinction between positive, neutral and negative analogies provides a handy tool for studying the analogical reasoning process. Positive analogy refers to those properties that the two analogues have in common, whereas negative analogy refers to the known differences between them. Neutral analogies refer to the properties whose commonality or difference has yet to be established and thus they provide, according to Hesse, the epistemic potential for further inferences and theoretical development. They suggest specific questions to study and the ways to extend the theory.

Hesse suggests also two other classifications. She distinguishes between material and formal analogies and “horizontal” and “vertical” relations. There is a formal analogy between two domains if the relations between certain elements within one domain are identical with, or at least comparable, to corresponding elements in another domain. Material analogies in turn require the two domains or analogues to have at least certain properties in common. For Hesse they are pretheoretic analogies between observables. As regard properties there can be “horizontal” and “vertical” relations between them. “Horizontal relations” refer to corresponding (similar) properties of the two domains, whereas “vertical relations”

²² It deserves to be noted that for instance in the writings of Hesse and Nersessian the ideas of models as analogues of the the real world targets and analogical comparisons between two domains of inquiry often coalesce. This is justified by the idea that an analogue from one domain can serve as a model of another domain, as Hesse’s well-known example of the billiard ball model of the “dynamical” theory of gases shows.

are relations between the properties within a domain. The two domains are formally analogous if there are similar with respect to their vertical relations.

Hesse's aforementioned distinctions come close to cognitive scientist Dedre Gentner's theory of analogy (1983). She distinguishes between attributes and relations and claims that an analogy does not become stronger only if the two analogues share more attributes. Instead she thinks that in analogy the key similarities are those that lie in the relations that hold within the domains viewing analogy as structure mapping between the base and target domains (see also Gentner and Markman 1997). In her stress on the connectedness of knowledge she focuses on Hesse's vertical relations. Both Gentner and Hesse stress the importance of the analogical transfer of the *relations* within the domain, which is what Gentner call systematicity. Such systematicity is a central feature of mathematical and computational models, which typically study the dynamic behavior of a system of interconnected variables (Knuuttila 2008).

Furthermore, Hesse seems to think that material analogy provides a basis for finding about formal analogies (cf. Bailer-Jones 2009; p. 58). With respect our case on genetic "circuits" (i.e. genetic regulatory systems) this is an important point: the genetic circuits have been conceptualized in terms of electric circuits, which analogical move has made genetic regulatory systems amenable to further conceptualization and formalization. Drawing material analogies to other kinds of systems, or employing theoretical concepts, such as noise, depicting certain kinds of systems or behaviors, seem to be needed to animate formal analogies and to give them theoretical content.

More generally, in the cognitive science and philosophical discussion on analogies one can discern following common features:

- Analogy is approached in terms of *similarity* and *familiarity*: one makes sense of a relatively novel domain in terms of an better known, familiar domain that is thought to be similar to the domain in question
- Analogy is conceptualized as a *mapping between target and source domains*

- As regards analogical reasoning the focus is on the *shared structure or dynamics of the two domains*, and on neutral analogies as kind of possibilities for further theoretical development
- *Analogy enables the mathematization* of the target domain in terms of the mathematical representation of the source domain

Although we find these points important and revealing, on the basis of our case study on synthetic biology we would like to make some amendments to this received view on analogical reasoning. Firstly, it seems to us that negative analogies carry much more epistemic weight than what the discussion on analogies lets us to expect (see however Morgan 1997). Secondly, we find the target-source pair too restrictive a unit of analysis: as we will show, the notion of noise subsumed a much more heterogeneous fabric of tools, methods, templates and concepts that were transferred from various disciplines to the study of gene regulatory mechanisms. Thirdly, as already pointed out, analogies function as means of mathematization, as a way to introduce a formalism from one domain to another, a point that was already made by Maxwell when he wrote that the aim of analogy is “to enable the mind to grasp some conception or law in one branch of science, by placing before it a conception or a law in a different branch of science, and directing the mind to lay hold of that *mathematical form which is common to the corresponding ideas in the two sciences* (1890, 219, emphasis by the authors of this article). However, limiting the focus on the respective representations of the source and target domains loses the sight of the genuinely cross-disciplinary computational templates that often provide the major tools of modeling (cf. Humphreys 2004). Synthetic biology like many other fields of computational science draw their theoretical templates from the repository of formal systems studied by the dynamical systems theory and applied in a variety of disciplines to a wide range of entirely different phenomena.

3. Noise in synthetic biology

In synthetic biology two different main branches can be distinguished: an engineering and application oriented branch and a basic science oriented branch. Within the two different branches different meanings of noise and its role in biological systems are predominant. Scientists in the engineering and application oriented branch of synthetic biology aim for the design of novel biological parts or even systems (e.g. Endy 2010, Purnick & Weiss 2009). On their agenda one finds the development of new vaccines by manipulating/changing the metabolic systems of bacteria (Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J., Modrich, M., Smolke, C., & Weiss, R., 2006) or the *engineering* of cancer killing bacteria (Lim 2010). What is interesting, is that the application and the basic science oriented branches of synthetic biology draw at times nearly opposing analogies to engineered artificial systems, as exemplified by the case of noise.

Following an engineering paradigm, the scientists in the application oriented branch view noise mainly as a disturbance interfering with the correct functioning of the designed biological system. In contrast, scientists in the basic science approach have started to consider seriously the differences between the functioning of artificial systems and biological systems. These scientists have often background in physics and they aim to gain insight into the organizational principles of biological systems. The design of synthetic models, that is, *simplified novel biological systems*, offers a new tool for the study of the possible design principles of biological systems. As an example we will discuss the *Repressilator*, an oscillatory genetic circuit designed exploring the regulating mechanisms underlying biological functions such as the day and night rhythm of organisms. The dynamics of such networks is often analyzed in terms of non-linear feedback mechanisms leading to complex behavior. Such complex systems can show surprising behavior exemplified for instance by how noise in the form of stochastic fluctuations can be beneficial for the behavior of the system. Even though many scientists presented this as a new and surprising result in the beginning of the 2000's, we will show that, on the one hand, external noise and internal noise in the form of stochastic fluctuations had

already been studied in the field of complex systems and biology. Many scientists of the basic science oriented branch of synthetic biology were already familiar with these results and/or started collaborations with experts on the effect of noise on complex systems.³ Thus the negative analogy concerning the role of noise in engineered artificial systems vis-à-vis biological systems was already anticipated by previous research and gave scientists conceptual and empirical resources to interpret the findings of synthetic modeling accordingly.

3.1. Electric vs. genetic circuits

In the field of synthetic biology noise became an intensively discussed topic as a result of synthetic modeling. Synthetic models are located in-between mathematical models and model organisms through their construction, which combines properties and features of both types of models. Using a mathematical model as a blueprint, synthetic biologists construct synthetic versions of mathematical models from biological components like genes and proteins. Synthetic models share with the mathematical models the structural features, such as architecture and dynamics, and with model organisms the right kind of materiality. The modeling practice of synthetic biology is combinatorial, synthetic modeling is triangulated with mathematical modeling and simulation as well as with experiments on model organisms. This is due to the different constraints of the three model types: although, for instance, synthetic models capture the structures and dynamics represented by certain mathematical models, they are realized in different media with highly different affordances.⁴ Synthetic biologists regard as the main advantage of this novel type of models the combination of the same materiality as that of the biological systems and the reduced complexity, which affords higher degree of control than model organisms do. Although one might thus think that synthetic

³ E.g. Michael Elowitz began a collaboration with Jordi Gracia Ojalvo.

⁴ For a more detailed discussion on the constraints introduced by the materiality of the model see Knuuttila, T., Loettgers, A. "Modeling and experimentation: The combinatorial strategy in synthetic biology." (http://philsci-archive.pitt.edu/8329/1/Modeling_and_experimentering_AL_TK.pdf)

models are superior to mathematical models, the flexibility of mathematical models, which is due to their symbolic mode, enables researchers to study many different possibilities with them far more easily than with synthetic models. In practice mathematical models and simulations are used to study the results of synthetic modeling. On the other hand, synthetic models being novel engineered biological systems, are not results of evolutionary process as naturally occurring biological organisms are that makes experiments on model organisms an ineliminable part of the research practice of synthetic biology.

An example of such a synthetic model is the *Repressilator*, an engineered oscillatory genetic network, which is one of the first and most famous of synthetic models (Elowitz & Leibler 2000). In the field of synthetic biology genetic networks such as the *Repressilator* are conceptualized in two different ways, with both rely on analogies drawn to electrical circuits as regards their basic structure and properties. What are these analogies? Genetic networks are typically modeled on electrical circuits in that that they make use of feedback mechanisms. Interestingly, however, the basic science oriented branch and engineering oriented branch make different uses of this original analogy. In both cases the purpose of the feedback mechanism is to perform control over specific functions. But there is an interesting twist to this.

Engineered systems use feedback mechanism to adjust the output to a given reference value. An example of this kind of feedback mechanism is provided by the thermostat, which measures the room temperature (input), compares it to a reference temperature (output), and in the next step changes the heater so that the room temperature is adjusted to the reference temperature. The *Repressilator*, in turn, is an example of the basic science approach in synthetic biology, which relies to a different kind of adjustment mechanism. In the feedback mechanism of the *Repressilator*, oscillations in protein levels provide the essential element of control. Consequently, the feedback mechanism works in a different way in engineered systems and in biological systems.

This difference between the artificial and biological control mechanisms was recognized already early on in the study of the circadian clock, which regulates the day and night rhythms of organisms. Brian Goodwin, whose model provided an

important basic template for the subsequent models of the circadian clock (including also the *Repressilator*) described this point in the following way: “The appearance of such oscillations is very common in feedback control systems. Engineers call them parasitic oscillations because they use up a lot of energy. They are usually regarded as undesirable and the control system is nearly always designed, if possible, to eliminate them.” (Goodwin 1963, 5). Consequently, at this point the analogy between electric and genetic circuits breaks down and the difference between them paves the way for new biology specific insights on the organization and functioning of biological systems. With the oscillations as the central element of control in biological systems the dynamics exhibited by these systems becomes a different one. This dynamics was probed by the *Repressilator*, which pointed also towards a further negative analogy between the control mechanisms of engineered and biological systems.

3.1 The Repressilator as an example of engineered genetic circuit

The *Repressilator* is an engineered oscillatory genetic network that consists of three repressor genes, each of which inhibits the expression of the following gene leading to oscillations in the protein levels. Figure (1) shows a simplified outline of the *Repressilator* commonly used in explaining the basic structure and dynamics of the system.

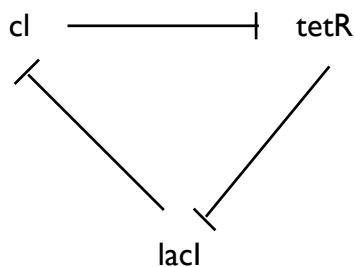


Figure 1. A diagrammatic depiction of the *Repressilator*. The three genes cl, tetR, and lacI repress each other.

In reality this system is far more complex. Already genes themselves are fairly complex entities. The same counts for the biochemical processes taking place in the network. Furthermore the synthetic model is not an isolated object but imbedded in an *E. Coli* bacterium. Whereby the assumption is made that the *Repressilator* is a module and not interacting with the rest of the cell environment and thereby not exposed to external noise coming from the rest of the cell. In constructing the *Repressilator* Leibler and Elowitz tried to find such design principles of biological systems that would lead to stable oscillations like those observed in the circadian clocks. The mathematical model on which they based the design of the synthetic model as well as the computer simulations performed with the mathematical model predicted the existence of stable oscillations: "the system may converge toward a stable steady state, or the steady state may become unstable, leading to sustained limit-cycle oscillations." (Elowitz and Leibler 2000, p. 335) The two scientists aimed for sustained limit-cycle oscillations but unfortunately the synthetic model they built did not exhibit the behavior predicted by the mathematic model used to guide its construction. The oscillations in the protein level—made visible by connecting one of the three repressor genes of the network to a green fluorescent protein (GFP) leading to oscillations in the intensity of the light emitted by the GFP—showed irregularities. Figure (2) is an example of such single cell observations. The arrow in the upper part and lower part of images (a) and (b)⁵ indicates a single *E. coli* bacterium over a period of time in which the population of bacteria did grow. The analysis of the intensity of light emitted from the GFP in the cell led to the oscillations depicted in Figure (3). Comparing the oscillations made visible by the fluorescence of single cells taken from a population allows the exploration the irregularities in the oscillatory behavior. Figure (4) shows the comparison of the oscillations of three different cells. One sees a shift in the phase of the oscillations and difference in the period of the oscillations.

⁵ The division into the two parts is not due to a difference in content but only to the way the pictures were taken.

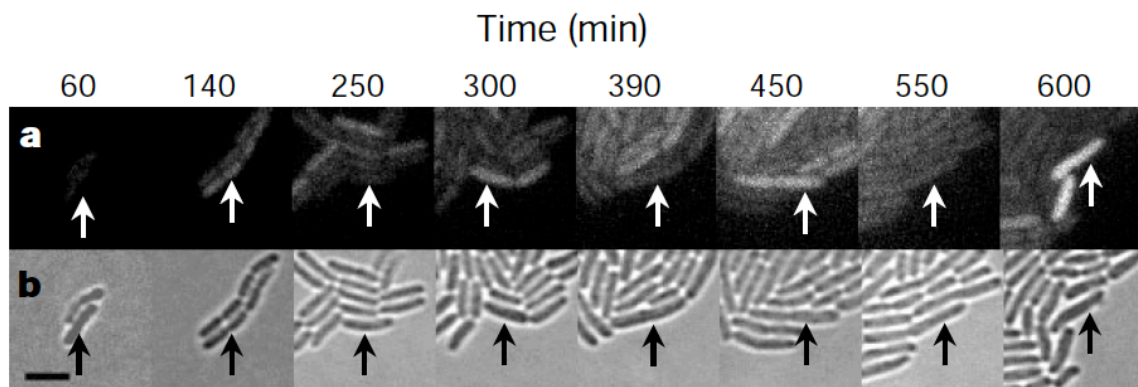


Figure 2. Observations on the oscillations in protein level of a single cell (Elowitz & Leibler, 2000)

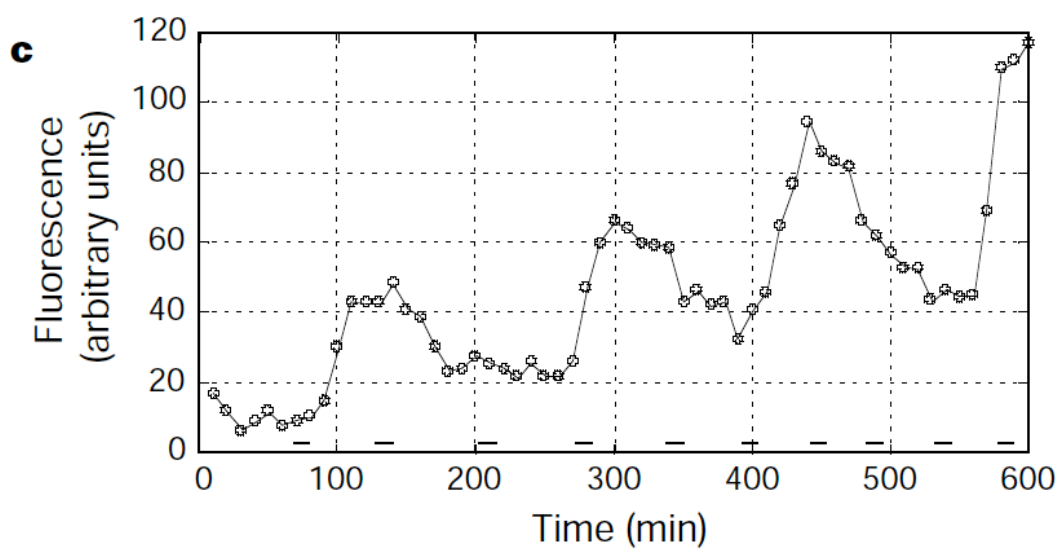


Figure 3. Diagrammatic representation of the oscillations in fluorescence (Elowitz & Leibler, 2000)

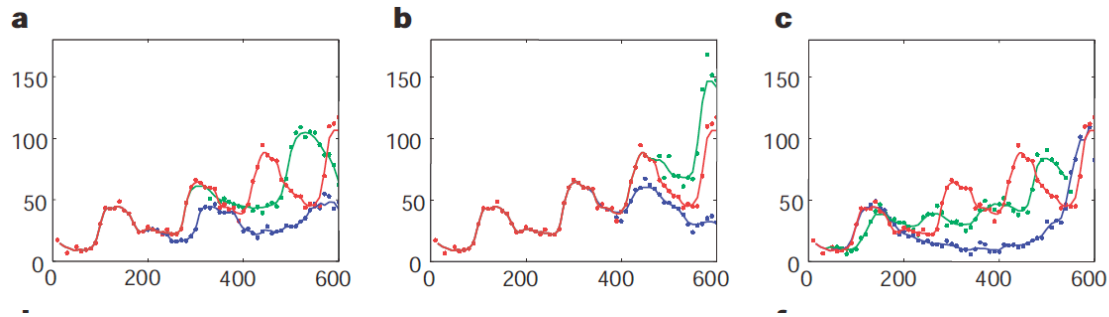


Figure 4. Comparison between the oscillations observed in single cells of a microcolony. (Elowitz & Leibler 2000)

The original mathematical model used was a deterministic model, which did not take into account stochastic effects such as stochastic fluctuations in gene expression due to the low number of molecules in many biological systems such as cells. Performing computer simulations on a stochastic version of the original mathematical model Leibler and Elowitz were able to reproduce similar variations in the oscillations as observed with the synthetic model. This led to the conclusion that stochastic effects may play a role in gene regulation meaning that the original analogy to electrical circuits, already adapted to biological systems by way of relegating oscillations a functional role, proved even more lacking. Indeed, what the *Repressilator* actually accomplished was to provide a test whether the analogy to electric circuits and the idea of negative feedback mechanism really worked in the case of actual biological systems. Are naturally evolved genetic regulatory networks constructed in such a way? The answer was only partly positive, but on the other hand the partial failure of the *Repressilator* provided also an important trigger for a new line of research exploring the different sources of noise and their effects on biological systems.

It should be noted that in the case of the *Repressilator* the stochastic fluctuations had still the character of a disturbance and the question became how to change the design of the *Repressilator* in such a way that it would lead to a more robust behavior. But what started to intrigue the researchers in the field was whether the stochastic fluctuations observed could also have a functional role. In

the subsequent research in the basic science oriented branch of synthetic biology, noise based on stochastic fluctuations gained a functional status. An example of the functional role of noise is provided by behavior of the bacterium *Bacillus subtilis*. In *B. subtilis* competence is a state in which the cell is able to take up DNA from the environment. The state is a transient state. After entering the state of competence *B. subtilis* will after some while exit this state again. In a population of identical cells not all will become competent but only 20 percent. This observation has raised the question of what kind of dynamics controls this probabilistic-transient dynamics. Experiments in which the dynamics of the underlying genetic circuits were studied by time-lapse fluorescence movies showed that “excitable dynamics driven by noise naturally generate stochastic and transient response” (Suel 2006, p.545). This specific kind of excitability has also been observed, as noted by the authors, in other systems such as neurons and it provides a mechanism in controlling transient states. Consequently, stochastic noise caused by variations in the gene expression of different cells, need not be a perturbation but rather an essential ingredient of the dynamics responsible for the uptake of DNA from the environment.

4. External and internal noise

Although the current intense discussion on the role of stochastic fluctuations, or noise, in controlling biological systems seems to give the impression that it is a novel topic, this is not the case. The origin of this idea can be traced back to at least to two different bodies of work, with both have had an impact on the development synthetic biology. Firstly, many synthetic biologists are physicists and draw their analogies to systems, which due to non-linear dynamics show complex behavior. Methods, techniques and computational templates developed in these studies have been transferred to synthetic biology. As we have seen, synthetic biologists often use the term noise to refer to stochastic fluctuations in gene expression (or in other processes).⁶ Such inherently random processes have been extensively studied in

⁶ Such stochastic fluctuations are caused by the low number of molecules in the cell.

statistical mechanics, both conceptually and experimentally. The functional role for noise was already evident in the more general field of the study of complex systems, where during the 1970's the influence of *external* noise on the behavior complex systems became an object of study. Secondly, also in 1970's biologists were studying the role of *internal* noise. In contrast to the studies on the external noise in complex systems, in the study of stochastic fluctuations intrinsic to biological systems theoretical modeling was interlinked with experimentation.

4.1 External noise enlarging the possible states of the system

The first time noise became recognized not only as a disturbance but something that could be beneficial was in the context of the research on complex systems. In the 1970's scientists such as the Russian physicist Alexander S. Mikhailov started to investigate noise-induced phase transitions in open systems such as biological, chemical, and physical systems. The main results of this research run against the common understanding according to which: "all relevant bifurcations⁷ are already present in dynamical equations and introduction of external noise or taking into account the internal statistical fluctuations leads only to some washing off of the sharp transition." (Mikhailov 1979, XX) Furthermore, and most interestingly, external noise was shown to be able to induce phase transitions: new stable states are created in the system states by applying external noise. Mikhailov came to these results by studying the classical Lotka-Volterra model to which he added a fluctuating food source representing the external noise. In this model two species are competing over the same type of food. One of the species denoted by N is stronger one, whereas the other one denoted by n is the weaker species. N and n represent the density of the respective species. The latter one is also able to move around and by doing so introduces a diffusion term to the system given by $D\Delta n$. The model is represented by the following differential equations:

⁷ Bifurcations in this context mean changes in the qualitative or topological structure in the solutions of differential equations.

$$\frac{dN}{dt} = (BM - A)N$$

$$\frac{dn}{dt} = (bM - a)n + D\Delta n$$

$$\frac{dM}{dt} = Q - GM - CN - cn + f(r,t).$$

The first two equations represent the rates of the increase of the populations. They are linearly dependent on the amount of food and negative when food is absent. The second equation describing the dynamic of the second species has an additional diffusion term $D\Delta n$ taking into account the mobility of this species. This makes the system a system with diffusion. The third equation describes the dynamic in the food density. Both species consume in a given unit of time an amount of food C and c , respectively. The food grows with a constant rate Q and even in the absence of both species the maximal food density is limited by a decay mechanism represented by the term $-GM$. By introducing a Gaussian random force $f(r,t)$ spatial and temporal fluctuations in the rate of the food growth are taken into account. This last term represents the external noise applied to the system. What is the effect of this additional noise source?

As Mikhailov shows, in the classical case without noise the system has only one steady state: competition leads to the extinction of the weaker species. In contrast, fluctuations in the system caused by the Gaussian random force makes the coexistence of the two competing species possible. New steady states are created meaning that new breeding regions can be created by fluctuations. This is a very important result. It shows that through external noise new states of the system can be created which then can be taken by the system.

Mikhailov sees the importance of his results for the study of other systems in other disciplines. He writes: "The cross-disciplinary significance of these results should be mentioned. Similar noise-induced phase transitions can occur in systems with chemical reactions and in physical systems with diffusion; they can be also of interest in discussions of genetic versatility." (ibid, p. 144) The results of the special

model have thus a more general character applicable to other complex systems. Mikhailov's statement is based on the assumption that systems belonging to this special class of complex systems share specific structural and dynamic elements, such as non-linear dynamics caused by feedback loops. Those elements are independent from the particularities of the respective system but are considered to be the basis for observing the functional component of noise and thereby, as some scientist argue, bring such models closer to the observed behavior in real systems. As our example of the *Repressilator* shows, exploring this functional component of noise turns out to be a rather intricate enterprise where among other things the specific *materiality* of synthetic models plays an important role.

4.2 Internal noise and its possible effects on the dynamics of complex systems

Scientists have been observing variations in the behavior of populations of cells since the 1940's . Max Delbrueck, for example, observed such variations in bacteriophage burst sizes (Delbrueck, 1945), E. O. Powell in cell division lengths (Powell, 1958), and Novick and Weiner in the concentration of β -galactosidase in the *lac* regulatory genetic network in *E. coli* (Novick & Weiner, 1957). These variations in the behavior of single cells gave rise to the question of what was causing these seemingly non-genetic variations. Spudich and Koshland (1976) explored this question in series of elegant experiments on the chemotactic behavior of a population of genetically identical *E. coli* bacteria in a homogenous environment.

Spudich and Koshland started their article by stating the importance of the question of the existence or non-existence of non-genetic variations in the following way: " Biological systems are constantly confronted with chance occurrences in their environmental conditions and internal processes. An elaborate biological apparatus has evolved to utilize chance in genetics for the survival of the species. Some feedback mechanisms, on the other hand, are designed to insulate the individual against chance fluctuations of environmental conditions. How much influence has chance in internal cellular processes in determining the metabolic state and behavior of the single cell? More specifically, if cells containing identical

chromosomes conditions, how similar would they be in their biochemical characteristics?" (Spudich and Koshland 1976, p. 467)

What Spudich and Koshland describe as chance in internal cellular processes is in the present literature on noise usually described as internal noise. However, it seems that there was not much overlap between the community of scientists studying the effect of external noise on complex systems such as Mikhailov and scientists focusing on internal noise such as Spudich and Koshland. Why this was the case needs to be studied in more detail, but it seems that the research on stochastic fluctuations was concentrated on non-genetic variations and therefore strongly linked to the research in biology, instead of physics.

In their experiments Spudich and Koshland studied the non-genetic individuality of cells by observing the chemotactic migration of bacteria. Bacteria such as *E. coli* make use of a unique system for sensing food sources in their environment. They alternate between a tumbling and a swimming phase. In the swimming phase, lasting only for a few seconds, the bacterium is moving straight, then it tumbles and changes direction. During these different movements it is sensing whether the conditions are getting better, meaning whether it is moving in direction of a food source, or worse, meaning it is moving to a source that could harm it. When the bacterium is sensing a food source it will start moving in direction of the food source and tumble less. This behavior is called chemotaxis. The movements are produced by flagella surrounding the bacteria and moving either counter-clockwise (swimming in straight line) or clock wise (tumbling). Spudich and Koshland investigated the non-genetic variability by using a population of genetically identical bacteria whose environmental conditions were changed by adding a chemical that would provide a food source for the bacteria. Sensing the chemical, the bacteria changed their swimming and tumbling behavior following the food gradient and moved in the direction of the food source. The exact mechanism by which the bacteria was changing its tumbling behavior as they sensed a food source was unknown but as the two authors proposed: "[...] can be rationalized as caused by changes in the levels of a tumble regulator [...]" (Spudich and Koshland 1976, p.467)

Spudich and Koshland observed the movement of individual bacteria by following the bacteria under the microscope and videotaped their behavior. They studied the following two situations:

1. The sensitivity to stimuli as measured by the recovery from temporal gradients.
2. The tumbling frequency when there is no gradient.

They found that individual cells reacted differently. Fig. (4) shows the reaction of 22 bacteria to a chemotactic stimulus over a time interval of 3.5 min. The lines mark brief tumbles of the bacteria and the x when the bacteria started to tumble again continuously. The figure shows clearly the heterogeneity of the behavior of different individuals. Spudich and Koshland explained this behavior by fluctuations in the tumble regulator changing the motion of the bacteria according to the changes in the environment.

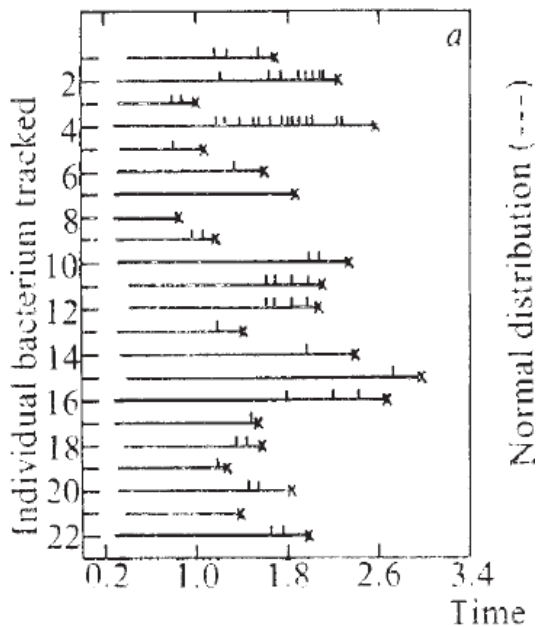


Fig.(4) The diagram showing the variation in the reaction of bacteria to chemotactic stimuli measured by Spudich and Koshland (1976)

But what was causing the fluctuations in the tumble regulator? Or as Spudich and Koshland formulated the question: "How could such individuality be produced?" (ibid, p. 470) They speculated that such individuality might be due to Poissonian variations in molecules present at low numbers in the cell. The variance of a Poisson distribution is equal to its mean. The standard deviation given by the square root of the variance becomes relatively smaller for larger numbers of molecules and larger for a small number of molecules. For example if there are 10^4 molecules of a special type, the standard deviation would be 100 molecules, giving a 1% deviation. If there are only 100 molecules, a 10% standard deviation would arise. At the end of their article the two authors asked the question: What could be the advantage of non-genetic variations for biological systems? They reason that non-genetic variations as observed in their experiment would provide a mechanism to react to random fluctuations in the environment: "Thus non-genetic variability would be a preferred mechanism for accommodation to random fluctuations in the environment and genetic variability the preferred mechanism for accommodation to long lasting environmental changes." (Spudich and Koshland 1976, p. 470). Here we already observe a functional explanation for the non-genetic variation but at this point it still remained a speculation.

The importance of the work of Spudich and Koshland lies, as M. D. Levin pointed out, in making possible the modeling of the possible origins of the observed variations (Levin, 2003, p. 135). Spudich and Koshland's work was indeed followed by a number of articles presenting theoretical models, which take into account the experimental results on the variations in the swimming behavior of E.coli. These articles discussed questions such as whether the variations in gene expression follow a Poisson distribution or not, or whether stochastic processes at the level of gene expression are linked to, for example, cell differentiation processes. Minoru S.H. Ko (Ko, 1991) introduced a model on stochastic gene induction with which he aimed to explain the experimental results showing that expression levels of individual inducible genes are very heterogeneous and not, as presumed before, identical to the averaged level of many copies. But he was very careful in relating these

stochastic processes on the level of gene expression to the observed stochastic manner in which cells in cell differentiation make commitment decision. Accordingly, Ko indicates an essential point in the development of the functional meaning of noise: Stochastic fluctuations in gene expression due to the low number of molecules in the cell do not link in any straightforward way to such processes as cell differentiation or the observed individuality in the behavior of cells in chemotaxis. Behind chemotaxis lies a complex signaling, perception and response mechanism. It is essential to identify the underlying mechanisms—their architecture and dynamics—which determine functions like chemotaxis or the effect of stochastic fluctuations in gene expression, or of other sources of noise on such mechanisms. As we have seen, novel imaging methods as well as synthetic models are helpful in this enterprise.

5. Discussion: Analogies, templates, and concepts

We have seen that the research on genetic regulatory networks made various uses of analogies. A basic positive analogy was drawn to electric circuits, making use of the negative feedback control mechanism in “genetic circuits”. However, the mathematical templates used to model the dynamics of genetic circuits were taken from other sources, notably from the study of complex systems. As we have shown, the progress and the direction of the research program was largely due to the negative analogies drawn between artificial and biological control systems. Firstly, already Goodwin (1963) suggested that oscillations (conceptualized as noise in artificial systems) in fact provide the means by which biological systems regulate themselves. A further negative analogy as regards to the role of noise in artificial vis-à-vis biological systems was drawn as a result of synthetic modeling. Although the mathematical model that was used as the basis of the construction of the *Repressilator* exhibited regular oscillations there remained the possibility that nature did it in a different way, which prompted the construction of the *Repressilator*. The observed irregular oscillations of the *Repressilator* made noise a pressing problem and an object of investigation. In principle there were two ways of

dealing with the observed noise, indicative of the differences between the application oriented and basic oriented branches of synthetic biology: Firstly, one could pursue the positive analogy between artificial and biological systems by treating the fluctuations as a disturbance and trying to find ways of making the system more robust by changing its architecture. This line is chosen by the application oriented branch of synthetic biology, which uses different strategies to isolate and get rid of different sources of noise. Such approach is justified by the aim of the application oriented branch of synthetic biology to engineer biological components for specific purposes: for vaccines, biofuels and cancer treatments.⁸

The basic science oriented branch of synthetic biology has chosen the opposite direction in drawing a further negative analogy to artificial control systems: Recognizing noise as an intrinsic part of biological systems researchers in the field have started to study the sources and impact of noise on biological systems. As a result of these studies noise has also been allotted a functional role in supporting biological systems.

The process in which noise gained a functional meaning in synthetic biology is closely linked to the transfer of knowledge, methods, and tools from the research on complex systems. In the Elowitz lab for example, via the cooperation with the physicist Jordi Garcia Ojalvo such knowledge, methods, and tools were imported into synthetic biology. Ojalvo came to the lab when research on the mechanism of the competence in *B. subtilis* started. He recognized parallels in the behavior of *B. subtilis* to excitable systems—lasers and neurons—which he had studied intensively in the context of complex systems (Lindner, Garcia-Ojalvo, Neiman, Schimansky-Geier, 2004).

It appears to us that paying attention to the cognitive importance of negative analogies makes visible the heterogeneous nature of analogical reasoning, which the traditional approach to analogical transfer largely neglects. Any negative analogies, by themselves, do not carry research much further, which might be the reason why so little attention has been given to them in the literature on analogical reasoning

⁸ This approach of the application oriented branch of synthetic biology can often be justified by the fact that on the population level the effect of stochastic fluctuations averages out.

(see however Morgan 1997). The situation becomes different if one enlarges the unit of analysis to cover other domains and bodies of work that could give negative analogies a tentative interpretation and point towards further study. In the case of the notion of noise, the noisy oscillations of the *Repressilator* made relevant the earlier studies on external and internal noise, which explains also in part why researchers in the field of synthetic biology were so quick to pick the idea of the functional meaning of noise. As we have shown, despite the reigning engineering connotations of the concept of noise, the new functional understanding of it had already emerged in physics from where synthetic biologists have adopted the majority of their modeling methods. Especially the work in statistical physics provided a well-understood concept of stochastic fluctuations and associated computational templates, such as Poisson variations, for the study of the non-deterministic fluctuations observed in biological systems (cf. Ozbudak, Thattai, Kurtser, Grossman, & van Oudenaarden, 2002). On the other hand, the functional meaning of noise had also already emerged in biology for instance in the studies of non-genetic variations although biologists did not label these phenomena noise.

Consequently, the notion of noise in the field of synthetic biology subsumes heterogeneous interdisciplinary influences between different fields of inquiry. In making use of analogies scientists do not merely focus on the similarities (or differences) between two domains as the mapping-accounts suggest but through drawing analogies they mobilize whole research fields including their theoretical results, tools, methods, templates and concepts. The modeling practice – which in the case of synthetic biology is particularly subtle, combining mathematical models, model organisms, and synthetic models - weaves all these heterogeneous interdisciplinary influences together.

Finally, our case shows that the mapping-account does neither provide an adequate treatment of mathematical representation. Namely, more often than not it is assumed that the formal structure used to describe *one domain* is transferred to another domain. For instance, Nersessian (2001) claims of Maxwell that he abstracted from specific mechanical models the dynamical properties and relations that continuum-mechanical systems, certain machine mechanisms and

electromagnetic systems had in common. However, in contrast to such abstraction account it seems that modelers typically make use of familiar computational templates in the sense of Humphreys (2002, 2004) that are then locally adapted to various subject matters.

The models that feature in our account involve several instances of such template use. In designing mathematical model underlying the *Repressilator* Elowitz and Leibler were seeking network architectures leading to stable limit-cycle oscillations. Such oscillations are self-sustained meaning that after small perturbations the system will move back to the limit-cycle trajectory. Systems showing such limited cycle behavior have been studied and analyzed in fields such as electrical engineering, biology, economics, and chemistry. One of the earliest examples is the van der Pol relaxation oscillator, which was found and analyzed by the Dutch engineer Balthazar van der Pol during the 1920's. Also Brian Goodwin explored different architectures. The Jacob-Monod operon model of gene regulation provided Goodwin the basic template for constructing his simple feedback network through which he studied the conditions under which sustained oscillations become possible. At the time when Elowitz and Leibler set out to design the *Repressilator* a large body of research on possible network architectures and oscillatory dynamics was already available. The design of the *Repressilator* is linked to the Goodwin model in that both models make use of feedback mechanism and only differ in the number of genes. Goodwin made use of one gene repressing its own function and Elowitz and Leibler used 3. They wrote: "The simplest such networks supporting limit-cycle oscillations are those containing a single repressor and a single activator, or an odd number of repressor exceeding 3." (2000, p. XX). Clearly, what is happening here is not adequately described in terms of *abstracting* a template from one domain and then applying it to another domain. The templates used here are largely subject-independent formal systems applicable across disciplines to wide range of problems.

Of course such templates can have an origin in a certain model describing a particular system. For instance, the Lotka-Volterra model, which Mikhailov used to study the effect of external noise was presented by Volterra as a model of a

particular predator-prey-system (although Lotka actually made use of template-oriented approach in his formulation of the model, see Knuuttila & Loettgers, in press). Since then the model has also been detached from its origin in fisheries and it has started to serve scientists and mathematicians in the field of dynamical systems theory as a popular model for studying non-linear dynamics. This template-oriented interest in the Lotka-Volterra model and its dynamic behavior was sparked by the advancements of computer technologies in the 1970's. Another kind of template is provided by Poisson variations, which Spudich and Koshland used to explore whether non-genetic individuality, the presence of a diversity of phenotypes in a genetically identical population, could be explained by Poissonian variations in the molecules present in a low number in a cell. Poisson variations differ from such a template as the Lotka-Volterra model which provides a *model system* for the exploration of certain kinds of dynamic behaviors and model specific behaviors. Poissonian fluctuations offer a mathematical *tool* for the analysis of intrinsic fluctuations in biological systems. It allows for example for an isolated analysis of the contribution of different sources of fluctuations such as fluctuations in the transcription and translation in protein production. (Ozbudak, Thattai, Kurtser, Grossman, & van Oudenaarden, 2002)

Yet despite the frequent use of computational templates in modeling the analogies between different domains are not entirely formal. Their use is motivated by theoretical considerations. In practice, as also the case of modeling genetic regulatory mechanisms shows, the templates are often introduced hand in hand with analogies and theoretical concepts. Hence it seems to us that Hesse is right in hinting that material analogies to systems like electric circuits and scientific concepts like noise are often needed to make templates tangible and to give them provisional theoretical content. And these analogies and notions are more bound to subject specific knowledge in certain scientific fields than the mere formalisms used. Therefore we suggest that a viable account of interdisciplinary exchange should combine insights concerning analogical reasoning with the study on how formal templates are used in scientific modeling.

6. Why “noise”?

The notion of noise has changed a lot since Shannon and Weaver’s (1963) notion of noise as an *undesirable uncertainty* introducing spurious information mixing with useful information.⁹ This useful information was to be cleaned up from the undesired wrong information or the source of noise was to be eliminated. We have shown how an entirely different approach to noise emerged in synthetic biology during the last decade discussing also its origins in different fields of study. In the research on complex systems in the 1970’s the notion of noise was mostly used in the studies on the effect of external sources of noise on complex systems. But only with the introduction of synthetic modeling and novel imaging methods did noise begin to cover all the unknown and known forms of fluctuations. These new methods made visible the non-genetic fluctuations on the level of single cells (Elowitz 2002). Such fluctuations had remained “invisible” in the previously studied population level where they typically average out. As the noisy behavior became observable on the molecular level, the earlier discussions on the function of non-genetic fluctuations in biological systems revived. New methods made visible a large number of before unrecognized/non-visible non-deterministic fluctuations and as a result researchers started to realize the complexity of the question of noise.

But why should the researchers continue talking about noise when referring to these non-deterministic fluctuations and their possible functional roles. Why was the notion of noise extended in such a way? This certainly reflects partly the influence of engineering sciences on biology and the background of many synthetic biologists in the research on complex systems. However, there seems to be also another more profound reason, which is related to the use of borrowed concepts and modeling methods. Namely, the application of engineering notions and modeling methods of physics to biology by way of analogical reasoning is not unproblematic. The sources of the fluctuations in biological organisms are largely

⁹ There is a long tradition in molecular biology of making use of metaphors and concepts from information theory (cf. Kay, 2000), but synthetic biology is taking it a step further by basing an *engineering approach* on these metaphors and concepts.

unknown in all but a few cases, as is their exact impact on the dynamics of biological systems. One reason for the use of the notion of noise then, we suggest, is exactly this uncertainty: noise functions both as an umbrella term and as a place holder for the emerging research on different forms of fluctuations, their sources and consequences for the dynamics of biological systems.

We have argued above that scientific concepts are often associated with specific modeling methods and computational templates. Yet, as the case of noise shows, they can also undergo semantic transformations and subsume new kinds of research agendas employing novel modeling tools.

Bibliography

- Bailer-Jones, D. (2009). *Scientific Models in Philosophy of Science*. Pittsburgh: University of Pittsburgh Press.
- Baker, D., Church, G., Collins, J., Endy, D., Jacobson, J., Keasling, J., Modrich, M., Smolke, C., and Weiss, R. (2006). Engineering life: building a FAB for biology. *Scientific American* June: 44-51.
- Barkai, N., & Leibler, S. (2000). Circadian clocks limited by noise. *Nature*, 403, 267-268.
- Cagatay, T., Turcotte, M., Elowitz, M., Gracia-Ojalvo, J., & Suel, G. (2009). Architecture-dependent noise discriminates functionally analogous differentiation circuits. *Cell*, 139, 1-11.
- Eldar, A., & Elowitz, M. (2010). Functional roles for noise in genetic circuits. *Nature*, 467, 167-173.
- Eldar, A., Chary, V., Xenopoulos, P., Fontes, M., Loson, O., Dworkin, J., et al. (2009). Partial penetrance facilitates developmental evolution in bacteria. *Nature*, 460 (7254), 510-514.
- Elowitz M.B., L. A. (2002). Stochastic Gene Expression in a Single Cell. *Science*, 297 (5584), 1183-1186.
- Elowitz, B. Leibler S (2000). A Synthetic Oscillatory Network of Transcriptional Translation. *Nature*, 403 (6767), 335-338.
- Delbrueck, M. (1945). The Burst Size Distribution in the Growth of bacterial Viruses (Bacteriophages). *J. Bact.*, 50, 131-135.
- Endy, D. (2005). Foundations for engineering biology. *Nature*, 438, 449-453.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Genter, Derde & Markman, Arthur B. (1997). Structure Mapping in Analogy and Similarity. *American Psychologist* 52(1): 45-56.
- Goodwin, B. (1963). *Temporal Organization in Cells*. London : Academic Press.

- Israel, G. (1993). The emergence of biomathematics and the case of population dynamics: A revival of mechanical reductionism and Darwinism. *Science in Context*, 6, 469-509.
- Humphreys, P. (2002). Computational Models. *Philosophy of Science*, 69, S1-S11.
- Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.
- Harre, R. (1970). *The principles of scientific thinking*. Chicago: University of Chicago.
- Hesse, M. (1966). *Models and Analogies in Science*. Notre Dame: Notre Dame University Press.
- Hesse, Marry (2001). Models and Analogies. In W.H. Newton Smith, ed. *A Companion to the Philosophy of Science*, Oxford: Blackwell, 200, 299-307.
- Kay, Lily E. (2000). *Who wrote the Book of Life?* Stanford: Stanford University Press.
- Knuuttila, Tarja (in press), Modeling and Representing: An Artefactual Approach. To appear in *Studies in History and Philosophy of Science*.
doi:10.1016/j.shpsa.2010.11.034
- Knuuttila, Tarja and Andrea Loettgers (in press). The Productive Tension: Mechanisms vs. Templates in Modeling the Phenomena. To appear in Paul Humphreys and Cyrille Imbert (eds.), *Representations, Models, and Simulations*. Routledge.
- Levin, M.D. (2003). Noise in gene expression as the source of non-genetic individuality in the chemotactic response of *Escherichia coli*. *FEBS Letters* 550, 135-138.
- Lim, W. A. (2010). Designing customized cell signalling circuits. *Nature Review Molecular Cell Biology*, 11, 368-372.
- Lindner, B., Garcia-Ojalvo, J., Neiman, A., Schimansky-Geier, L., (2004) Effects of noise on excitable systems. *Physics Reports* 392, 321.
- Loettgers, A. "Model organisms, mathematical models, and synthetic models in exploring gene regulating mechanism." *Biological Theory*, 2, (2007), 143-142.
- Loettgers, A. "Synthetic biology and the emergence of a dual meaning of noise." *Biological Theory*, 4, (2009), 340.
- Nersessian, N. (2002). Maxwell and "the Method of Physical Analogy": Model-based reasoning, generic abstraction, and conceptual change. (D. Malament, Ed.) *Essays in the History and Philosophy of Science and Mathematics*, pp. 129-166.
- Novick, A., & Weiner, M. (1957). Enzyme induction as an all-or-none phenomenon. *Proc. natn. Acad. Sci. USA*, 43, 553-566.
- Mikhailov, A. S. (1979). Noise-Induced Phase Transition In A Biological System With Diffusion. *Physics Letter*, 73A (2), 143.
- Morgan, M. (1997). The technology of analogical models: Irving Fisher's monetary worlds. *Philosophy of science*, 64.
- Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., & van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nature Genetics*, 31, 69.
- Paldi, A. (2003). Stochastic gene expression during cell differentiation: order from disorder? *Cellular and Molecular Life Sciences*, 60, 1775-1178.
- Powell, E. O. (1958). Criteria for the growth of contaminants and mutants in continuous culture. *J. gen. Microbiol.*, 18, 382-417.

- Purnick, P. E. M., Weiss, R. (2009). The second wave of synthetic biology from modules to systems. *Nature Reviews Molecular Cell Biology*, 10, 410-422.
- Suel, G. G.-O., Garcia-Ojalvo, J., Liberman, L., Elowitz, M. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. . *Nature*, 23 (440), 545-550.
- Spudich, J. L., Koshland, D. E. (1976). Non-genetic individuality: chance in the single cell. *Nature*, 262, 467-471.
- Thattai, M. Oudenaarden, A v. (2001). Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci USA*, 98, 8614-19.
- Weaver, W., & Shannon, C. E. (1963). *A Mathematical Theory of Communication*. Univeristy of Illinois Press.