Measurement in biology is methodized by theory

Maël Montévil

Noname manuscript No.
(will be inserted by the editor)

Measurement in biology is methodized by theory*

Maël Montévil

the date of receipt and acceptance should be inserted later

Abstract We characterize the access to empirical objects in biology from a theoretical and philosophical perspective. Unlike physical objects, biological objects are the result of a history and their variations continue to generate a history. This property is the starting point of our notion of measurement. We argue that biological measurement is relative to a natural history which is shared by the different objects subjected to the measurement and is more or less constrained by biologists. We call symmetrization the theoretical and often concrete operation that leads to consider biological objects as equivalent in a measurement. Last, we use our notion of measurement to analyze research strategies. Some strategies aim to bring biology closer to physics, by studying objects as similar as possible, while others build on biological diversification.

*Keywords:*Biological measurement, experiments, evolution, systematics, strains, symmetry

1 Introduction

Science and more specifically biology and medicine are facing a crisis because systematic attempts to reproduce experiments published in reputable journals fail in the majority of cases (Begley and Ellis 2012; Baker 2016). Much has been said on the issue, especially about the management and organization of scientific institutions. For example, the pressure to publish does not foster good practices (Begley and Ioannidis 2014; Lancet 2018), and new technologies can enhance the exchange of experimental protocols (Teytelman 2018).

In the case of experimental biology, theoretical and philosophical analyses can play a role to respond to this crisis (Nadin 2017). There are aspects proper to biological experiments that should be analyzed systematically in the light of the current understanding of living beings. This discussion is particularly relevant now that the scientific focus on (big) data analysis bears the risk of forgetting that data are obtained in specific empirical conditions (Leonelli 2014). Data detached from these conditions do not carry a genuine scientific meaning.

A scientist cannot assume that her access to reality is the one of an omniscient daemon. Understanding what it means to observe natural phenomena is fundamental. This question is not just technical, and part of the answer should be principled, on the basis of a theory. As Einstein puts it: "whether you can observe a thing or not depends on the theory which you use. It is the theory

* Paper submitted to Biology & Philosophy.

Address(es) of author(s) should be given

which decides what can be observed" (A. Einstein quoted in Salam 1990). In physics, measurement is described in theories and is a fundamental part of their formulation. The theoretical notion of measurement provides the link between the output of a measurement and the theoretical and mathematical description of the object of study. For example, measurements in classical mechanics provide approximate results. The notion of measurement is more basic, general and elementary than the notion of experiments.

Biologists often use physical concepts, and measurement is no exception. The notion of measurement in classical mechanics is widely used in biology, sometimes implicitly. This situation leads us to ask whether biology requires a specific notion of measurement. In the literature, there is at least one such account: following the informational metaphor, molecular biology generally considers measurement as a classical measurement applied to *finite, discrete* structures: the sequences of nucleotides (Crick 1958). Then, the measurement can *in principle* achieve "perfection" (Turing 1950). Classical measurement has a limited precision, but knowing finite, discrete structures with a sufficient *finite* precision means knowing them *exactly*. The discrete topology transforms the notion that measurement is always approximate to a measurement that can be perfect. If we do not introduce other considerations, the same reasoning applies *mutadis mutandis* to other discrete structures such as the topology of networks which are increasingly important in biology (Huneman 2018).

This point of view is in contrast with experimental methodologies which are very rich and sometimes subtle (Weber 2004; Kohler 1994). We will argue in this paper that at least some aspects of experimental methodologies in biology are not just practical, but pertain to *principles*. The biometrician Bookstein (2009) contrasts measuring stable elements and contingent ones: "the biometrician is like the pilgrim in Friedrich's painting Der Wanderer über dem Nebelmeer, uncertain as to whether to measure the mountains or the clouds. The mountains stand for contingent history, the clouds for the subset of the data most closely matching controlled experiments suitable for quantitative biometric summary. Biometrics applies to the clouds, not the mountains". The problem thickens when we consider ideas stemming from evolution, such as the contingency thesis of Beatty (1995). According to this thesis, biological regularities ultimately stem from a contingent history, which implies that the separation between contingent history and stable properties described by Bookstein is not absolute.

We will use the principles proposed recently for a theory of organisms (Mossio et al. 2016; Montévil et al. 2016; Soto et al. 2016a) in order to analyze biological measurement. This framework provides a conceptual continuity between the understanding of organisms and evolution. In particular, our approach emphasizes historical analyses.

In this framework, biological objects cannot be defined theoretically like physical objects. The theoretical definition of physical objects is mathematical. Despite quantitative differences, the transformations of a well-defined object are assumed to follow an underlying mathematical structure. Such structures are typically defined by invariants and invariant preserving transformations, also called symmetries (Longo and Montévil 2014). For example, a falling stone follows the same equation during its fall despite its changes of position and velocity, and a falling log would follow the same equation. As a result, physicists can talk about the generic phenomenon of falling bodies. Physical notions of measurement apply to generic objects, and the reproducibility of physical experiments is guaranteed, at least statistically, once the same generic conditions apply.

By contrast, biological objects are historical in the sense that their organizations stem from an evolutionary and individual history and continue to produce a history. We have developed this idea theoretically and called it the principle of variation (Montévil et al. 2016). To an extent, this principle is in line with earlier ideas, in particular, the contingency thesis of Beatty (1995) and the centrality of historicity defended by Gould (2002, chap. 11). For example, a falling tetrapod is not a purely physical notion since "tetrapod" is a biological concept. In the atmosphere, tetrapods do not just fall, some fly and others are gliders. All these behaviors require different equations, and these changes of equation depend on the underlying evolutionary history. This basic example

illustrates the general idea that biological objects should not be conceived as generic and are prone to deeper changes than physical objects, including the appearance of new possibilities (Montévil 2018). Moreover, biological objects are contextual in the sense that their organizations depend on their past and current contexts.

In a nutshell, biologists manipulate objects that are the result of a history and continue to produce a history over time: diachronic objects. With these ideas stemming from the theory of evolution in mind, experimental reproducibility is not straightforward. Biological objects tend spontaneously to vary whereas perfect reproducibility would require fixed physiology and development. The main idea that we will defend is then that biological measurement is relative to a history.

In section 2, we introduce how several physical theories define measurement and the epistemological and theoretical roles this notion plays. Section 3 discusses the theoretical nature of biological measurement. Biological measurements describe a way to handle natural histories and contexts, not just quantities. Section 4 explores several ramifications of our framework. In particular, we classify different research strategies to handle the specificities of biological measurement.

2 Physical measurement

In order to introduce the question of measurement, we will discuss a few examples of how physical theories conceptualize it. By measurement in physics, we mean "obtaining quantities" in experiments or observations, and the theoretical descriptions of this process. These descriptions capture the status of measurement *in principle* and not just specific experimental situations. Physical notions of measurement are sufficiently general to be valid for *any* practical situation in the corresponding theory. Despite their general nature, they have practical and theoretical consequences.

2.1 Classical measurement

In classical mechanics, a system has a pointwise state in the space of possible states. The empirical access to this state is approximate: measurements have a finite precision, ϵ, which can *in principle* be arbitrarily small. Thus, the state of a system is a point, and the result of the measurement is an interval. Classical measurement is a *metrical* notion: it is based on the notion of distance.

Classical dynamics are deterministic, but depending on the *dynamics*, measurements may or may not allow to predict the subsequent trajectory with a reasonable approximation. Unpredictable dynamics such as chaotic dynamics are called sensitive to initial conditions. The notion of measurement articulates determinism and randomness in the sense of theoretical impredictability (Gillies 2012; Longo and Montévil 2017). This example shows that a simple notion of measurement can have far-reaching consequences.

2.2 Quantum measurement

In quantum mechanics, measurement has a highly original and fascinating role because it changes the object and leads to quantum randomness. Informally, a quantum state can be decomposed *for a given measurement* as the superposition (the sum) of different states called eigenstates. Each of them corresponds to a single obtainable result. Performing the measurement means that the state of the system becomes an eigenstate associated with the quantity obtained. The other eigenstates in the initial superposition disappear. Quantum measurement has an *algebraic* (and geometric) nature.

There is an internal consistency to this notion. Performing the same measurement twice in a row will lead to the same result because the state of the object is already an eigenstate associated with this result. For this specific measurement, there is only one possible outcome: the one already

obtained. Another aspect of quantum measurement is that it is an *irreversible* process: the initial state collapses to a single eigenstate, and the other components of the initial state disappear irremediably. Another critical point is that the decomposition is *not* necessarily the same for different observables. An eigenstate that corresponds to a specific position, for example, does not correspond to a specific velocity and the other way around. Then, measuring the position, measuring the velocity and then measuring the position again will not necessarily lead to the same positions twice. Lastly, some authors argue that, in an experiment, a measurement is needed to put the system in a known initial state (Mugur-Schächter 2002). The typical theoretical structure of an experiment is then: measurement, time evolution (Schrödinger equation), measurement.

2.3 Reference frame

Experimenters choose space-time reference frames arbitrarily to represent concrete situations and describe features such as position quantitatively. Relativity (Galilean, special and general) states how the description of a situation in one reference frame can be transformed into the description of the same situation in another reference frame and ensures that they are coherent. Thus, relativity overcomes the arbitrary choices of reference frames.

2.4 Conclusion

The concepts of physical measurement we described are very diverse: measurement understood as imprecision, as changing the state of the object and as dependence on arbitrary reference frames. Their common point is that they all describe the role of the experimenter and its instruments in an abstract and very concise way. These notions of measurement facilitate the reproductibility of experiments by making these roles explicit, including the arbitrary choices of the experimenter and the alteration of objects in quantum measurement.

3 A theoretical account of biological measurement

To describe our theoretical notion of biological measurement, we will rely mainly on the principle of variation (Montévil et al. 2016). This principle builds on evolutionary biology and states that biological objects can vary in a stronger sense than physical objects. Physical objects change, but physicists understand their changes by underlying stable mathematical structures. Instead, biological variations in the strong sense require changing mathematical structures. Biological objects are generated by a cascade of such variations which leads to a fundamental notion of historicity and of contextuality, corresponding to the variations associated with the context. Figure 1 summarizes this perspective which guides our analysis of biological measurement.

Physical objects can be highly simplified and remain physical objects. For example, it is sound to study a material composed only of iron. In biology, this is not the case. For example, looking at one or several molecules alone pertains to biochemistry, not biology. In biology, the features of organisms or cells measured, such as the concentration of molecules or the shape of tissues, are measured in organisms or cells and are generally produced and maintained by them (Mossio et al. 2016; Montévil and Mossio 2015). Therefore, our discussion of biological measurement is not limited to the parts observed *per se*. Instead, our approach of measurement accommodates both the parts observed and the organisms associated. Both are elementary components of biological measurement.

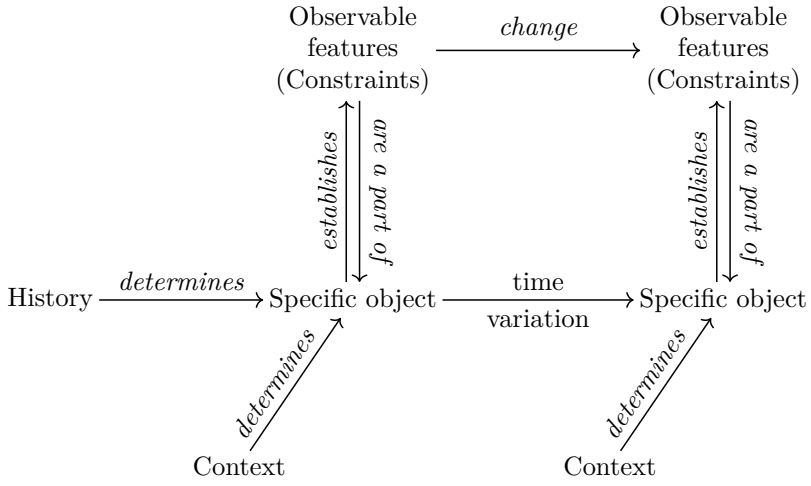

Figure 1 *Theoretical structure of biological objects.* In biology, an organism is not a generic object that follows invariants and invariant preserving transformations (symmetries). Instead, their regularities are constraints that come from an history and that collectively maintain each other in a given context. Moreover, these constraints can change over time as objects continue to generate an history over physiologic, developmental and evolutive time scales.

3.1 Phylogenetic classification and nomenclature of biological objects

Reporting a biological measurement starts with describing the organisms observed and naming them. The standard, general way to name organisms is to use the results of systematics. Biologists always use this method, even though other elements can be used as discussed in the following sections.

We want to emphasize two aspects of this method that impact the concept of measurement. The first is the definition of the names themselves and the second is the phylogenetic classification of living beings (de Queiroz 1992; Lecointre and Le Guyader 2006).

In order to provide stability to the meaning of the names used to describe living beings, systematics establish and follow strict rules to describe new species and other clades (e.g., genus, family). Nomenclature codes use the principle of typification (1999; 2012). Typification means that defining a name requires a type. For example, the definition of a name at the family level requires a genus-level type, and a genus-level name requires a species type. Last but not least, describing a new species (or subspecies) requires referring to one specimen (holotype) or several specimens (syntype) which are kept in a collection[1] (CZN International 1999; McNeill et al. 2012, art. 72.3 and 40 resp.). Name-bearing types are required to be in a biologically inactive state (McNeill et al. 2012, art. 8.4). Typification implies that the definition of biological names ultimately depends on specific, static, material objects. This situation should be contrasted with the definition of physical units which is theoretical in the International System of Unit. For example, a meter is the distance traveled by light in vacuum in 1/299792458 seconds. This definition refers to matter but does not need the conservation of a specific object. Instead, it uses the generic, theoretical object called "light in the vacuum" which has an invariant velocity according to both special and general relativity[2].

Names associated to specific material objects (types) are not sufficient for scientific practices. In order to endow names with a more general meaning, systematics uses the phylogenetic classification

[1] Exceptions exist when this is not possible. Then, an illustration can be used.

[2] Historically, the definition of a meter has first been theoretical, then it used a standard prototype. Last the current definition is again theoretical.

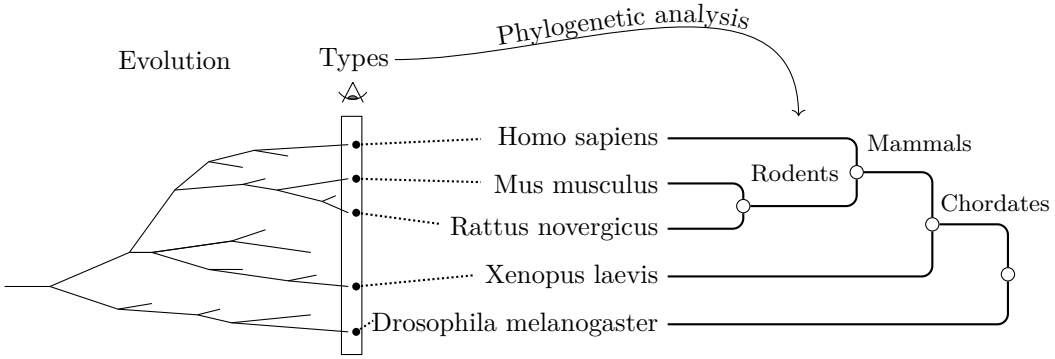

Figure 2 *Principle of the phylogenetic classification.* Left : a schematic representation of the genealogy of a few species over evolutionary timescales. This genealogy is not accessible as such. Middle: the consequence of evolution is the presence of diverse life forms, some of which are used by biologists as types that are kept in collections and define names. Right: the characters that the specimen share and do not share are the basis to assess their evolutionary proximity. In order to simplify the representation, we did not include fossil name-bearing types.

method, also called cladistic (de Queiroz 1992; Lecointre and Le Guyader 2006). The principle of this method is to classify living beings by estimating their genealogy. The genealogy is a theoretical concept that stems from the theory of evolution; however, the genealogy of current organisms spans billions of years, and human observers cannot access it. As a result, the phylogenetic classification uses different concepts than a genealogical tree. For example, it is impossible to determine whether a fossil species is an ancestor of a current species, but it is possible to establish that they are closely related. The phylogenetic method distinguishes a theoretical level and an observable level which is reminiscent of the distinction between a state and what can be observed in physics.

The phylogenetic classification assesses the evolutionary proximity between different organisms. Systematists start with the characters characterizing the different organisms. These characters are used by a computational method which provides a nested hierarchy of groups, see figure 2. These methods typically assume that the more likely situation minimizes the number of evolutionary changes, and in particular the appearance of novelties. This analysis leads to a classification. In the classification, acceptable groups are the descent of a common theoretical ancestor and are called monophyletic groups or clades. Proposing a classification leads to a critical discussion of its evolutionary implications. For example, a classification implies that some characters are analogies instead of homologies[3]. Last, the classification can be used for taxonomic purposes. Of course, evolutionary considerations guide the choice of the characters and the computational method used, and these choices are commonly debated.

The names provided by systematics rely on a historical analysis. Clades are defined by their origin and not by their ecological status or their current physiology. For the classification, specimens are described by the characters that they possess and not by what these characters do. Since the definition of clades is based on the analysis of natural histories, it accommodates well the diversity and diversification of living beings. For example, the famous goat (*Capra aegagrus hircus*) discussed by West-Eberhard (2003) is a paradigmatic example of developmental plasticity because it is bipedal. Despite its peculiarities, this specimen is still part of the subspecies *C. aegagrus hircus*.

[3] Homologous characters have a common origin such as the limbs of mice and bats. Analogous characters have similarities but do not have a common origin such as the limbs of mice and the legs of flies. An important criterion to distinguish between the two situations is the appearance of new developmental processes (Wagner and Lynch 2010).

Biological observations typically refer to a specific clade, and usually a specific species or subspecies. Biologists experiment on different specimens of this clade. By definition, this only ascertains a relatively recent (at evolutionary time scales) shared ancestor and many shared characters abstracted from their function (i.e., abstracted from what they do). This common past goes with functional similarities between the specimen studied but does not guarantee that the properties of interest will follow the same relationships.

3.2 Observed and controlled genealogy

The design and description of many biological experiments use genealogical elements that go beyond what systematics can provide. This genealogical knowledge stems from the direct observation of the lineages leading to the specimens studied and can be more or less complete. For example, in the case of cells in culture, usually only a common ancestor and an estimate of the number of generation from this common ancestor are known. Moreover, direct genealogical knowledge is limited to the historical period where biologists follow the necessary methods, that is to say, roughly a century at best.

Usually, direct genealogical knowledge is not limited to observations and includes more or less control over the genealogy. In the case of organisms reproducing sexually, there are two main strategies to control genealogies: establishing inbred or outbred strains, see figure 3A. Inbred strains stem from several generations of inbreeding. By enforcing this behavior, biologists obtain a genetically homogeneous population. Inbred strains still change over time at least as a consequence of genetic drift. These changes lead to the definition of substrains that have biologically relevant differences and sometimes lead to confusions (Simpson et al. 1997). By contrast, breeding outbred strains aims to maintain heterozygote populations while keeping as much genetic homogeneity as possible. These strains are more genetically labile than inbred strains and are often considered more variable phenotypically (Chia et al. 2005; Festing 2014, however see Jensen et al. 2016).

A specific nomenclature for strains completes the nomenclature stemming from systematics. For example, a widespread strain in biomedical research is the inbred mouse strain C57BL/6 (Black 6) (Festing 2014). Naming strains to report an experiment includes the breeding institution. For example, C57BL/6NCrl are Black 6 mice which come from the National Institutes of Health (N) and are bred by Charles River Laboratory (Crl) (Sacca et al. 2013).

In the case of cells, the situation is more straightforward since it is possible to produce a clonal population by starting from a single cell, at least in some cases, see figure 3B. Overall, the situation is similar to the case of animals. Cell lines and sub-lines are established, named, and exchanged between laboratories. For example, the first laboratory immortal human cell line, the HeLa cell line, originated from a patient, Henrietta Lacks (who died in 1951). This cell line is widely used, and more than 99000 references in PubMed mention it (08/2018). Biological experiments using this cell line use cells that have a common historical origin since they stem from an individual organism. In many cases, biologists use a finer-grained control over the genealogy of the cells such as using clonal populations.

We mentioned that even inbred strains vary over time. The same applies to clonal populations of cells. However, in the latter case, higher stability can be obtained by the use of frozen samples, see figure 3B. For frozen samples, the biological time has stopped. Biologists use them to obtain populations of cells that are genealogically closer to the common ancestor of a clonal population than cells which are proliferating with variation in culture.

Both strains and cell lines can be modified for research purposes, either by artificial selection for a specific trait or by genetic engineering, a subject extensively discussed by Kohler (1994) in the case of *Drosophila melanogaster*. The modifications of these populations include ruling out animals with spontaneous but problematic mutations.

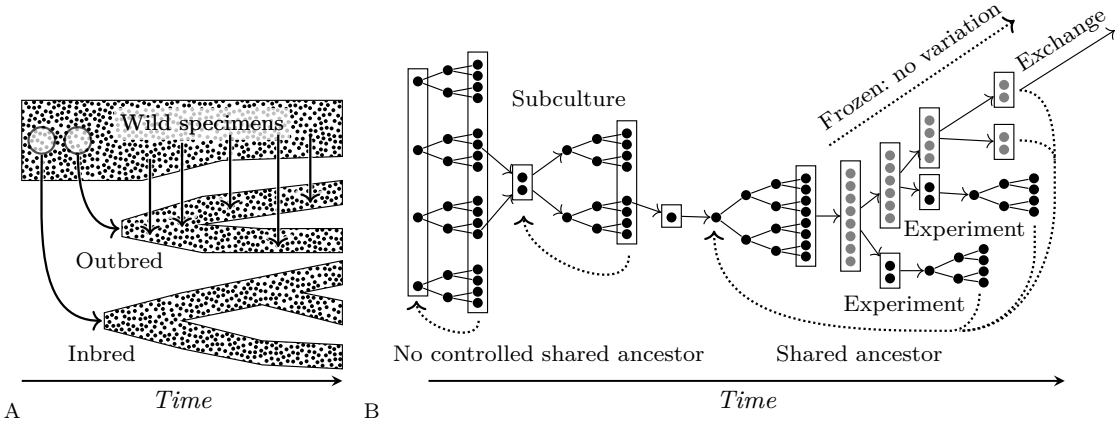

Figure 3 *Observed and controlled genealogies.* A: *A schematic representation of strain breeding.* Biologists use wild or domesticated specimens to start controlled strains. In the case of inbred strains, there is no crossing with specimens external to the strain. In the case of outbred animals, some diversity is regularly introduced (Chia et al. 2005). Substrains may be defined, either because they are the result of genetic manipulations, selection in outbred strains, or just as the result of genetic drift. B: *Control of genealogies in the case of cells.* The initial population has unknown or loosely known ancestry. Doing a standard subculture is not enough to ensure that the population shares a recent common ancestor. To ensure recent shared ancestry, biologists typically perform a highly diluted subculture which isolates a single cell. This cell and its descent proliferate, and this proliferation leads to a new population. This population can then be frozen in order to stop biological processes, and in particular to stop proliferation with variation (Soto et al. 2016a). Subsets of this frozen population can be used to perform experiments and shared with other laboratories. The cells obtained using this method share a known, recent common ancestor. Even in experiments performed years later, the number of generations separating these cells from the common ancestor of the whole line is limited because biologists can use the frozen cells.

Moreover, it is common practice to communicate live sample between research laboratories or between breeding institutions (Kohler 1994, chap. 5). Communicating live samples is required for biologists to ensure that the specimens studied in different laboratories are close genealogically and carry the same modifications if any. Commitment to perform these exchanges is required to publish in many journals. Replicating an experiment using specimens from a controlled genealogy requires an exchange of matter, a point that we discuss in section 4.1.

Genealogies are not limited to cell division and sexual reproduction: viruses lead to horizontal transfers, biologists use a diversity of manipulations, such as chimera obtained by the fusion of different zygotes, and some authors consider that microbiomes should be considered as parts of organisms which implies that several lineages come together to form a holobiont (Gilbert 2014). These examples are beyond the basic concept of genealogy but fit a broader concept of genealogy which describes the historical origins of specimens.

The use of controlled strains and cells lines is not universal in biological experiments. For example, cells may come directly from recent human samples, and animals may come from captures in the wild. However, the practice of using sometimes very tightly controlled genealogies is widespread, in particular in biomedical research. The active control of genealogies, including modifications, leads to a situation where the natural history of the specimens is entangled with the human history of biological sciences (see Kohler 1994, for a discussion in the case of *D. melanogaster*).

The knowledge and control over part of the recent genealogy of the specimens experimented upon is a complement to the phylogenetic method of classification. It ensures that the specimen studied have a recent shared past. Even though this control is tighter than with the description of systematics alone, the same theoretical and philosophical limitations apply: the description is historical and does not ensure that the specimens have the very same organizations. Nevertheless, several methods provide a partial control over biological organizations. For example, inbred strains

are homogeneous genetically, and some aspects of animal phenotypes are controlled regularly in breeding institutions. In conclusion, these methods provide tight control over the historical origin of the specimens studied and limited direct control over their organizations.

3.3 Historical context

Knowledge and control of the past of organisms and cells are not limited to their genealogy. Their past contexts are also relevant. By context, we mean the environment, including the possible interactions with other organisms. The control of past contexts can go from the timescale of many generations to the timescale of ontogenesis or even the shorter time scales preceding the experiment.

In the case of cell culture, the control and knowledge of the context correspond first to the use of a standardized medium, the control of temperature and the avoiding of contaminations. Even the choice of supplies such as centrifugal tubes used with the medium can have dramatic consequences on cellular behaviors (Soto et al. 1991; Colborn et al. 1996).

Another critical parameter is the density of cells. When this density is too low, the lack of quorum effect can change cellular behaviors. On the opposite, when the density is too high the cells constrain each other's proliferation. Moreover, cells typically need time to adjust to a change of conditions such as a change of medium composition. All these factors are important since they determine the status of the cells subjected to the experimentation. In order to perform controlled experiments, experimenters choose an initial status that can be obtained consistently in a cell population (homogeneity) and in different replicates (reproducibility). The most straightforward condition that can be obtained and sustained consistently is unconstrained proliferaion.

In the case of animals, the situation is similar to that of cells. In laboratory conditions, the control of the context includes typically the temperature, light cycle, the nature and quantity of food, avoiding pathogens, and the number of animals per cage. For example, Heindel et al. (2015, section 2.6) describe the context in which animals are raised before and during a large scale experiment. However, their past context can be considered problematic. This work aims to study the effects of the endocrine-disrupting chemical bisphenol A (BPA). The animal experimented upon are raised in BPA free cages, but they originate from strains which are raised in polycarbonate cages by the animal provider, and polycarbonate leaks BPA. Moreover, the exposure of pregnant females to BPA have known effects spanning two generations ("grandmother effect", Susiarjo et al. 2007) and there may be unknown epigenetic factors.

Understanding the importance of past contexts requires a short theoretical discussion on heredity. Under the assumption that DNA sequences are the only form of heredity, past contexts do not matter beyond one generation, where contexts impact development. However, this assumption is not valid in general, and epigenetic inheritance is a widespread phenomenon (Jablonka and Raz 2009; Jablonka et al. 2014; Danchin et al. In press). Let us introduce a simple example that does not require recent progress in epigenetics. MMTV is a retrovirus which can be inherited exogenously from the milk of an infected host to another animal, usually its descent (Dudley et al. 2016). If, say, inbred mice are fed milk from contaminated mice of another strain, then these mice will carry MMTV and transmit it to their descent. A contaminated female will lead to a substrain that is genetically identical to the original inbred strain (as long as the retrovirus does not alter mice DNA) but has critical immunological and oncologic differences.

Many strategies such as working with inbred strains or clonal cell populations target genetic uniformity. These kinds of strategies could be extended formally to known forms of epigenetic heredity. However, our point is that the knowledge and control of past contexts over several generations is an *indirect*, partial way to control known and unknown epigenetic heredity, in combination with the control of genealogies. As a conclusion, past contexts over several generations are relevant.

The context at the time scale of one generation is even more relevant. The context impacts profoundly the development which leads to the concept of ecological developmental biology (Gilbert

and Epel 2009). Even the position of a fetus with respect to its male and female siblings in the uterus has a measurable impact (Ryan and Vandenbergh 2002). The context matters at shorter timescales too. For example, to measure heart rate or blood pressure on a rat, biologists need to take into account the memory and anticipation associated with the procedure (Longo and Montévil 2011b; Nadin 2017, for conceptual frameworks). In this particular case, the stress induced by the measurement impacts the heart rate, and it is possible to limit these changes by training the animal, that is changing its anticipations (Gross and Luft 2003).

The context in which organisms and cells live before the experiment matters from the time scale of several generations to the timescales of development and physiology. The control on the context is a complementary method to control the past of the specimens studied during an experiment and to ensure that some interactions do or do not take place.

3.4 Synchronic aspects of measurement

The aspects of measurement discussed above are mostly diachronic: they pertain to the past of objects. By contrast, this section analyses aspects relevant during the observation of intended features.

3.4.1 Current context

Overall, the discussion in the previous section applies also to the context during an experiment. The context contributes to the definition of the specimens and quantities observed. This contribution is both practical and theoretical. It is practical because it describes the necessary operations required to perform the same measurement beyond using the same apparatus and reading its results. It is theoretical because the meaning of the results depends on these operations.

To illustrate the importance of the context, let us consider the example of mammal metabolism observed by the oxygen consumption rate. This rate seems to be a simple empirical quantity; however, it depends on the activity of the organism observed and its relevant components. To compare the metabolism of different organisms, biologists define different kinds of activity that organisms can follow. The target activities have to be meaningful and achievable for all the organisms considered, which may be diffult when measurement applies to the many different species of a large clade. In all cases, the meaning of the results depends on the nature of the activity chosen, see figure 4. Biologists developed several approaches for the metabolic rate (Longo and Montévil 2014, chap. 2 for a review).

- The field metabolic rate (FMR) corresponds to the activity of organisms in an ecosystem, without constraints from the observer (figure 4E1).
- The basal metabolic rate (BMR) considers an activity which levels down the specificities of the specimens considered. This activity reduces temporarily the impact of the different variations that took place in natural history (figure 4E3). In common biological terms, the basal metabolism corresponds to an undisturbed, non-sleeping organism in a thermoneutral environment and in a post-absorptive state. It is not always simple to instantiate this definition. For example, ruminants are always digesting; therefore it is not possible to put them in a post-absorptive state (figure 4E2).
- The maximum metabolic rate (MMR) considers the maximum level of sustainable activity. By focusing on the upper boundary of the metabolism, only the determinants of this boundary are relevant and not the various characters involved in biological activities (figure 4E4).

By choosing different contexts, biologists co-determine what is observed even when the same measurement apparatus is used to observe the same part. The case of the BMR and MMR show that it is even possible to choose observations that focus on characters that are common to different

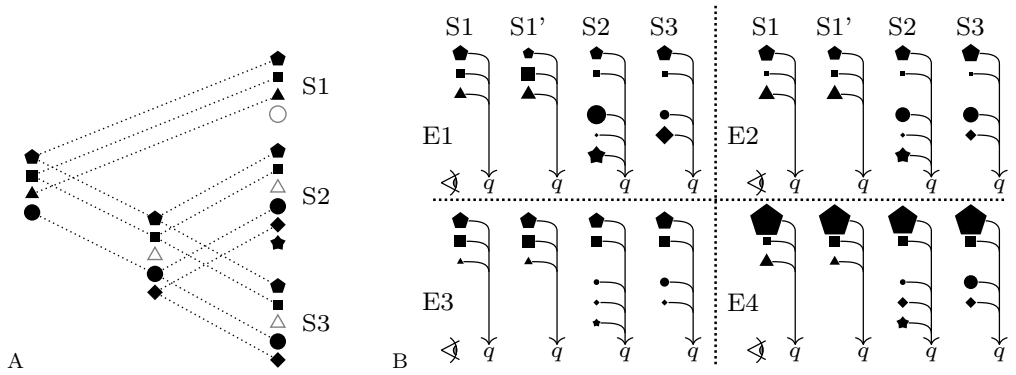

Figure 4 *Different measurements of the same quantity.* A: *A schematic representation of the appearance and disappearance of relevant characters.* Dotted lines represent relations of homology. White shapes are characters that disappeared. B: *Four different ways to measure a quantity q.* S1 and S1' are two specimens of the same species. All represented characters impact q. The size of the symbol representing a character shows its impact on q in the given context. E1: A measurement performed without specific care for the characters contributing to q, e.g., the field metabolic rate. E2: A measurement performed in a standardized way for S1 but not for the other species. E3: The animal performs no specific activity which reduces the weight of several characters, e.g., the basal metabolic rate. In this case, only homologous characters remain quantitatively relevant. E4: A constraint dominates the determination of the measured quantity despite the diversity of relevant characters, e.g., the maximum metabolic rate.

species, by leveling down the organizational diversity stemming from history and impacting the measured quantity. We discuss this situation further in section 4.

3.4.2 Choosing or eliminating individuals

Filtering out individuals is a method to control strains: breeders disregard animals with mutations, diseases, or other peculiarities. Moreover, sometimes only minimal control over the past context and genealogy is possible. For example, in humans, most methods above would be unethical. Choosing individuals having specific characteristics and eliminating individuals with unwanted characteristics is an alternative method of control on the organisms studied.

Filtering individuals is possible during experiments; however, filtering impacts the meaning of the results. For example, in the case of a toxicological experiment, unexpected variations should be reported since they may be relevant to understand the effect of the chemical studied. However, if we want to study the "normal" physiology of insulin after long-term exposure to high-sugar diet, then it is acceptable to rule out animals having diabetes. Of course, the quantities of interest cannot be measured at the expected time point in the case of individuals who meet an untimely death, which is an uncontrolled filter.

Filtering individuals on the basis of their properties is a complementary way to control biological objects. Performing this kind of filtering enables biologists to discard specimens which have gone through unwanted variations, or which have not gone through expected variations. Criteria can range from developmental anomalies, mutations, pathologies to an animal that is frightened during measurement.

3.4.3 Data acquisition

Lastly, biological measurements typically provide quantities, and this process has an anhistorical dimension that is comparable to physics. More precisely, as pointed out in the introduction, the notion of measurement of classical physics is relevant in biology. When measuring a continuous

quantity, the measurement will never be exact and will provide an interval instead of a single quantity see section 2.1. Other physical notions such as reference frames can also be relevant for biology. Since these aspects are not proper to biology, we will not develop them further here.

Let us mention that the principle of variation implies that an observed feature can become ill-defined or acquire a different meaning. For example, the heart rate is defined by beat to beat intervals, but pathological situations such as torsade de pointes escape the standard definition of a heartbeat, and the notion of heart rate becomes ill-defined. Similarly, if a mutation impacts coding DNA, then the concentrations of the corresponding mRNA with and without the alteration can be seen as the same variable or not depending on the biologist' perspective.

Last, most experimental protocols in biology use control groups which are not subjected to the transformations studied (Johnson and Besselsen 2002). Control groups have a statistical role. However, in line with the discussion by Weber (2004, chap. 4) on control experiments, their role goes beyond statistics. Control groups enable experimenters to assess the organization of specimens subjected to almost the same conditions that the organisms experimented on and are therefore a way to assess whether the results stem from the context, spontaneous variations, or conditions tested. Biological objects are labile, and control groups are the closest reference point possible to the objects tested.

3.5 Irreducibility of biological variation

Despite the use of methodologies providing tight control over biological objects, the principle of variation entails that there are always possible remnant variations. Variations can impact the observed features directly, making them variable, changing their meaning or even possibly making them ill-defined. When observing a given feature among several specimens, biologists report "not applicable" (NA) for a specimen when qualitative variations are too important for an observed feature. For example, pathological heartbeats that do not follow the same sequence of events that regular heartbeats lead to beat-to-beat intervals that do not have the same meaning. This kind of departures appear for theoretical reasons and not just as a result of experimental errors.

Observable, qualitative variations can be shown experimentally even for clonal cells, for example as a result of asymmetries in cellular division (Cai et al. 2006; Stewart et al. 2005; Lindner et al. 2008; Soto et al. 2016a) or from dynamical reasons (Braun 2015). Of course, the development of multicellular organisms also leads to a high level of variations. Variations occur even when comparing an individual with itself at another time point, even in the case of close time points. For example, many physiological time series are non-stationary (see Longo and Montévil 2014 chap. 2 and West 2006). Stationary time series follow the same distribution over time. In this case, the mean is a stable quantity and observing the phenomenon over a sufficient duration provides an estimation of it. By contrast, non-stationarity means that assessing an average quantity over time twice will not necessarily provide the same results. As a consequence, it is not possible to characterize an organism by precise values of physiologic quantities, and precise results are only valid at a specific time.

3.6 Recapitulation

To sum our theoretical approach up, we propose the following principles :

1. The measurement is relative to/constituted by a history and a context. History, here, means a cascade of context-dependent, qualitative variations. A biological measurement is a specific way to manipulate and describe these contexts and natural histories.
   (a) *Genealogy* handles an uncontrolled history that is *shared* by the different organisms studied. Methods include the phylogenetic classification (§ 3.1) and direct genealogical control in the case of strains and cell lines (§ 3.2).

(b) *Past and current contexts* (environment/interactions) can be (partially) known in the field or controlled in laboratories or breeding institutions . Relevant contexts include relatively recent past contexts, for example over several generations, during the development or shortly before observations (§ 3.3), and current contexts, during the experiment and observations (§ 3.4.1).
(c) *Choosing or eliminating individuals* can be used to observe or eliminate specific histories or variations (pathological cases, unwanted behavior, ontogenetic or phylogenetic histories, . . . § 3.4.2).

2. Uncontrolled variations can always impact the measurement, including the very definition of the features observed (§ 3.5).

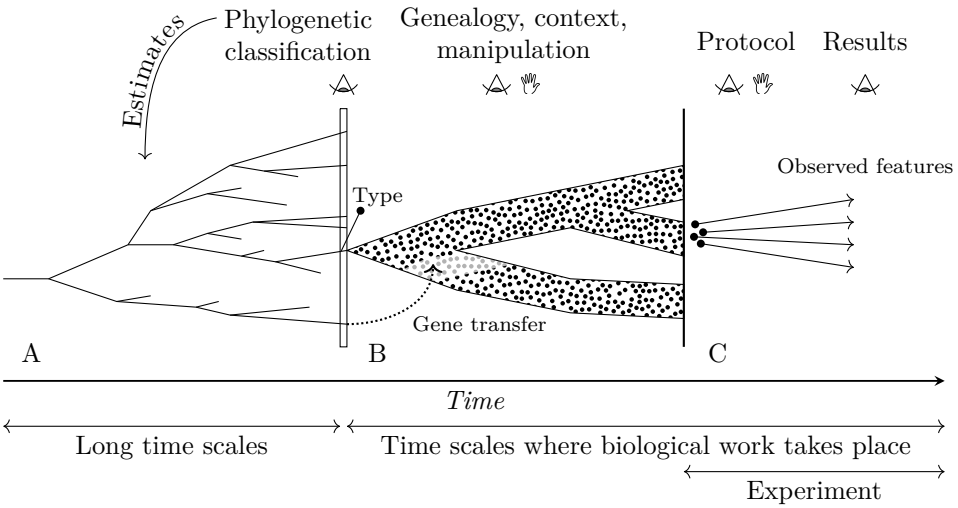

Figure 5 *Recapitulation of the diachronic elements used to define the objects of a typical experiment.* The whole construct illustrated is required to describe the measurement performed. A: These objects are the result of an evolutionary history, which is not directly accessible but can be estimated by the phylogenetic classification method. B: Specimens of a given species can be used to start a strain raised and bred in controlled condition. C: Elements of this strain are used in an experiment to obtain data.

4 Discussion

In this section, we discuss the philosophical and practical consequences of our notion of biological measurement.

4.1 The materiality of biological phenomena

The notion of the materiality of biological phenomena states that biological knowledge is more intertwined with matter than physical knowledge (Soto et al. 2016b). Our discussion on biological measurement illustrates this idea. First, biological names, in systematics, are not associated with theoretical constructs and are associated instead to specific specimens called name-bearing types

(§ 3.1). These specimens are material objects kept in collections. Then, experimenting with individuals of a species associated with this name means experimenting on individuals which descend from a material, common ancestor, shared by both the specimens experimented upon and the name-bearing type. Even though these specimens' lineages may have diverged, they possess a diachronic, material continuity made explicit by their genealogy. The same reasoning applies to the controlled strains and cell lines often used in experimental biology. The exchange of living specimens between laboratories is the further materialization of this philosophical idea.

In general, we can distinguish between different categories of theoretical situations in order to reproduce observations.

1. The description of objects is generic, and the same theoretical object can be instantiated empirically twice without communication of matter, as discussed by Feynman and Gleick (1967, chap. 4). It is typically the case in physics.
2. The object's behavior is the specific result of a history.
   (a) Scientists use the permanence of the material object studied. The name-bearing types mentioned above are a good example: they are the specific result of a history and serve as static references for future observations.
   (b) The objects can reproduce. These objects provide an exponential amount of objects sharing a common past. The study of living organisms and cells falls typically in this category (exceptions are case study such as types above).

4.2 Symmetry and symmetrisation

Symmetries play a central role in physics (Feynman and Gleick 1967; Van Fraassen 1989; Longo and Montévil 2014) and will enable us to provide a more in-depth analysis of biological measurement. Symmetries are transformations that do not change the relevant aspects of a given object. For example, the equation describing free fall does not change for an experiment performed one century ago or today. Time translation is a transformation that does not change the theoretical description of the object: a symmetry. Another fundamental symmetry, here, is that the same equation applies regardless of the nature of the object. If experimenters replace an iron bead with another one, or a copper or wood bead, the phenomenon remains the same: permuting (interchanging) these objects is a symmetry. Note that symmetries can be either exact, in the sense that they stem from fundamental principles, or approximate.

The very concept of experimental reproducibility is a notion of symmetry. The reproducibility of an experiment means that the same set of observations can be performed by different observers, usually on different material objects and at different times and places. Moreover, in a given experiment, biologists typically use different specimens exposed to the same conditions in order to perform statistical tests. The tests assume that these specimens follow the same probability distribution, that is the tests assume that behind the quantitative variations for the observed features there is a single abstract mathematical object (the probability distribution): this is again an assumption of symmetry.

However, biological objects are the result of a history and continue to generate a history. Interpreting this notion in terms of symmetries leads to assert that when time flows, the changes of biological objects can require changes of symmetry that do not stem from the description of the initial objects. These changes are the core of the principle of variation (Longo and Montévil 2011a; Montévil et al. 2016; Montévil 2018). As hinted to in the introduction, these variations conflict with the aim to perform reproducible experiments. In biology unlike in physics, the symmetries associated with reproducibility are not granted theoretically. Instead, they depend on the measurement as summarized in section 3.6.

The concepts of symmetry and symmetrization are useful to understand biological measurement. Since biological regularities are more labile than physical ones, symmetries are not provided

directly by the theory. Instead, they are co-established by the measurement process and the biological objects used — and these two aspects are intertwined. We propose to call this practical and theoretical operation "symmetrization". Biologists typically work on specimens of the same species or more generally specimens with a shared common past. In experiments, they assume a partial equivalence between these specimens and the way in which they are organized. In other words, biologists establish an approximate symmetry between their organizations and their response to experiments. Control over past contexts is also a symmetrization of the specimens studied. For example, in section 3.3, we have discussed how biologists aim for cells *in vitro* to be in a consistent state over time, that is to say, symmetrize them. The different methods described in section 3 should be seen as different symmetrization.

Different symmetrizations can be performed during an experiment or even during data acquisition. Choosing a symmetrization or another endows the results with entirely different biological meanings. Figure 4 illustrates this idea and shows different ways to make organisms equivalent. In figure 4E1, by being in the field, organisms express their historically (evolutionary) relevant activities and these activities are diverse. In figure 4E3, different organisms are mostly restrained to activities that are common to them. The experimenter performs a more controlled symmetrization by limiting the characters involved in the determination of the observed quantity. The method in figure 4E4 is similar but targets a specific common aspect of biological organizations and ensures that it overwhelms the contribution of other characters.

We can distinguish two kinds of symmetrization: concrete and epistemic symmetrization. Concrete symmetrization involves the action of biologists on objects. For example, establishing inbred strains or the symmetrization performed in figure 4 are concrete symmetrizations. By contrast, epistemic symmetrization does not change material objects and is limited to choosing what is considered equivalent, that is to say, symmetric by permutation. For example, we mentioned that the position of a fetus in utero has measurable consequences. Taking this aspect into account or not corresponds to different epistemic symmetrizations. Epistemic symmetrization is particularly relevant for statistical analyses and subsequent biological discussions.

We hypothesize that epistemic symmetrization is always relevant because biologists have to symmetrize organisms that are not genuinely symmetric as a result of the principle of variation. Concrete symmetrization is relevant in most experiments. It is not typically relevant in observations without experiments such as the observation of specimens in systematics. Performing concrete symmetrization constrains methodological choices, and thus the kind of biological objects studied. For example, it is far easier to symmetrize cell culture by maintaining unconstrained proliferation than to use an alternative symmetrization with consistent periods of proliferation and saturation.

In conclusion, biological measurement as summarized in 3.6 describes both the concrete and epistemic symmetrization performed to obtain experimental results and endow them with biological meaning.

4.3 Measurement strategies

Baxendale (2018) proposes to map scientific practices on a continuum of strategies defined by their stances concerning reductionism. In this section, we use a similar idea concerning measurement. Our concept of biological measurement leads to the notion that measurement depends on symmetrizations, but symmetrizations can be performed more or less tightly and at different levels. To represent these strategies, we propose to organize measurement strategies along three axes as illustrated in figure 6. In this section, we discuss only the strategies of different measurements, and we do not imply that they necessarily succeed or that they prevent the use of other strategies to address the same problem.

The first axis describes whether the measurement is variable or on the opposite reproducible. Here, reproducibility means that the measurement generates data consistently with different specimens. For example, using inbred strains leads to more reproducible results than wild specimens.

The second axis describes whether the measurement targets singular or general objects. Working on the metabolic rate of mammals is more general than working on a single species by measuring wild specimens. Both are more general than outbred strains and *a fortiori* inbred strains. Reciprocally, inbred strains are more singular than outbred strains and so on.

The last axis assesses whether the measurement defines objects coherent with their evolutionary past or instead whether the objects are more or less profoundly altered. For example, inbred strains are homozygotes for all genes. Similarly, the basal metabolic rate is not representative of a species past evolutionary context, and the field metabolic rate is more coherent with the natural history of the organisms studied.

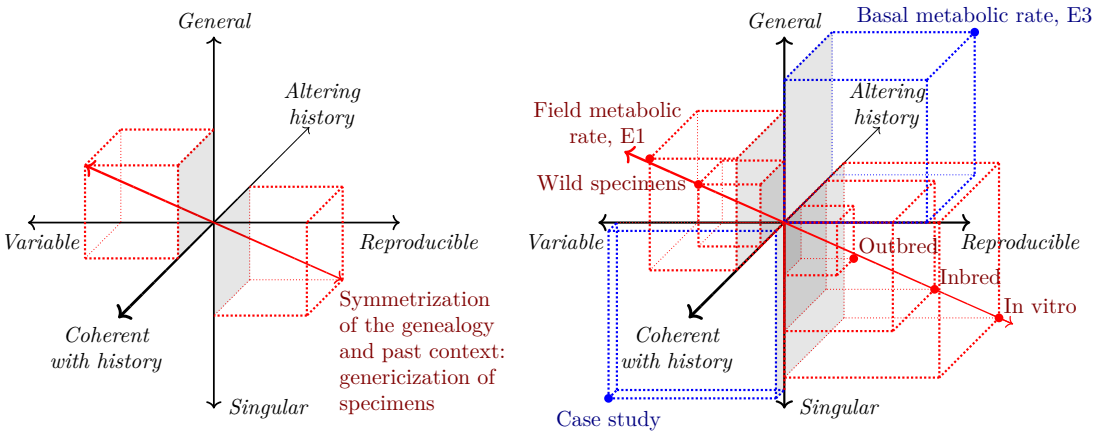

Figure 6 *Different measurement strategies.* The three axes correspond to measurements that are reproducible vs. variable, general vs. singular and coherent with their evolutionary past vs. altered by experimenters. Left: axis where many strategies lie. On one end of the spectrum, these strategies aim to produce specimens that are as close to each other as possible by controlling them tightly. On the other end, experimenters relax this control and aim to make more general measurements. Right: different cases are represented in the space describing measurement strategies. Most measurements are on the axis represented on the left, but departures from this axis are equally interesting since they represent other ways to approach biological phenomena empirically such as case studies.

A qualitative axis emerges in this three-dimensional space, see figure 6. This axis is given by the strategies which increase simultaneously the reproducibility, singularity, and alterations associated with the measurement. In other words, these strategies lose generality and alter the specimens in order to increase the reproducibility of the measurement. At the limit, these strategies aim to generate specimens that are as symmetrized as possible, that is specimens that would be close to a single generic object like physical objects are: these strategies aim the genericization of biological organizations and use many methods to reduce diversity. For example, in the case of cells, samples are frozen to prevent spontaneous variations between experiments. The focus on model organisms at the level of the research community is a collective strategy of genericization. In situations like clinical trials, on humans, genericization is very limited for ethical reasons: populations that are too similar is evidence of malpractice Bolland et al. (2016). In contrast to genericization strategies, other strategies on the same axis aim to gain generality and coherence with evolutionary history but at the cost of more variability in the results (i.e., less reproducibility).

In order to face the reproducibility crisis and to obtain significant results with fewer animals, it is common to promote strategies genericizing specimens (Festing 2014; Chia et al. 2005). However,

these strategies bear the cost of studying singular behaviors: the results obtained may not even be representative of the species studied. For example, Black 6 mice have singular features such as their nociception (sensation of pain) (Mogil et al. 1999). Isaacs (1986) tested the incidence of tumors in rats exposed to the carcinogen DMBA and found that this incidence is 0%, 15%, 40% and above 90% depending on the strain used. These examples show how profoundly the choice of strain can impact results. Moreover, the conditions of the laboratory reduce exposure to pathogens in order to symmetrize the life history of animals studied, which is part of the alteration axis. However, this situation leads to immunological functions that differ from wild animals (Abolins et al. 2010).

The genericization of specimens aims, at the limit, to study a single, reproducible organization and is thus highly singular. Results may depend on the specificities of these organizations and their contexts in unknown ways. Therefore, these strategies can be vulnerable to minor departures from the genericization performed initially. For example, performing measurement in different laboratories always involve a change of context despite the explicit control of many factors. Genericization aims reproducibility in the sense of specimens that are very similar, but the reproducibility of experiments is made difficult by the lack of generality guaranteed by the measurement.

In figure 6, there are only two cases which are far from this axis. The first corresponds to measurements like the basal metabolic rate, see figure 4E3. This measurement is reproducible and nevertheless general. The downside is that the organisms are put in a specific state to level down their differences impacting the feature observed. Genericizations above apply to the whole organizations of specimens. By contrast, the strategy of E3 only targets a few specific aspects of these organizations.

Case studies are the second kind of strategies departing from the main axis. Case studies focus on a single individual and reproducibility is not a goal. For example, Patterson and Linden (1981) study the intelligence of non-human primates. To do so, they did not develop standardized conditions and protocols. Instead, they taught sign language to several specimens and focused on a particularly gifted gorilla, Koko, who mastered up to 2000 symbols. Other examples are works on types in systematics, the study of the bipedal goat discussed by West-Eberhard (2003, see also section 3.1) and the cloned sheep, Dolly, which is one success among 277 attempts and which remained a single success for a long time (Wilmut et al. 1997). Note that while the study of types does not involve alterations, teaching Koko or cloning a sheep do: case studies are diverse for the third axis.

Case studies are sometimes neglected by experimenters who strive to design reproducible experiments in order to study mechanisms. For example, the success of cloning Dolly without reproducing this feat led to an intense debate, especially when evidence accumulated that Dolly was indeed cloned from an adult cell (Solter 1998). However, in our conceptual framework, case studies have a specific epistemic role. They are sufficient to prove the existence of a possibility in a context where biological possibilities are not predefined (Montévil 2018). The bipedal goat exemplifies developmental plasticity and studying a type is sufficient to defend the existence of a new species. Let us emphasize that case studies can be very detailed. For example, the anatomy of the bipedal goat provided insights on developmental plasticity and Koko provided insights on Gorilla's thought processes. Moreover, the study of types plays a crucial role in the general architecture of biological knowledge.

This short discussion shows that it is fruitful to represent different measurement strategies by the symmetrization operated. These strategies are different responses to the difficulties raised by the historical and changing nature of biological objects.

5 Conclusion

Our theoretical notion of measurement accommodates the way in which biologists manipulate immensely complex objects, organisms and cells typically, which are the result of a history and

continue to produce a history by generating qualitative variations. Our notion of measurement states that a measurement is relative to a history and context. To develop reproducible experiments, biologists observe specimens with a shared past. This shared past is ascertained by systematics and by direct knowledge and control of both their genealogy and past contexts. In the study of objects defined by their history, the objects which can be considered equivalent are objects stemming from the same history. In this context, we call symmetrization the concrete and theoretical operations which establish and posit the equivalence of different objects with more or less tightly controlled shared pasts. Symmetrization also includes the operations performed during the observation which can either lead to variability or reproducibility.

The notion of biological measurement is compatible with different research strategies and leads to a framework to map them. In this framework, we find two opposite poles. In one end, strategies aim at the genericization of biological organizations at the cost of studying singular organizations and altering them. To implement these strategies, biologists developed a plethora of methods. They expose objects to similar contexts and ensure that they have recent, controlled common ancestors. In some cases, biologists freeze samples to prevent them from undergoing variations between experiments. On the other end of the spectrum, the objects studied are more general (e.g., diverse genetically) and coherent with their evolutionary history, but they are also more variable. There are strategies which escape this opposition, for example, case studies or methods to level down relevant diversity to symmetrize a part without symmetrizing the organization as a whole.

Having a clear notion of what it means to access biological objects empirically is crucial for biological knowledge. In this paper, we provide only an outline of biological measurement, and this notion deserves further discussions, focusing on both general and specific situations. Nevertheless, since our notion builds on solid ground, namely the theory of evolution, we hope our work can provide a stable starting point for further research. We have shown that biological measurement has critical differences with the notion of measurement in physics. Acknowledging these differences should provide a more profound and systematic way to approach biological experiments and observations critically and ultimately to promote experimental reproducibility.

References

Abolins SR, Pocok M, Hafalla J, Riley E, Viney M (2010) Measures of immune function of wild mice, *Mus musculus*. Molecular Ecology 20(5):881–892, 10.1111/j.1365-294X.2010.04910.x

Baker M (2016) 1,500 scientists lift the lid on reproducibility. Nature 533:452–454, 10.1038/533452a

Baxendale M (2018) Mapping the continuum of research strategies. Synthese 10.1007/s11229-018-1683-1

Beatty J (1995) The evolutionary contingency thesis. In: Conceptual Issues in Evolutionary Biology., MIT Press, Cambrdige, MA, p 217

Begley C, Ioannidis J (2014) Reproducibility in science. Circulation Research 116(1):116–126, 10.1161/CIRCRESAHA.114.303819

Begley CG, Ellis LM (2012) Drug development: Raise standards for preclinical cancer research. Nature 483(7391):531–533, 10.1038/483531a

Bolland MJ, Avenell A, Gamble GD, Grey A (2016) Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. Neurology 10.1212/WNL.0000000000003387

Bookstein F (2009) Measurement, explanation, and biology: Lessons from a long century. Biological Theory 4(1):6–20, 10.1162/biot.2009.4.1.6

Braun E (2015) The unforeseen challenge: from genotype-to-phenotype in cell populations. Reports on Progress in Physics 78(3):036602, 10.1088/0034-4885/78/3/036602

Cai L, Friedman N, Xie X (2006) Stochastic protein expression in individual cells at the single molecule level. Nature 440(7082):358 – 362, 10.1038/nature04599

Chia R, Achilli F, Festing M, Fisher E (2005) The origins and uses of mouse outbred stocks. Nature genetics 37(11):1181, 10.1038/ng1665

Colborn T, Dumanoski D, Myers J, Murden M (1996) Our stolen future. Dutton

Crick FH (1958) On protein synthesis. In: Symp Soc Exp Biol, vol 12, p 8

CZN International (1999) International code of zoological nomenclature. International Trust for Zoological Nomenclature, London

Danchin E, Pocheville A, Rey O, Pujol B, Blanchet S (In press) Epigenetics as the hub guiding nongenetic inheritance to genetic assimilation. Biological Reviews

Dudley J, Golovkina T, Ross S (2016) Lessons learned from mouse mammary tumor virus in animal models. ILAR Journal 57(1):12–23, 10.1093/ilar/ilv044
Festing M (2014) Evidence should trump intuition by preferring inbred strains to outbred stocks in preclinical research. ILAR Journal 55(3):399–404, 10.1093/ilar/ilu036
Feynman R, Gleick J (1967) The character of physical law. MIT press
Gilbert SF (2014) A holobiont birth narrative: The epigenetic transmission of the human microbiome. Frontiers in Genetics 5(282), 10.3389/fgene.2014.00282
Gilbert SF, Epel D (2009) Ecological developmental biology: integrating epigenetics, medicine, and evolution. Sinauer Associates Sunderland
Gillies D (2012) Philosophical theories of probability. Routledge
Gould S (2002) The structure of evolutionary theory. Harvard University Press
Gross V, Luft F (2003) Exercising restraint in measuring blood pressure in conscious mice. Hypertension 41(4):879–881, 10.1161/01.HYP.0000060866.69947.D1
Heindel J, Newbold R, Bucher J, et al. (2015) NIEHS/FDA CLARITY-BPA research program update. Reproductive Toxicology 58:33 – 44, 10.1016/j.reprotox.2015.07.075
Huneman P (2018) Diversifying the picture of explanations in biological sciences: ways of combining topology with mechanisms. Synthese 195(1):115–146, 10.1007/s11229-015-0808-z
Isaacs JT (1986) Genetic control of resistance to chemically induced mammary adenocarcinogenesis in the rat. Cancer Research 46(8):3958–3963
Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. The Quarterly review of biology 84(2):131–176, 10.1086/598822
Jablonka E, Lamb MJ, Zeligowski A (2014) Evolution in Four Dimensions, revised edition: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life. MIT press, London
Jensen VS, Porsgaard T, Lykkesfeldt J, Hvid H (2016) Rodent model choice has major impact on variability of standard preclinical readouts associated with diabetes and obesity research. American journal of translational research 8(8):3574
Johnson PD, Besselsen DG (2002) Practical aspects of experimental design in animal research. ILAR journal 43(4):202–206, 10.1093/ilar.43.4.202
Kohler RE (1994) Lords of the fly: Drosophila genetics and the experimental life. University of Chicago Press
Lancet (2018) Uk life science research: time to burst the biomedical bubble. The Lancet 392(10143):187, 10.1016/S0140-6736(18)31609-X
Lecointre G, Le Guyader H (2006) The tree of life: a phylogenetic classification, vol 20. Harvard University Press
Leonelli S (2014) What difference does quantity make? on the epistemology of big data in biology. Big Data & Society 1(1):2053951714534395, 10.1177/2053951714534395
Lindner A, Madden R, Demarez A, Stewart E, Taddei F (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. Proceedings of the National Academy of Sciences 105(8):3076 – 3081, 10.1073/pnas.0708931105
Longo G, Montévil M (2011a) From physics to biology by extending criticality and symmetry breakings. Progress in Biophysics and Molecular Biology 106(2):340 – 347, 10.1016/j.pbiomolbio.2011.03.005
Longo G, Montévil M (2011b) Protention and retention in biological systems. Theory in Biosciences 130:107–117, 10.1007/s12064-010-0116-6
Longo G, Montévil M (2014) Perspectives on Organisms: Biological time, symmetries and singularities. Lecture Notes in Morphogenesis, Springer, Dordrecht, 10.1007/978-3-642-35938-5
Longo G, Montévil M (2017) Comparing Symmetries in Models and Simulations, Springer. 10.1007/978-3-319-30526-4
McNeill J, Barrie F, Buck W, et al. (2012) International Code of Nomenclature for algae, fungi and plants. Koeltz Scientific Books, Konigstein, Germany
Mogil JS, Wilson SG, Bon K, et al. (1999) Heritability of nociception i: Responses of 11 inbred mouse strains on 12 measures of nociception. Pain 80(1):67 – 82, 10.1016/S0304-3959(98)00197-3
Montévil M (2018) Possibility spaces and the notion of novelty: from music to biology. Synthese 10.1007/s11229-017-1668-5
Montévil M, Mossio M (2015) Biological organisation as closure of constraints. Journal of Theoretical Biology 372(0):179 – 191, 10.1016/j.jtbi.2015.02.029
Montévil M, Mossio M, Pocheville A, Longo G (2016) Theoretical principles for biology: Variation. Progress in Biophysics and Molecular Biology 122(1):36 – 50, 10.1016/j.pbiomolbio.2016.08.005
Mossio M, Montévil M, Longo G (2016) Theoretical principles for biology: Organization. Progress in Biophysics and Molecular Biology 122(1):24 – 35, 10.1016/j.pbiomolbio.2016.07.005
Mugur-Schächter M (2002) Objectivity and descriptional relativities. Foundations of Science 7:73 – 180, 10.1023/A:1016095424229
Nadin M (2017) Rethinking the experiment: necessary (r) evolution. AI & SOCIETY pp 1–19, 10.1007/s00146-017-0705-8
Patterson F, Linden E (1981) The education of Koko. Holt, Rinehart, and Winston

de Queiroz K (1992) Phylogenetic definitions and taxonomic philosophy. Biology and Philosophy 7(3):295–313, 10.1007/BF00129972

Ryan BC, Vandenbergh JG (2002) Intrauterine position effects. Neuroscience & Biobehavioral Reviews 26(6):665 – 678, 10.1016/S0149-7634(02)00038-6

Sacca R, Elder B, Wasson K (2013) The C57BL/6 Mouse: The Role of the C57BL/6N Mouse in the Creation of Future Genetically Engineered Models. Charles River Laboratories International, Inc

Salam A (1990) Unification of Fundamental Forces. Cambridge University Press

Simpson EM, Linder CC, Sargent EE, et al. (1997) Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. Nature genetics 16(1):19, 10.1038/ng0597-19

Solter D (1998) Dolly is a clone—and no longer alone. Nature 394(6691):315

Soto AM, Justicia H, Wray JW, Sonnenschein C (1991) p-nonyl-phenol: An estrogenic xenobiotic released from "modified" polystyrene. Environmental Health Perspectives 92:167–173, 10.1289/ehp.9192167

Soto AM, Longo G, Montévil M, Sonnenschein C (2016a) The biological default state of cell proliferation with variation and motility, a fundamental principle for a theory of organisms. Progress in Biophysics and Molecular Biology 122(1):16 – 23, 10.1016/j.pbiomolbio.2016.06.006

Soto AM, Longo G, Noble D, et al. (2016b) From the century of the genome to the century of the organism: New theoretical approaches. Progress in Biophysics and Molecular Biology, Special issue pp 1–82

Stewart E, Madden R, Paul G, Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. PLoS Biology 3(2):e45, 10.1371/journal.pbio.0030045

Susiarjo M, Hassold TJ, Freeman E, Hunt PA (2007) Bisphenol a exposure in utero disrupts early oogenesis in the mouse. PLoS genetics 3(1):e5, 10.1371/journal.pgen.0030005

Teytelman L (2018) No more excuses for non-reproducible methods. Nature 560:411, 10.1038/d41586-018-06008-w

Turing AM (1950) Computing machinery and intelligence. Mind 59(236):433–460

Van Fraassen B (1989) Laws and symmetry. Oxford University Press, USA

Wagner GP, Lynch VJ (2010) Evolutionary novelties. Current Biology 20(2):R48 – R52, 10.1016/j.cub.2009.11.010

Weber M (2004) Philosophy of experimental biology. Cambridge university press

West B (2006) Where medicine went wrong: Rediscovering the path to complexity. World Scientific, Teaneck, NJ

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press

Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH (1997) Viable offspring derived from fetal and adult mammalian cells. Nature 385(6619):810–813, 10.1038/385810a0