

Chapter 2

Invariances in Theory

Abstract This chapter surveys the multifaceted roles that invariants play in theorizing, from physics and mathematics to biology and neurobiology. The question “What is an invariant of behavior?” is posed, and some alternatives are proposed and discussed: genes, neuroanatomy, and reflex theory. From that, the cybernetic take on the issue is introduced and placed in an evolutionary context, in which single behaviors are identified in respect to the goals they achieve and how they subserve the organism’s viability. The search for invariants of behavior is framed as a search for mechanisms. This search is far from trivial, as assumptions play a prominent role.

2.1 Invariants in Physics and Mathematics

2.1.1 Invariants Are That Which Remains When Something Else Changes

Invariants are usually thought of as conserved quantities, relational structures, constant qualities or functions; in essence, invariants are properties of a system that are conserved irrespective of some manipulation. They are something describable about the system, which, often enough, are abstract properties existing nowhere but in the language of theory. But despite this dubious ontological status, invariants are invaluable as well as fundamental to theories. The discovery of an invariant may reveal connections between sorts of a theory previously thought unrelated. When something is conserved when the system is manipulated, that something becomes a latch for the other variables in the system to change around. It becomes a source of explanatory power, equating other variables in its terms.

2.1.2 Theoretical Invariants: Speed of Light, Planck’s Constant

Without invariants both modern physics and mathematics are scarcely conceivable. Most appropriately the term originates from mathematical disciplines. Examples of

explanatory invariants are profuse and pervasive. Heisenberg's uncertainty principle draws a relation between our knowledge of a particle – how fast it moves, and where it is. It is bounded (above) by a number, Planck's constant, representing the amount of knowledge we may extract in an experiment. Planck's constant is an invariant in quantum mechanics that reappears in the resolution of physical paradoxes, such as the radiation of a black body. It gives a stance of solidity even in uncertainty. Another example of a theory that is anchored in assuming invariants is Einstein's special theory of relativity. It is based on the axiom that the speed of light in a vacuum with respect to an inertial reference frame is always constant, a *theoretical invariant*. In the test of this assumption many physical paradoxes found resolution, e.g., the Michelson–Moreley experiment. Incidentally, Einstein eventually regretted having called the theory *Relativitätstheorie*, having later preferred *Invariantentheorie*. Both the speed of light and Planck's constant are theoretical invariants; no manipulation affects the conserved quantity, the former determined by postulate, the latter by deduction.¹

2.1.3 Theoretical and Relational Invariants: Energy

A giant leap in physical science followed the inception of the concept of energy of a system (Leibniz was the first to propose a formulation), an abstract quantity that is conserved through transformations of the components of a physical system. As an object falls from some height, the increasing velocity results from the transformation from potential energy to kinetic energy. As the object hits the ground, it produces sound, heat, and plastic alterations of the object, which are all expressible in terms of energy. Through the principle of conservation of energy we know that the heat, sound, and shape deformations produced all relate to the same quantity, which remains unaltered throughout the process. At a given time, the total amount of energy of a closed system is constant, and the invariant relates all components and their states within a system. Energy is the common currency across transformations, and is invariant for a closed system. If the system is open, i.e., other components are injected or removed, we know that the energy of the new state will increase or decrease by an exact number, representing the energy fluxes in and out.

2.1.4 Empirical and Relational Invariants: Law of Gases

Other examples of invariants relate quantities, such as the Boyle–Mariotte law in thermodynamics, which relates the ratio of pressure and temperature of an ideal gas to a constant. The gas law was found experimentally; it is not absolute. An ideal

¹ Other purely theoretical examples exist, such as the entities in the study of topology. For example, a torus – a donut – will always have one hole irrespective of deformations, because its shape identity is defined by its topological properties. The invariant property is a part of the entity's definition.

gas is an abstraction; there are variations dependent on particular gases. But all gases follow the rule, within some tolerance bounds, that pressure and temperature are inversely proportional. Similarly, the number of atoms in a volume, given temperature and pressure, is constant. They are related by a constant, first encountered in experiments, then derived analytically (among others, by Einstein [49]). Avogadro's number specifies how volume relates to the number of molecules in a system. Like in Boyle's law, Avogadro's number closely follows reality; it is also an idealization conjoining empirical measurements. It works for ideal gases, which do not exist, but serves as a template for all real gases. Empirical invariants are constants and proportions and may indicate general principles underlying theories. Another example fitting this category is the thermal expansion coefficient, and it also fits in the next category. Both the thermal expansion coefficient and the gas constant are context-dependent, as both depend on the material or gas.

2.1.5 *Context-Dependent Invariances: Sum of Internal Angles of the Triangle*

Some invariants are absolute, as Planck's constant, and no conceivable manipulation affects the quantity bound by it. However, this is the exception; most invariants are context-dependent, requiring a *ceteris paribus* condition: that everything else should be kept constant. Some invariants are so only within a frame or a context (e.g., speed of light *in a vacuum*). Even in mathematics there are examples of context-dependent invariants Table 2.1. In Euclidean geometry the sum of the internal angles of a triangle is an invariant number (180°). When Euclid's parallels postulate is removed, the sum of internal angles ceases to be an invariant. This is a simple example of how invariants are dependent on the theories that define them, even in mathematics. Likewise, Planck's constant, λ , is only valid within quantum mechanics, becoming meaningless when interpreted within, for example, Newtonian mechanics. The invariants from different theories may be incompatible, or irreducible to each other, or even devoid of meaning.

Table 2.1 Some examples of invariants in physics and mathematics

Example	Characteristics
Speed of light	Theoretical, postulate
Planck's constant	Theoretical, formally deduced
Avogadro's number	Empirical, relational
Thermal expansion coefficient	Empirical, idealization
Energy	Theoretical, relational, idealization
Gas constant	Empirical, relational
Sum of internal angles of a triangle	Theoretical, context-dependent

2.1.6 Invariants Prefigure Theories

Hence, invariants are bound by the assumptions of the theories that define them. Invariances are describable properties of a system that putatively remain unaltered irrespective of some classes of manipulations, or perturbations, or interventions in the system. Both in physics and in mathematics, invariants are the hallmarks of constancy, and often the axiomatic basis of a system of explanation, as in Einstein's theory of relativity. This has also been phrased as a theorem stating that whenever there is symmetry, a quantity is conserved [44]. When there is symmetry, variables are coupled, and their interdependence begets predictability and explanation. The theorem finds support in a vast number of examples in physics. Invariants prefigure theories, and often theories are born when invariants are discovered.²

2.2 Invariants in Biology

2.2.1 Biology Borrowed Invariants from Physics and Mathematics

How about biology? Modern biology motivated by the successes in physics has also made extensive use of invariants, more or less explicitly. In biological theories, invariants are as pervasive as – if less persuasive than, because they are less general – physical principles. In biology, mostly everything seems to be in constant change, contexts vary fluidly, boundaries are not well defined: there are few absolutes.

2.2.2 Invariants in Biology Are Context-Dependent

But there are some constancies that, although not absolute, subsist healthily within particular contexts given that some conditions for existence are met. Through them order can be retrieved. *Relational invariants* figure prominently in models of systems biology, of which there are several, including high-level concepts such as order, organization, hierarchy, and structure. Explanations in biology, and particularly in systems biology, gravitated towards concepts such as balance, equilibrium, and exchange. Many models have been accordingly proposed. Relational invariants in biology produce cycles, equilibrating processes, oscillatory reactions, and the like. Numerous biological models deal with networks of processes assessed for their stabilities and asymptotic states. As examples we have homeostasis (Ashby) [4], the hypercycle (Eigen and Schuster) [18], autopoiesis (Maturana and Varela) [63], and gene regulatory networks [37]. In these cases, constituents of cyclic processes are

² Noether's theorem informally stated: differentiable symmetries generated by local actions have correspondent conserved current. An English translation of the original paper is found in [45].

causally interrelated, where *this begets that, begets that*, and so on, in a systematized cycle of causes. One studies the bounds of operation of the networks, in terms of stability and in terms of the networks of interactions between constituents. The constancies resulting from networks of coupled processes are akin to self-organization, the idea that some arrangements of matter are endowed with certain equilibrating potentialities.

Because of the necessity of assuming particular contexts and outlining a subpart of the phenomenon, most invariants in biology have a more tentative, hypothetical nature. Many are pure abstractions, such as the hypercycle or homeostasis. The further biology becomes from physics and chemistry, the flimsier, the more context-dependent its invariants. They lose the status of principles and axioms as in physics, and become conceptual ancillaries for characterizing the boundaries and identity of a biological phenomenon.

Therefore, to employ invariants in biology, more than in physics, one is committed to a keen understanding of contextual influences, boundaries, and *ceteris paribus* assumptions. In biology, invariants dwell on very particular levels of description/analysis/explanation. Somewhat paradoxically, invariants in biology are most useful when contextual (and conceptual) boundaries are clearly discernible.

2.2.3 Three Examples: Life, Form, and Behavior

To illustrate this, I will discuss three examples of the search for invariants in biology: life, form, and behavior. Each example will bring in an important set of considerations, historical, logical, epistemological, and ontological. From those we derive conclusions that inform and specify the search for invariants of behavior, which shall occupy us for most of the rest of this book.

2.2.3.1 Life, as an Invariant of the Living, Is a Tautology

Some theories assume abstract invariant properties coextensive with what is to be explained. In this case, they become dry tautologies, and fail to provide an explanation. Such is the case, for instance, of life as an invariant property of the living.³ Tautologies, unless they can be framed as an axiomatic basis of a theoretical system, are dry in explanation. Difficult questions such as what is the fundamental property of the living led to many misconceptions, such as the supplanted idea of *elán vital* (Bergson) and *entelechies* (Driesch) [38]; in trying to distill the general property of life, too dilute a concept evaporates. This happens because the abstract property

³ *Coextensive*: If P is an invariant property of a system, then for all members of set A , it is true that A has P . In the example of life, if all the living A have life (property L), then $\forall A : L(A) \rightarrow L(A)$, a tautology. An explanatory principle such as this has vaporous foundation, and no logical derivations are possible. Nothing can be learned; life becomes magic.

(life) has too many instantiations (organisms); taken individually, each instantiation is fundamentally different, and therefore, life as a property can only refer to all instantiations of it when in a highly abstracted form. Attempting such overhauling levels of generality is disappointingly unproductive. The trouble with excess generalization in the search for invariants is that the result often becomes a self-referential tautology – devoid of predictive power – one with which one can have no hope of constructing a powerful theory.

2.2.3.2 Autopoiesis: An Invariant of Living Things

Tautologies will be explanatory when they can be made pillars of axiomatic theories, when they are self-evident truths at the foundation of explanations. If the tautology is not an end in itself, but may be composed of other tautologies, it can become a principle of a system of explanation. In *autopoiesis*, Maturana and Varela's theory of the living, the abstract invariant property "living," being defined only loosely, is translated into more explanatory relations that take the actual organization of an organism into consideration. Maturana and Varela propose that living should be defined in terms of cyclical processes: processes producing more processes like the originals exhibit the hallmark of the living [61–63]. It is this smart move that replaces the idea that life is something independent of the organisms. Life becomes the organization of processes which maintain themselves by recurrently producing themselves. This is the abstract property that becomes the explanatory invariant. In studying life then, instead of a mystical search for some extra factor that confers life, one focuses the search on the kinds of processes that confer autopoietic properties to the living. Consequently, the study of the living becomes the search for processes that maintain themselves metastably under perturbations. This is a valuable reduction, for it frames the approach at large, and determines the further set of questions to be addressed.

2.2.3.3 Genes as Invariants of Form

Many successful biological theories are founded upon invariances. Ever since Linnaeus, and perhaps Aristotle, it is a set of morphological traits of a particular species that defines its identity. It was clear that there must have been some reason for the invariance of morphogenesis, although the reason itself was unclear. The whole exercise of taxonomy, despite its pitfalls, strives for the most explanatory level of morphological invariance; those properties that congregated define a species identity [39]. The discovery of the inheritance mechanism being a crystal lodged in all the cells of an organism (DNA) infused biology with optimism. There was finally a reason for constancy of form and behavior. The whole endeavor of molecular phylogeny results from the realization that genomes can be taken as invariants of species. But for form and behavior, the history was not complete until the cell and the environment had been brought back into the morphogenesis equation.

2.2.3.4 Developmental Systems Theory Critique: Environment as a Determinant of Form

The idea of the gene as the quintessential invariant of ontogeny has recently been called into question by researchers [17, 22, 24, 27, 33, 34, 47, 54].⁴ Developmental systems theory has emphasized the role of reciprocal interactions with the environment in ontogeny, in an attempt to bring the context as an integral part of explanation. Modern theories emphasized the role of developmental processes, such as triggers of biochemical environment, cellular environment dependent protein production, energy consumption, intracellular and extracellular transport, and consumable materials. All of these help scaffold a form. If the genes are regarded as invariants of form, they are so only within very complex contextual frames.

2.2.4 *Ontogeny Needs the Environment*

Morphogenesis of some social insects, such as ants and termites, is a striking example of contextual dependency of form. It turns out that individuals with the very same genetic material, clones in every respect, will develop strikingly different morphologies and roles depending on the chemical context of the nest's chamber in which they are born. And those morphological differences will be further expressed in the individual's role in the colony, some becoming queens, some becoming winged soldiers, some becoming workers. Although not quite debunking the gene of its prominent role in generation of identity, developmental systems theory enforces contextual thinking in what may at first seem the stablest stance of invariance in biology. In the least, it becomes obvious that genes are only coresponsible for morphology and function. Because they are contextual, the genes are, at most, contingent invariants of form. In the extreme case, there are examples showing that even complex forms may be generated from a single cell, utterly devoid of genetic material, purely scaffolded by physiochemical processes [24].

2.2.5 *Development and Biophysics*

The biophysical category of explanatory principles include examples such as adhesion, viscosity, permeability, rigidity, and elasticity, and in general the relational forces generating potential for rearrangements and self-organization, and by extension the organizing of constancies. Many of these properties are only remotely connected to genetic prescription. Brian Goodwin [24] describes the interesting case of the unicellular macroscopic alga *Acetabularia*, which despite being only one cell

⁴ See also [15] for an insightful book review on Oyama's cycles of contingency.

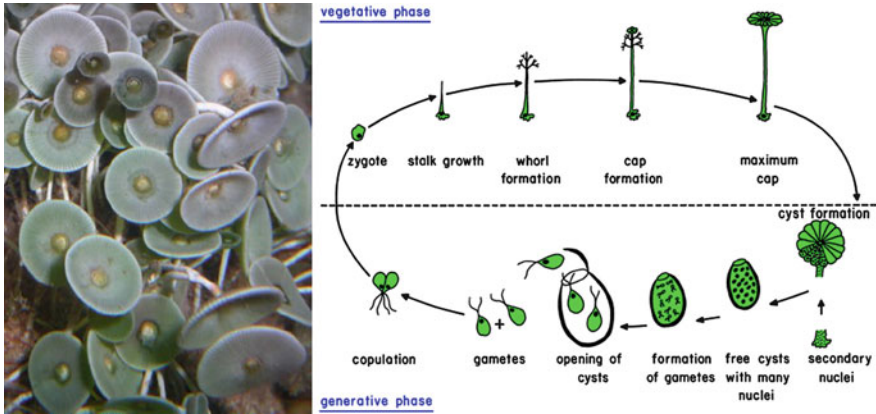


Fig. 2.1 *Acetabularia* and its life cycle

has complex form, resembling an umbrella (Fig. 2.1). Owing to complexity of its morphology, one would initially be inclined to attribute this complexity to genetic assembly. But a simple experiment precludes that conclusion. If the alga is cut at its stem, the nucleus of the cell remains at the bottom, presumably containing all genetic material with the information about the shape of the organism. There are now two parts, one with the nucleus at the bottom and no umbrella, and the other with no nucleus but with the umbrella. The surprising fact is that both parts regenerate a fully fledged version of the umbrella, irrespective of the presence of the nucleus. That is, the severed umbrella grows a second umbrella, at the severed side. So, the production of the material and the conditions that grow it are irrespective of the nucleus and its genetic material. Brière and Goodwin [25] presented a model of the scaffolding of *Acetabularia*'s tip (the umbrella), where the assumptions are the physical properties of the cell's substrate and time, excluding any further role for genetic production of material. This example illustrates well how complex form can originate solely from the physics of cell processes, even without gene transcription.

2.2.6 Evolution Operates on Organisms, Not on Genes

The appearance of an invariant trait or feature in an organism is not a function of genes alone. So, instead of absolute invariants of form, it is perhaps more adequate to frame genes as invariants of cellular processes, as in gene regulatory networks, because at this level we are much more competent in defining a process in terms of components, dimensions of interdependence, as well as susceptibility to contextual influences. It is more parsimonious to keep genes at a proximal sphere of influence. For the same reason, it is a category mistake (*sensu* Gilbert Ryle) to say that evolution operates on genes [14]; an idea that finds much receptivity in the

scientific community. Evolution is dependent on inheritance mechanisms; whether these are genetic or not is, in principle, irrelevant. Certainly, the gene pool correlates with evolutionary history, and has the potential of giving insight into the history of certain strains of beings. But it is a muddled conclusion that evolution operates on genes. Evolution operates on the level of the organism as a whole.

2.2.7 Genes as Invariants of Species Identity

Arguably, genes are at their best with *molecular phylogeny*, the search for the branches of *phyla* across time. Comparing genetic data from fossils with those of modern organisms, similarities across genes may help reconstruct the tree of life, tracing back mutations and thereby finding branches that belong together. To an extent, genes are defensible as invariants of phylogenetic identity. Even so, there are many examples where the evolution of adapted traits is essentially regulated by the environment, such as the adaptive shape and color of some seashells [26]. Moreover, there are many mechanisms of epigenetic inheritance, such as chromatin markings, which are only now being demonstrated [33]. Inheritance is about any transmitted trait; if cellular material transmitted to progeny determines features of progeny, then there is epigenetic inheritance. Owing to other forms of inheritance, molecular phylogeny is also susceptible to misclassifications, and genes are not fail-free invariants of species identity.

2.2.8 Invariants in Biology: Only Within Narrow Contextual Boundaries

As we proceed in our explanation, it turns out that most candidates for invariants in biology will follow this general rule; invariants in biology, if defensible, are so within narrow conceptual boundaries. Genes are the paradigmatic example. To say that the gene is an invariant of the form of an organism is to exclude too many formative levels. Genes are indeed crucial invariants of development and behavior, but are not almighty. By forcefully excluding the cell and the environment and their multifarious influences on the development of stable forms, the scope of explanation shrinks. In biology, it is counterproductive to try to attribute too much to one element of explanation. For the same reason, although genes and form do covary in certain experimental paradigms, it is a mistake to attribute form solely to genes.

2.2.9 Genetic Triggers Do Not Build an Organism

So, genes do not hold the position of absolute invariants of form, as the gene's causal prowess dilutes in development when the environment (also cellular) is taken into the morphogenesis equation. Molecular biology is at a stage where it can often

say *what* a gene triggers, but remains silent and circumspect to the question *how*. Molecular biology errs if it confuses causes and triggers, because although there are many ways to cause some sort of function – by knocking a gene, for example – there are fewer to build one. To explain how something is built, it does not suffice to say which gene triggers what, rather the story must be spelled out throughout the development cascade, up until the generation of a full organism. The organism is not made up of a set of genetic triggers, because triggers are not stuff [28, 55].

2.3 Invariants of Behavior

2.3.1 *What Is an Invariant of Behavior?*

Following on the definition of invariants, an invariant of behavior is that which remains constant with respect to an identified behavior. But what is it, and more crucially, is there such a thing? What is invariant when conspecifics execute similar motor behaviors? Further, will homologous behavior across *phyla* have similar invariances? Before these questions may even be asked, we must endure a demanding assumption: that it is possible to define functional behaviors.⁵ Start out with behaviors instinctive or stereotypical, preferably of simple organisms, and search for what may be the invariant concomitants of those. Think of a fly rubbing its hands, or a fish masticating, or a dog swallowing; simple behaviors. Ask what is invariant in those. At first glance, for the simple behaviors there are several candidates, of which three are discussed: genes, neuroanatomy, and reflexes.

2.3.2 *Genes as Invariants of Behavior*

2.3.3 *Genes Are Untenable as Invariants of Behavior, for They Are a Diluted Cause*

In Sect. 2.2.6, we discussed the role of the genome as an invariant of form. We found that although genes are certainly fundamental for phylogeny and inheritance, they are insufficient as invariants of form. Cellular processes underlying morphogenesis exist which are independent of the genes. Hence, the environment (also cellular) has a formative and causal role in the creation of a form, where not all processes are directly caused by the genes. The creation of a living organism is not merely a sequence of genetic triggers, but a complex scaffolding process of interdependencies. The same argument can be extended to include behavior. Enthusiastic with

⁵ Our working definition of behavioral function: when a goal can be attributed to a behavior, then the behavior is a functional behavior. Contrast that with behavior *in toto*.

the stability of genes, some have argued that the genome carries information, not only about species identity, but also a full program for the creation of a form, complete with modes of behavior.

At this point, however, we address a matter of scope. Genes certainly have a crucial role in the development of the structure of an organism, and the structure in turn will produce, with environmental interaction, behavior. However, genes as explanatory entities fail to give proper answers to the “how” question. Genes can be associated with “which” behavior or form, but they alone cannot answer “how.” To say that genes are invariants of behavior, therefore, is an unwarranted exaggeration; one that simply vaults over too many levels at a time.

A simple example will clarify the claim that genes are not the best level of analysis for behavior. Some species of fish are born without eyes. Genes have recently been found that trigger the formation of the eyes [43, 50]. If the eye-controlling motor neurons are present, then the fish uses them; so the presence of eyes modifies behavior. But the changes in behavior are better explained in terms of the newly acquired eyes than in terms of the gene that triggered the production of eyes. Behavior cannot be explained as a set of genetic triggers, for the genes do not see, the organism does.

The form of an organism and its modes of interaction with the environment are more immediate to behavior than genes. To be sure, the structure is a consequence of development, genetically guided, so through circuitous routes, one may attribute behavior to genes. But this is like saying that the Sarajevo crime caused the First World War. Although it may have triggered it, all the conjuncture that made it a propitious trigger was given at a much larger scale of interdependences. Triggers are descriptions of linchpin transitions, which require quite a bit of context to be activated. A set of genetic triggers does not explain a developmental process, as much as an organism is not a set of triggers. Therefore, to pin a gene as responsible for behavior is to give too summarized a story about the behavior. Important in their own right, it is a hopeless game to exhaustively list all the genetic triggers of behavior. Invariances of behavior are to be found in the engaged interaction between the being and the world (see Fig. 2.2).

2.3.4 Neuroanatomy: Invariant Connections Between Architecture and Behavior

That there must be a connection between neuroanatomy and behavior is nowadays beyond question (in the time of Aristotle it was not so, since he thought that the brain had a cooling function). Organisms of the same species with roughly the same behavioral traits also have, at certain scales, similar neuroanatomies. So, there are architectural characteristics that correlate with function and behavior. Occasionally, conspicuous features of similar neuroanatomies can be traced to function. However, this is the exception rather than the rule: although there are many distinguishable architectonic features, they do not always have obvious relations to function.

Braitenberg writes, “Very likely these quantitative relations reflect some general principles of the architecture of neuronal networks.” When quantified, many features of neuronal circuitry will display invariance in a statistical sense, such as average length of dendritic spines, average number of synapses in a given cortical area, and a long list of other features [6, 11, 19, 20].

The number of distinguishable levels of neural organization correlates with the complexity of the organism. Nervous systems of simpler organisms may have wiring that looks almost fully prespecified, as the wiring of insect ommatidia [32] (the fly is also a neat example, especially in the plates in [8]), whereas in complex organisms the degree of variation is much greater. Still, for both, there are gross⁶ levels of organization that appear highly patterned and repeatable (as in Fig. 2.3). As Braitenberg writes, this gives clues that, at some level, there must be some invariance of neuroanatomy responsible for constancy in behaviors.

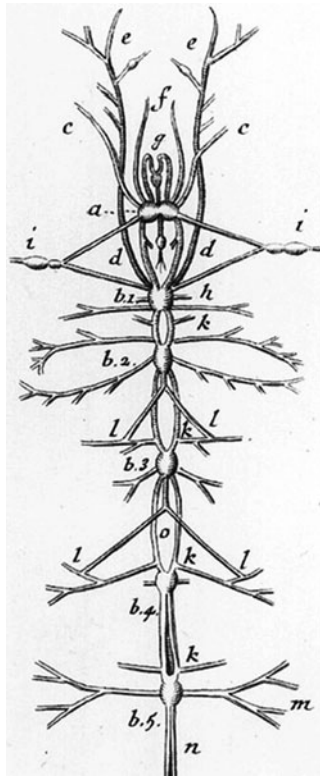


Fig. 2.3 Invariance at the coarse level of architecture in the nervous system of the caterpillar. The brain is marked with *a*. (From Swammerdam [57])

⁶ Gross in the sense of coarse, at first glance.

2.3.6 *Some Features of Gross Anatomy Can Be Traced to Behaviors*

As an example, Braitenberg [9], a neuroscientist by training, derived from neuroanatomical observations a model of behavior now known as a Braitenberg vehicle that abstracts salient (gross) features of neural architecture, such as contralaterality or ipsilaterality, to explain simple functional behaviors, such as following or avoiding a light (see Fig. 2.4). This is a simple and powerful instance of how simple features of architecture could be taken as invariants of behavior (provided, of course, there is the body and the environment).

Similarly, simpler organisms with only a few neurons (in the hundreds) may have almost identical wiring, to the extent that neuroscientists can find unique identifiers for every single neuron. At times, single neurons will be responsible for particular behaviors, such as escape responses, or simple reflexive behaviors, in which simple analysis of connectivity (sometimes one neuron, as in the case of zebra fish [58]) explains constancy in behavior. But there is no general method that relates architecture to function. Cortical layers, despite their similarity in construction, are often associated with very distinct functionality.

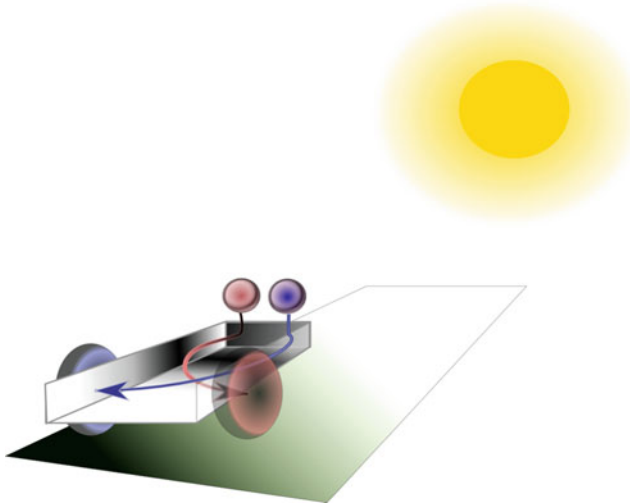


Fig. 2.4 A Braitenberg vehicle with phototropism. The contralateral connections establish negative feedback producing light-following behavior. Colors indicate contralateral pathways

2.3.7 Connections Between Anatomy and Function Do Not Always Exist

Such a clear connection between features of anatomy and functional behavior only exists for either the coarse level of the neuroanatomy or for very simple organisms. As architectures and organisms become more complex, this kind of direct relation between architecture and behavior becomes increasingly difficult to retrieve. Already in insects there is a considerable difficulty in tracing features of architecture to behaviors. Although many distinguishable regularities may be observed, it is not even clear whether for each of those there will be a function, or whether that architecture is the only way to realize the function. If it is the case that different architectonic features could express the very same analogous function, then an argument can be made against the case for anatomy as an invariant of behavior. An invariant of behavior, by definition, must be constant when the function is constant. So, if there are multiple neural implementations of a particular function, then those neural implementations *reflect* the invariant, but *are not* the invariant.

What is more, although we may describe the brain in very fine architectural features, the relationship between those features and function can be elusive. For historical reasons the anatomical characterization of the human brain is either independent of function or based on historical hypotheses (as in Broca's aphasia). Wernicke's areas, for example, are described in terms of the different cytoarchitectures and gross features of anatomy such as sulci and gyri. Even if, on occasion, some localization of function is attempted (see also Sect. 3.2.5 on functional magnetic resonance imaging and epistemic magnetic resonance imaging), the matching between function and structure inherits the arbitrariness and biases of the architectonic description. Incidentally, localizing function in the architecture informs little about how the function is executed (as will be discussed in Sect. 3.2.5), which is our primary interest.

2.3.8 Neuroanatomical Variation and Constancy of Function

If it is difficult to map constancy of function to architecture, which has an apparent order; it is even more difficult to account for the extreme neuroanatomical variation in complex organisms. Even in simple organisms there are many possible neuroanatomical implementations that express similar behaviors. In insects it is often a single neuron which realizes a complete pathway, and the variation across individuals is large. This suggests that although gross neuroanatomy *may* be taken as an invariant, it is so only at a coarse scale. One is hard-pressed to isolate features of geometry that are invariant with respect to function. However, this conclusion does service in constraining the search for an invariant of behavior, in that it adds one important requirement. An invariant of behavior has to give an account of which variations in neuroanatomy may express similar behaviors, and how (as we shall see in Part II).

2.3.9 *One to Very Many Mappings from Anatomy to Function*

Even if there is a level where gross neuroanatomy correlates with some simple functions (and there is enough reason to believe there is), the case is very different for finer neuroanatomical details and more sophisticated behavior. The problem is that for more sophisticated brains, mapping between function and neuroanatomy is one to *very* many. In one individual, one neural architecture maps to all the functions an organism potentially expresses. Therefore, except for very particular cases, neuroanatomical features cannot explain the breadth of behavior. As organisms become more sophisticated, they express a widening breadth of behavior [56], not necessarily traceable to gross features of the architecture. In a sense, the architecture “lodges” all possible behaviors an organism can express, and hence is not very informative about specific behaviors. This a strong indication that neuroanatomy is not the most adequate level to search for invariants of behavior.

This many-to-one function-to-structure mapping is manifested in current neuroscientific literature, where particular anatomical structures are hypothesized to perform a number of functions, often incommensurable or even contradictory. To take an extreme example, the hippocampus has been associated with a large number of cognitive and behavioral functions. Examples include localization [46], declarative memory [13], emotion and motivation [59], face recognition [51], sequencing and planning [7], and language production [53].

This nonexhaustive list of functions attributed to the hippocampus betrays subjectivity in the attribution of its role, and more than that, it befuddles brain theorists as to whether it is possible to attribute to the hippocampus a unique function subsuming all the instances of function above. As Braitenberg puts it (p. 7 in [8]) (my emphasis):

We will not be able, in most cases, to explain the peculiarities of a certain brain structure by invoking the rules and constraints of the mechanism that synthesizes brains out of neurons, but will always have to consider explanations *in terms of the function it performs*.

The problem is that for any behavior that has the hippocampus as an active component, the role of the hippocampus may change, dynamically and within behavioral context, introducing a difficulty for any static functional label.⁷

2.3.10 *Requirements for an Invariant of Behavior*

A theory of behavior certainly has to explain how constancies in the gross architecture level lead to constant function, but it cannot stop there. We may derive some

⁷ Schema theory (see more in Sect. 2.3.12) exemplifies the attempts to overlay the brain with a structure of functional components, “boxes,” to aide in explanation of overall behavior. But although being a tenable level of functional analysis, it encounters the problems of arbitrariness. That is because, in an important sense, a function only exists while it is being executed, and only within the frame of a theory. Often, there will be no fundamental way to distinguish between different overlaying schemas which one is better or best.

requirements for an invariant of behavior, in that (1) it has to explain the appearance of gross features of neuroanatomy, (2) it has to explain its “capacity for behavior,” the fact that it may lodge a large number of “latent behaviors,” and (3) it has to explain how constancy in function appears despite variability in neuroanatomy.

Much like genes, neuroanatomy subserves behavior but does not fully exhaust its explanation. The fact is that behavior exploits neuroanatomy in the same way that it employs the body, by being embedded in a context. For a Braitenberg vehicle to follow a light, there must be, in addition to the neural architecture, also a light source and a body that turns to it. Invariants of function rely on contingencies of the environment and are only definable in those terms. The behaviors expressible by the Braitenberg vehicle are only so when proper stimuli are present, which usually requires that the environment should be introduced into the explanation.

Reflex theorists such as Skinner, Pavlov, Thorndike, and Herrnstein took this to heart. They proposed that the environment could be distilled into a number of stimuli, each evoking an action, where the sum of all reflex–stimulus pairs could make up an explanation of behavior. This was soon shown to be too reduced a picture to encompass the multitude of adaptive behavior [40].

2.3.10.1 Reflexes as Invariants of Behavior

Although it has not been framed as such (a search for invariance), the most prominent example of the search for invariants in behavior is reflex theory. It departed from the observations that some stimuli were stereotypically associated with particular motor responses. At its inception, it seemed like a productive front of research, infused with optimism, but as the research program developed, dissatisfaction arose in the community. It soon became obvious that an organism is more than a finite list of stimulus–response reflexes [40].

Reflex theory is attractive because it hints at a possible systematic search of stimulus–response couplings. It may have caused some researches to believe incorrectly that it is possible to compile a comprehensive, if not complete, list of stimulus–response pairs. Summarizing Merleau-Ponty, the animal is more than the sum of its behaviors. The invariant reflex is too circumscribed, too contingent on precise definition of *ceteris paribus* conditions, to succeed as a theory of behavior. But in spite of its shortcomings, it was instrumental in the search for mechanisms behind behavior. The search for an exhaustive list of reflexes was an original and insightful attempt to explain the sources for the apparent invariances in behavior, albeit somewhat rudimentary.

2.3.10.2 Limitations of Reflex Theory

Reflex theory faces two problems, one epistemological, one empirical. Empirical, because the laboratory setup is a far cry from natural behavior. As Merleau-Ponty (pp. 44–46 in [40]) writes, “few pure reflexes are found [because] laboratory reflexes

are not *biological realities*.” Epistemological, because there was no obvious way to reintegrate the individual stimulus–reflexes in a whole picture of behavior. This highlighted the ludicrous nature of the original project of reflex theory. Because the environment is immensely complex, it is not possible to exhaustively list stimuli–reaction pairs. It is also not possible to assume a null context, because every stimulus presentation has a different context (even if that difference is merely time elapsed).

2.3.11 *Cybernetics, Reafferece, and Sensorimotor Loops*

The nervous system and the automatic machine are fundamentally alike in that they are devices which make decisions on the basis of decisions they made in the past.

Norbert Wiener, *Cybernetics*

There would be no adequate framework to resolve the quandaries from reflex theory until the appearance of cybernetics. Cyberneticists realized that single stimulus–reaction pairs were insufficient to explain behavior. Theirs was the insight that if one adds *feedback* from the environment, then a recurrent structure emerges in time, where a stimulus–reaction pair entails the next stimulus–reaction pair. The organism then has an active/positive role in producing its own stimuli, and so on, in a cycle. Behavior appears as stimulus–reactions are sequentially and smoothly chained. Moreover, cyberneticists acknowledged that the organism had internal states, which rendered it a “nontrivial machine,” *sensu* von Foerster [21].

That suggested that the structure of the nervous system was organized as a “negative feedback with a purpose” (see Sect. 8.5.3). Mittelstadt, von Holst [31], Ashby [4], Arbib [1, 2], Wiener [52, 64], and others [9, 23, 60] provided a rich set of examples of organisms whose neural structures could execute the function of negative-feedback controllers. Examples ranged from the flight of a fly, to the eye movements of frogs. Their seminal contribution spawned many fields of research [30].

2.3.11.1 **Cybernetics Extends Reflex Theory**

As in Sect. 2.3.10, although in some cases the underpinnings of reflexes could be traced to anatomy and network structures, in most cases behavior still defied a mechanistic explanation. Although an important start had been made, at the time a general theory of behavior was still missing. That would be the case until cybernetics introduced the concepts of reafferece, recursion, and the sensorimotor feedback. The notable contributions of cybernetics provided new ground towards an explanation for the immense variation of complex nervous systems.

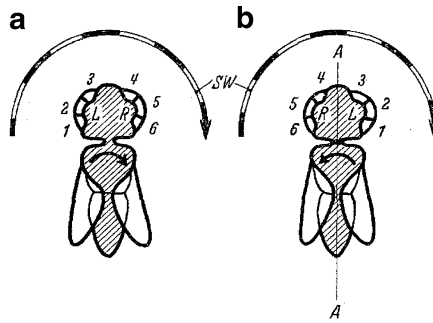


Fig. 1a u. b. Verhalten eines Insekts (*Eristalis*), um das eine Streifenwand *SW* nach rechts bewegt wird. a normales Tier; b nach Drehung des Kopfes um die Achse *A—A* um 180° , *R*, *L* rechtes, linkes Auge (die Ommatidien sind numeriert). Der Pfeil im Tier zeigt dessen aktive Drehreaktion an.

Fig. 2.5 An example of an explanation based on refference. (From [31])

Despite much variation in neural structure, constancy in behavior was too pervasive for the concept of reflex to be relinquished altogether. Admittedly incomplete, on the upside it provided a stepping stone to the explanation of behavior and invariances, passive of further research on how specific mechanisms could underlie multifarious behavior. Cybernetic views enhanced reflex theory as refference theory, and suggested a path for a more sophisticated take on a theory of invariances in behavior Fig. 2.5. It was to be searched for in the engaged interplay between body, nervous system, and environment.

2.3.11.2 Functions, Mechanisms, and Purpose

The principal insight of cybernetics was to suggest that the organism could be regarded as a complex mechanism whose functions emerge as the organism *becomes* (*sensu* Varela and von Foerster) in its environment. Wiener and contemporary cyberneticists proposed that purposeful behavior is the outcome of a mechanism; that which a machine is built to do. It presupposes an abstract principle of construction – a mechanism – that entails function when in a sensorimotor loop. A machine that acts and corrects towards a goal is behaving purposefully, and its mechanism(s) for doing so is(are) the behavior invariant of that goal. In that reductionist sense, the simplest example in the literature of a machine with a purpose is a thermostat. Its construction is such that it compares two temperature values, a reference with a measured value, and sends corrective signals (which may be heating, or cooling), to reach a temperature equilibrium. More than an analogy, this was proposed as purposeful behavior. Many more such examples exist, as the watt regulator, or radar control, where negative feedback is the operating principle. Distasteful to some [36], the proposal that an organism acts as a feedback mechanism has profound implications. For if a certain mechanism entails a certain function, then it is the mechanism

itself which is the invariant of function. The mechanism is, so to speak, coextensive with its functions. Moreover, it acknowledges the role of the environment as an inextricable part of function: for the function to be exerted, the environmental context must be given.⁸

The thermostat is the standard example of a cybernetic device: a system in the world, endowed with sensors and actuators that enable it to make feedback corrections to attain its goal. In that, it is similar in description to an organism that pursues a willed goal by deploying a set of actions and corrections to achieve it. The thermostat, however, has a goal in a rather narrow sense; in fact, it has only one goal, about which it is not cognizant. If, for example, the window is opened, the thermostat may huff and puff, but will not do more than keep the heater (cooler) on (off). Hence, compunctions to call its behavior multitudinal (or even adaptive) are more than justified. Multitudinal and adaptive behavior are hallmarks of organismic behavior of even insignificantly small mammals that do all sorts of cool stuff, such as finding a mate, finding a home, playing, fleeing, growing, finding food, killing food, and feeding, and all that flexibly, adaptively, and purposefully.

Hans Jonas, in contemptuous disagreement, was quick to point out the analogy of purposeful behavior as a mechanism is misapplied [36]. Concerning the relationship between goal and purpose, in a mechanism, goal is that what the mechanism is built to achieve. To construct a machine is an exercise of defining a goal and constructing mechanisms to achieve it. An organism, on the other hand, seems to define its own goals. Its behavior not only exhibits purpose, but is also guided by it. That is the problem Jonas stresses: a machine has its goals attributed to it by us, an organism defines its own goals by existing and being intrinsically motivated (the concept of mediacy; pp. 106–114 in [36]). Opposite to a machine, an organism is not a set of stapled functions. Rather, an organism entails functions, which are exerted or not depending on the internal context and external affordances. Goals for the organism appear as the organism behaves, and as it does it changes, and consequently its goals change, and so on recursively.

One may not discard Jonas's criticisms lightly, but a closer look may help bring the two, machine and organism, together. First, we have it as a given that a mechanism may be taken as an analogy to an organism; after all, we initially built machines to accomplish purposes often copying those of organisms. We should not forget, as Jonas reminds us, that they are indeed our purposes, not the machine's.

But as a model for behavior, a cybernetic device is more than a cold machine. It outlines a set of prescriptions to assemble purposeful-like behavior, with a savory small set of assumptions. And it does not take too much dialectic effort to bring purpose back into the explanation, not as an analogy, but something which is embedded in the control loop, generating the specifications for behavior – even if it fails to show how the drives feel to their owner. My purpose here is to show that a mechanism is a correct, albeit hopelessly incomplete, model of the purposes, or

⁸ Environmental context is meant broadly: a cell can be an environment for DNA, as blood can be an environment for a cell, or the body can be a contextual environment for a neural system.

drives, or urges, of an organism. My reasons for the incompleteness claims are the same as Merleau-Ponty's: "to build a model of an organism would be to build the organism itself" (p. 151 in [40], also Wiener [64]). This does not mean that we should give up building models, but that our models are not organisms – although they may help us understand the functions of an organism. And ultimately, it is not precluded that we will be able to assemble such a resourceful contrivance that will *act* as we do, although it may not have the same type of (organismic, motivational) reasons for its drives.

If we bring evolution into the picture, the issue gains perspective. Goals appear as evolutionary viability imposes order on self-organizing structures, endowed with intrinsic potentialities. In evolution, mechanistic, functional behavior bestows organisms with enhanced viability. It is an orderly exploit of chaos, a selection, that what begets function. Life evolves towards function because function enhances viability. Very simple function, mechanical behavior, such as chemotaxis in a bacterium, was described as purposeful long before the advent of cybernetics as a theory. A bacterium as a complex organism has the purposeful behavior of a mechanism. All the behaviors a bacterium can express are latent in its structure, and are expressed when it encounters contexts, the environment. To the extent that we can disinter the mechanistic underpinnings of the behaviors of a bacterium, we can understand what its mechanisms are. Because an organism is a mechanism coextensive with all its potentials for functional behavior, a quest for function is fulfilled with the description of a mechanism. But the problem is a little deeper than that.

2.3.12 *Schema Theory and Functional Overlays*

If we are compelled by the idea that behavior is subserved by a dynamical concoction of mechanisms operating in concert and context, the simplest way to study brain function is by finding a match between mechanisms and functions. One intuitive method would be to *overlay* the neural structures with hypothetical functional boxes, whose actions would produce the analyzed functional behavior. This method is essentially a product of systems theory applied to the study of neural function. The idea that function can be localized to particular substructures has been prolifically employed by cognitive scientists and neuroscientists. Unfortunately, this appealing project suffers from similar critiques as reflex theory. Whereas reflex theory falls short of explaining how the sum of functions builds the whole behavior, schema theory strains to show how brain areas shift between functions. As the same brain area may subserve different functions at different times (and even simultaneously), a static attribution of function to brain area is bound to be partial and incomplete.

A functional overlay is a metaphor of localization of function, in which an idealized anatomical map of a brain, often hierarchical, is overlaid with functional boxes representing a mapping between brain areas and functions (an example is given in Fig. 2.6). The functional boxes are usually connected with arrows representing some kind of *well-formed information exchange*. The boxes are in fact black boxes

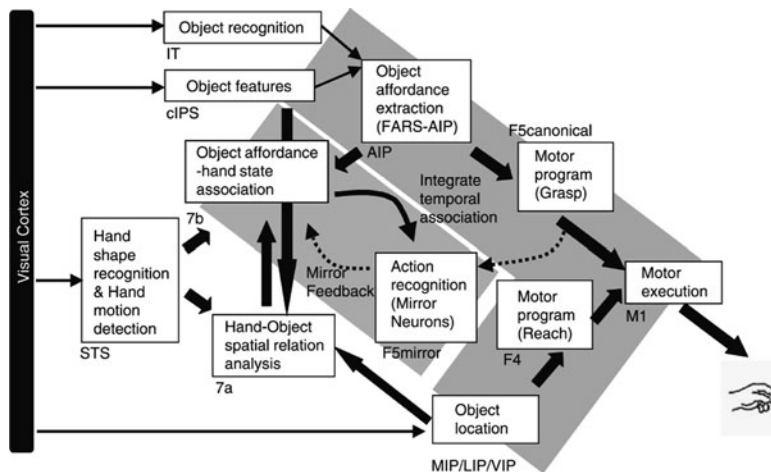


Fig. 2.6 A schema as an overlay of functional boxes on brain areas. This particular schema proposes explaining the functional roles of brain areas involved in action recognition. Abbreviations: *IT* – intratemporal cortex; *cIPS* – caudal Infra Parietal Sulcus; *AIP* – anterior intra parietal cortex; *STS* – Superior Temporal Sulcus; *MIP/LIP/VIP* – medial, lateral, ventral intra parietal cortex. (From [48])

of function, implemented neurally somehow – perhaps instantiated with models, or with models pending. A structure of such boxes representing function–area associations is able to generate predictions about interdependencies, for example, in lesion studies. Outlining functional boxes is a basal tool for theory (and hypotheses) making, and in some incarnations is called schema theory [1–3, 16]. A schema is a relational structure of functional boxes standing for an explanation of a certain behavioral or cognitive ability.

Hence, systems theory applied to brain function incurs a few critical difficulties. First, a function is a relational abstraction, subjectively outlined as in [5, 42]. Second, as discussed in Sect. 2.3.9, the same neural structure, such as the thalamus or the hippocampus (can there be a unique function for these structures?), may underlie different functions (also simultaneously!). Third, (vide Sects. 3.3 and 10.1–3) a neural structure may participate in two different functions in distinct ways. Fourth, to any functional question (what for?), there will be answers (for *x*), good or bad, as I will analyze in more depth in Chap. 3 about empirical assessments of invariance. Fifth, (as in Chap. 8.5.3) there is compelling evidence indicating that brain function is dynamically organized: a function is the product of dynamics that only really exist while the function is executed (else only as a potential). These points are individually addressed in the sections mentioned, and lead to the upcoming conclusion that these epistemological difficulties appear because the brain is a *hermeneutic machine*, one that can only be comprehended in terms of function, which only exists (in the eyes of the theory) once it is named. One finds (or does not), the function one names.

A functional overlay map is similar to a labeled and tagged map used for navigation, where depending on one's interests, restaurants, hotels, and entertainment venues, would be highlighted, depending on the search. A functional overlay is like a meta-data-laden navigation map, where a restaurant and a hotel may occupy the same place simultaneously. The analogy is wanting, because although restaurants and hotels exist even when not shown, functions are only there when they are executed. For instance, it is not productive to analyze object recognition in the absence of the object, nor to analyze speech production in the absence of words.

2.3.12.1 Descriptions of Schema Functions Are Abstractions Over Instances

When a neuroscientist asks questions about cognitive or behavioral function, he or she asks questions about memory, perception, understanding, abstraction, categorization, attention, judgment, decision, volition, and learning. Essentially these abilities are abstract clusters derived from single instances. All these functions have a relational structure, and all these functions are "of" something. As Brentano [12] pointed out (in 1874), for all these cognitive and behavioral functions, there are objects. Memory is "of" some event, perception is "of" some thing, and so forth. The functions listed above are abstract categories that cluster a large number of instances in which similar purpose is recognized. Descriptions of function derive from categorizations across instances which inherit properties of the individual instances, which may be contradictory or even mutually exclusive. This is the source of many a misunderstanding in the search for mechanisms underlying function. Hence, the search for a mechanism implementing a function stumbles upon a problem equivalent to Minsky's frame problem in artificial intelligence [41]. The more detailed the description becomes, the more it excludes. Between excessive abstraction and excessive specificity schema theories of behavior fail to display the flexibility necessary to explain the relations between structure and function.

The important lesson to be drawn is that the search for brain mechanisms must fully embrace the multifunctional character of the brain, as it dynamically shifts from function to function. The explanation of a brain function remains the search for a mechanism, but one that lodges functions only as potentials until they are put to use. An explanation of organismic function based on mechanisms and functional boxes has to explain more than a single function at a time; it must also address the multifarious potentialities for function that one neural structure may deploy, in different contexts and moments. Because the functional role of the neural substrate is often dynamically organized, and thus only exists in the context of behavior, box structure overlays become invalid in different moments.

Dynamical systems explains how this can be the case, as structures assume different roles in different contexts, vis-à-vis dynamical parameterizations. Invariants of behavior arise within dynamical contexts, and admit only very flexible functional overlays. At the expense of sounding repetitive, behavioral functions are entailments of an organism. Schema theory confuses *component functions* with *de facto* structural functionality.

2.3.13 Behavioral Function and Invariants of Behavior

A *functional behavior* is a *phenomenon*, in the proper sense of the term: the object of a person's perception; what the senses notice or the mind notices [35]. Functional behavior is *immanent* in organisms; whereas observers describe and categorize behavior (after Merleau-Ponty). When purposes and goals can be attributed to a set of actions, this set of actions can be outlined as a functional behavior. Consider an organism as a complex mechanism with the potential for many *component functions*. Deploying these capabilities in a concerted manner to achieve a particular goal creates an example of *functional behavior*. The potential for behavior is fulfilled in interaction with the environment, in context. An *invariant of behavior* appears with the minimum set of mechanisms with the inherent potential for those functions indispensable for a functional behavior. Incidentally, akin to other invariants of biology, the idea of an invariant of behavior obviates subjectivity in the definition of function and context.

Behaviors have functions as prerequisites. Consequently, the search for an invariant of behavior begins with the description of a behavior, parsed into component functions necessary for that behavior. The act of outlining the functional components of behavior is, by necessity, a subjective activity. Correspondingly, functional descriptions may be contradictory or fuzzy – the best heuristics being parsimony; as when attempting to avoid excessive anthropomorphism, notwithstanding its ineluctability. In a rough description, a behavioral function of a fly such as (*pursuing a mate*), requires components to (*localize a moving dot in visual field*) and (*flight control*). Those descriptions are shorthand implicitly assuming the existence of mechanisms that enable these functions (note, moreover, that the mechanism underlying both functions may be one and the same).

The invariant of a behavior is the set of mechanisms fundamentally necessary for a behavioral function. In that, it is inextricable from the descriptions of the component functions. It implies the required functions for one functional behavior, and defines the outline of mechanisms to be explained.

In reflex theory the atom of explanation is a stimulus–reaction pair. In my (constructivistic) conception, the atom of explanation is the association of a functional behavior (such as feeding), with all its component functions (e.g., the capacity for hunger, foraging, mastication), and the minimal substrate that can implement that behavioral function, within an organismically relevant frame.

The invariant of behavior is a many-to-one association from behavioral function to mechanism(s). Because functions can be implemented differently, the substrates implementing an invariant of behavior may vary widely. The less similarity between organisms, the less likelihood that the mechanisms will be similar in implementation. Conversely, the more similar the organisms, the more probable that the implementation will evince similarities in mechanisms. Evidently, the mechanisms leading to feeding behavior in an elephant are more similar to those of a hippopotamus than to those of an ant. Still, even between conspecifics there will be variation in implementations. That which remains the same across compared organisms is the

invariant of behavior, which may be a null set (e.g., a palm and an elephant) or an equivalent set (e.g., two worker ants).

Different organisms with similar capabilities for behavioral function will in one way or another implement the same component functions. Thus, findings about behavioral invariants from one organism may tell tales about the selection of component functions of the behavior of different organisms. In that sense, invariants of behavior can be conceptualized as the overlapping regions of constancy, across the variable implementations of functional mechanisms across organisms. Invariants of behavior are conceptual shorthand for the definition of behavioral function and its possible implementations/instantiations. Invariants of behavior map to the space of (abstract) mechanisms with the potential for behavioral function.⁹

What is similar between an elephant feeding and an ant feeding? Simple answer: the description. Feeding is the function of acquiring energy through ingestion of food. Quite obviously, an ant and an elephant will have quite different mechanisms for feeding, but both share the description. This is the difficulty of the study of many a cognitive function with model organisms, the description is shared, but the mechanism is not. The main problem with that conception is also the necessary frame for any solution: language. A function is a description of a purpose. In the elephant as in the ant, the description for feeding is the same, because the goal is the same. At the abstract level of teleological analogies, the invariant of behavior is a description of that behavior and the possible analogies between component functions/mechanisms [23]. I participated recently in a round table of neuroethologists which included specialists in mammals, fish, and insects. At the event, results of operant conditioning in *Drosophila* were presented. In the paradigm, *Drosophila* flies were made to couple an unconditioned stimulus with a conditioned stimulus. The source of disagreement was the number of trials until the response of the *Drosophila* was conditioned. Learning took place in very few trials (two to seven). Mammal experts pointed out that this process is much shorter than in mammals, where operant conditioning takes many more trials (more than 50) until the conditioned response is exhibited. Consequently, the mammal experts argued (with some support from fish experts) that what happened in *Drosophila* could not be called operant conditioning. Much to the dismay of the insect experts, who pointed out that apart from the number of trials, the descriptions of the experiments were essentially the same.

The lesson to be learned from this event is that descriptions are dangerous. I take it that the problem behind the disagreement was that nothing learned about operant conditioning in insects could be imported to mammals. All the same, because the mechanisms are likely to be fundamentally different, a certain strictness concerning terminology is utterly justified.

Component functions of a behavioral function may be more or less dependent on each other, and this dependence may vary between different organisms. Therefore, mechanisms may be more or less *modularizable* than others, reflected in the

⁹ This is one of the main difficulties with analogies in ethology and neuroethology. Descriptions of functions are shared across different organisms, whereas mechanisms are not.

behavioral invariants. Lesion experiments test for that. In monolithic mechanisms, component functions may share the substrate with which they are implemented – the same neural structure may underlie two simultaneous functions. As an example, certain cephalopods use their visual system both to guide escape reactions and to select skin patterns for camouflage [29]. Because the visual systems underlie both abilities, damage to the common substrate impairs both behavioral functions. Although the substrate is the same, it may be employed by different subsystems in distinctive modes (as the experiment described in Sect. 10.3.2 will illustrate).

2.3.13.1 Invariants of Behavior from Convergent Evolution

Because organisms share the world, along with some basic physics and biochemistry, some behavioral functions may be stumbled upon anew in different evolutionary lines. The phenomenon denominated “convergent evolution” may lead to similarities of form and of behavior, or both. A conception of evolution as procuring a maximum denominator in behavioral function may rely on invariants of behavior for a test bed of hypotheses. Invariants of behavior may be defined within individuals, across conspecifics, and beyond phyla. The amounts of overlap are informative about similar problems different organisms may have encountered in their evolutionary paths, leading to convergence in behavioral function, and structure.



Invariants of behavior are proposed as a conceptual shorthand in the understanding of the multifarious behavior of organisms. In this chapter I exemplified how invariants of behavior affect the search for a mechanism for behavioral function. The procedure is simple: name a behavioral function, partition the behavioral function into necessary component functions, and finally, search for the minimum set of mechanisms with the capacity to subserve the behavioral function named. Invariances appearing as problems are solved similarly. It is in this sense that invariants of behavior will be henceforth employed.

2.4 Conclusion

At the expense of having belabored the obvious, I argued among other things that to compose a powerful theory of one has to find those invariants that are most telling about a phenomenon. When aiming at an explanation of behavior, there are many stances not fully suitable for the task. It is unproductive, for example, to say that genes are invariants of behavior. Although that may be true, it is also very coarse. Unless genes show the mechanisms of behavior, they are improper candidates for explanatory invariants; likewise for neuroanatomy regarded in isolation. Neurons, networks, and their properties vary widely, even in conspecifics; and yet, most functions remain unaltered. The space of variation may have to be defined “outside” the

range of neuroanatomy – for instance, in dynamical terms. So the most adequate level of description sees neuroanatomy as subserving the dynamics that mediate the environment and organism.

The derivation of lawlike relationships from correlations is a class of inferences common in biology (this is a neutral statement of fact). Nonetheless, the considerable difficulties in measurement and control conditions must be acknowledged. Biologists are often forced – often owing to empirical limitations – to abduct relations between variables that may be quite distant in reality. Circumventing the difficulty by quantifying relations between behavior and neurons, or behavior and genes, does not quite solve the problem. Behavior and genes may be related, but not immediately related. Between them there is too big a gap, even for a leap of faith. It is better to build bridges. That bridge is constituted by the environment and the body as links between genes, neurons, and behavior.

This chapter contained an analysis of epistemological aspects involved in a search for invariance in behavior. The following questions were asked: What is an invariant? What forms does it take? What is the nature of invariance in the study of behavior? What forms do an invariant take? How have theories employed invariants? The next chapter deals with invariants arising from empirical measurements, and how invariant function is measured in brains. Essentially, all empirical methods of brain measurement rely on experimental paradigms that attempt to isolate the invariants in brain function. In neurophysiology, invariants are used to indicate the relevant dimensions of change of a phenomenon. If something simply does not change when something else changes, no distinction can be drawn, hence nothing can be learned. At the other extreme, if it varies too much, it is also not very enlightening. Science operates on distinctions and on gradations of change and relationships between variables. Therefore, the susceptibility of a hypothetical invariant to change of context offers a qualitative assessment of the quality of that invariant. The relevance of an invariant is assessed in terms of its dimensions of change. Thus, the next chapter discusses methods of empirical assessments of invariance in brain and behavior.

2.5 Summary

The argument of this (somewhat philosophical) chapter is as follows:

- Invariants are that what remains when something else changes (Sect. 2).
- Invariants underlie theories and are pervasive in physics and mathematics, as well as in biology (Sect. 2.1.1).
- Invariants come in different forms, fitting assorted categories. We named five: theoretical, empirical, relational, idealization, context-dependent (Sect. 2).

- Invariants in biology are context-dependent, genes are taken as an example. Although genes help build an organism, and thus may be framed (somewhat narrowly) as invariants of form, they do not offer an appropriate level of analysis for behavior, for behavior appears in the engaged interaction between an agent and the environment (Sect. 2.3.2).
- Similarly, neuroanatomical features can be taken as invariants of functional behavior, either in simple organisms or at the level of gross architecture, but they fail to address three crucial aspects (Sect. 2.3.4).
 1. The mapping between neuroanatomical features and function is not always clear.
 2. One neuroanatomy subserves many behaviors.
 3. Different neuroanatomies subserve similar behaviors.
- The reflex is regarded as a hypothetical invariant of behavior. Reflex theory assumed that behavior was composed of stimulus–reflex pairs (Sect. 2.3.10).
 1. However, behavior is more than concatenated stimulus–reflex pairs.
 2. Moreover, a characterization of behavioral function is inherently subjective (Sect. 2.3.9).
- For cybernetics, purposeful behaviors result from the interaction of complex feedback mechanisms (reafferences) with the environment (Sect. 2.3.11).
- An invariant of behavior is the set of the possible mechanistic implementations of a behavioral function (Sect. 2.3.13).

References

1. Arbib MA (1972) *The metaphorical brain, an introduction to cybernetics and brain theory*. MIT Press, Cambridge
2. Arbib MA (1982) *Machine Intelligence 10*, Chichester: Ellis Horwood, chap Rana Computatrix, an evolving model of visuomotor coordination in frog and toad, pp 501–517
3. Arbib MA, Fellous JM (2004) Emotions: From brain to robot. *Trends Cogn Sci* 8(12)
4. Ashby W (1960) *Design for a brain: The origin of adaptive behavior*, 2nd edn. Chapman & Hall, London
5. Bateson G (1972) *Steps to an ecology of mind*. University of Chicago Press, p. 533
6. Benavides-Piccione R, Hamzei-Sichani F, Ballesteros-Yanez I, DeFelipe J, Yuste R (2006) Dendritic size of pyramidal neurons differs among mouse cortical regions. *Cereb Cortex* 16(7):990–1001
7. Berns GS, Sejnowski TJ (1998) A computational model of how the basal ganglia produce sequences. *J Cogn Neurosci* 10(1):108–121
8. Braitenberg V (1977) *On the texture of brains: An introduction to neuroanatomy for the cybernetically minded*. Springer, New York

9. Braitenberg V (1984) *Vehicles, experiments in synthetic psychology*. Bradford Book, Cambridge
10. Braitenberg V (2001) Brain size and number of neurons: An exercise in synthetic neuro-anatomy. *J. Comput. Neurosci.* 10(1):71–77
11. Braitenberg V, Schüz A (1998) *Cortex: Statistics and geometry of neuronal connectivity*. Springer, Berlin
12. Brentano FC (1874) *Psychologie vom empirischen Standpunkte*. Duncker & Humblot, Leipzig
13. Cohen N, Squire L (1980) Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* 210(4466):207
14. Dawkins R (1976) *The selfish gene*. Oxford University Press, New York
15. Di Paolo E (2002) Book review: *Cycles of contingency*. *Artif Life* 8(2)
16. Dörner D (1999) *Bauplan für eine Seele*. Rowohlt, Reinbek
17. Edelman G (1988) *Topobiology: An introduction to molecular embryology*. Basic Books
18. Eigen M, Schuster P (1978) The hypercycle. *Naturwissenschaften* 65(1):7–41
19. Elston G, Rockland K (2002) The pyramidal cell of the sensorimotor cortex of the Macaque Monkey: Phenotypic variation. *Cereb Cortex* 12(10):1071–1078
20. Elston GN (2005) Cortex, cognition and the cell: New insights into the pyramidal neuron and prefrontal function. *Cereb Cortex* 13(11):1124–1238
21. von Foerster H, von Glaserfeld E (2005) *Einführung in den Konstruktivismus*, 9th edn. Piper Press, Munich
22. Fox Keller E, Harel D (2007) Beyond the Gene. *PLoS ONE* 2(11):e1231
23. Glaserfeld Ev (1990) Teleology and the concepts of causation. *Philosophica* 46(2):17–42
24. Goodwin B (2001) *The evolution of complexity: How the leopard changed its spots*. Princeton Academic, Princeton
25. Goodwin B, Briere C (1989) A mathematical model of cytoskeletal dynamics and morphogenesis in acetabularia. *The Cytoskeleton of the Algae*. CRC Press, Boca Raton, pp 219–238
26. Gould SJ, Lewontin RD (1979) The spandrels of San Marcos and the Panglossian paradigm: A critic of the adaptationist programme. *Proc R Soc Lond* 205:581–598
27. Griffiths PE, Gray RD (2000) *Darwinism and developmental systems*. MIT Press, Cambridge
28. Griffiths PE, Stoltz K (2007) *The cambridge companion to the philosophy of biology*. chap Gene, Cambridge University Press, Cambridge, pp 103–119
29. Hanlon R (2007) Cephalopod dynamic camouflage. *Curr Biol* 17(11):400–404
30. Heylighen F, Joslyn C (2001) Cybernetics and second-order cybernetics. In: Meyers R (ed) *Encyclopedia of Physical Science and Technology*, 3rd edn. Academic, New York
31. von Holst VE, Mittelstaedt H (1950) Das Reafferenzprinzip. *Die Naturwiss* 37(20):464–476
32. Homberg U, Paech A (2002) Ultrastructure and orientation of ommatidia in the dorsal rim area of the locust compound eye. *Arthropod Struct Dev* 30(4):271–280
33. Jablonka E, Lamb M (2005) *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. MIT Press, Cambridge
34. Jablonka E, Lamb M, Avital E (1998) ‘lamarckian’ mechanisms in darwinian evolution. *Trends Ecol Evol* 13(5):206–210
35. Jewell E, Abate F, McKean E (2001) *The new Oxford American dictionary*. Oxford University Press, Oxford
36. Jonas H (2001 (1966)) *The phenomenon of life*. Northwestern University Press, Evanston, IL
37. Kauffman S (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22(3):437–67
38. Mayr E (1961) Cause and Effect in Biology Kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science* 134(3489):1501–1506
39. Mayr E (1976) *Evolution and the diversity of life*. Harvard University Press, Cambridge
40. Merleau-Ponty M (1963 (translation), 1942) *The Structure of Behavior*. Duquesne University Press, Philadelphia
41. Minsky M (1975) *The psychology of computer vision*, chap A Framework for representing knowledge. McGraw-Hill, New York
42. Nagel E (1979) *The structure of science: Problems in the logic of scientific explanation*. Hackett Publishing, USA

43. Niven J (2008) Evolution: Convergent eye losses in fishy circumstances. *Curr Biol* 18(1):27–29
44. Noether E (1918) Invariante variationsprobleme. *Gott Nachr* 235
45. Noether E, Tavel M (2005) Invariant variation problems. Arxiv preprint physics/0503066
46. O’Keefe J, Dostrovsky J (1971) The hippocampus as spatial map: preliminary evidence from unit activity in the freely moving rat. *Brain Res* 34:171–175
47. Oyama S (2000) *The ontogeny of information: Developmental systems and evolution*. Duke University Press, Durham
48. Oztop E, Kawato M, Arbib M (2006) Mirror neurons and imitation: A computationally guided review. *Neural Netw* 19(3):254–271
49. Pais A (1982) *Subtle is the Lord. The science and the life of A. Einstein*. Oxford University Press, Oxford
50. Porter J, Baker R (1997) Absence of oculomotor and trochlear motoneurons leads to altered extraocular muscle development in the Wnt-1 null mutant mouse. *Dev Brain Res* 100(1): 121–126
51. Quiroga R, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435(7045):1102–1107
52. Rosenblueth A, Wiener N, Bigelow J (1943) Behavior, purpose and teleology. *Philos Sci* 10: 18–24
53. Ryan L, Cox C, Hayes SM, Nadel L (2008) Hippocampal activation during episodic and semantic memory retrieval: Comparing category production and category cued recall. *Neuropsychologia* 46(8):2109–2121, DOI <http://dx.doi.org/10.1016/j.neuropsychologia.2008.02.030>, URL <http://dx.doi.org/10.1016/j.neuropsychologia.2008.02.030>
54. Smith JM, Burrian R, Kauffmann S, Alberch P, Campbell J, Goodwin B, Lande L, Raul D, Wolpert L (1985) Developmental constraints and evolution. *Q Rev Biol* 60(3):265–287
55. Smith PG (2007) *The cambridge companion to the philosophy of biology*. Cambridge University Press, Cambridge, chap Information in Biology, pp 103–119
56. Sterelny K (2005) *Thought in a hostile world*. MIT Press, Cambridge
57. Swammerdam J (1737) *Biblia Naturae, Sive Historia Insecto*, vol 1. IDC (Leiden)
58. Ton R, Hackett J (1984) Neural mechanisms of startle behavior, Springer, Berlin, chap The Role of the Mauthner Cell in fast starts involving escape in Teleost Fishes
59. Tracy A, Jarrard L, Davidson T (2001) The hippocampus and motivation revisited: appetite and activity. *Behav Brain Res* 127(1–2):13–23
60. Turchin VF (1977) *The Phenomenon of Science: a cybernetic approach to human evolution*. Electronic URL <http://pespmc1.vub.ac.be/POSBOOK.html>
61. Varela F (1979) *Principles of biological autonomy*. North Holland, New York
62. Varela F, Maturana H (1987, 1998) *The tree of knowledge*, 1st edn. Shambala, Boston, MA
63. Varela F, Maturana H, Uribe R (1974) Autopoiesis: the organization of living systems, its characterization and a model. *Curr Model Biol* 5(4):187–96
64. Wiener N (1961) *Cybernetics: or the control and communication in the animal and the machine*, 2nd edn. MIT Press, Cambridge



<http://www.springer.com/978-1-4419-8803-4>

Invariants of Behavior

Constancy and Variability in Neural Systems

Negrello, M.

2011, XX, 252 p., Hardcover

ISBN: 978-1-4419-8803-4