Philosophia Scientiæ

Travaux d'histoire et de philosophie des sciences

19-1 | 2015

Logic and Philosophy of Science in Nancy (II)

From Malfunction to Mechanism

Bertold Schweitzer

Electronic version

URL: <http://journals.openedition.org/philosophiascientiae/1034>

DOI: 10.4000/philosophiascientiae.1034

ISSN: 1775-4283

Publisher

Éditions Kimé

Printed version

Date of publication: 1 March 2015

Number of pages: 21-34

ISBN: 978-2-84174-707-8

ISSN: 1281-2463

Electronic reference

Bertold Schweitzer, « From Malfunction to Mechanism », *Philosophia Scientiæ* [Online], 19-1 | 2015, Online since 21 April 2015, connection on 03 November 2020. URL : <http://journals.openedition.org/philosophiascientiae/1034> ; DOI : <https://doi.org/10.4000/philosophiascientiae.1034>

Tous droits réservés

From Malfunction to Mechanism

Bertold Schweitzer

University of Dundee (United Kingdom)

Résumé : Le dysfonctionnement, la défectuosité et les erreurs sont des phénomènes qui peuvent aider à mieux comprendre les entités affectées par ces incidents. L'analyse approfondie du fonctionnement et du dysfonctionnement facilite en particulier la découverte, l'explication et la modélisation théorique des structures, fonctions et mécanismes propres à un système. Dans certains cas, les dysfonctionnements sont le seul moyen d'accéder aux processus internes d'un certain système. Le présent essai analyse les méthodes de diverses disciplines tenant compte de dysfonctionnements tels que les mutations, les illusions optiques, les erreurs d'action ou les lapsus linguistiques et identifie les motifs communs aux stratégies de recherche basées sur le dysfonctionnement. Ces stratégies contribuent à la décomposition de systèmes, à la reconstruction de séquences d'événements, à la localisation de sous-systèmes, ainsi qu'au développement et au test d'hypothèses mécanistes.

Abstract: Malfunctions, deficits, and errors provide considerable insight into key features of the entities in which they occur. In particular, the careful analysis of patterns of functioning and malfunctioning facilitate discovery, explanation and theorizing about structure, function, and underlying mechanisms of a system. In some cases, malfunctions even supply the unique probe into the internal workings of a system. This essay analyzes methods used by various disciplines involving malfunctions such as mutations, visual illusions, action slips, or speech errors, and identifies common patterns of malfunction-based research strategies that contribute to the decomposition of systems, the tracing of sequences of events, the localization of subsystems, and the generation and testing of complex mechanistic hypotheses.

1 Introduction

This paper advances the idea of exploring the intersection, largely neglected thus far, of two fields of research—on *malfunction* and on *mechanism*—and of combining the more practical and usually discipline-specific results of how to gain insights about normal structure and function from the analysis of malfunctions with the more theoretical and systematic understanding of how to examine mechanisms in general.

Malfunctions—including deficits, errors, anomalies, breakdowns, or failures¹—which are observed in the behaviour of systems may not only contribute to understanding how and why the system is malfunctioning, i.e., diagnosis, or the means by which a normal, functioning state could be restored, i.e., therapy, but also to discover the underlying causes, structures, and mechanisms of the systems affected, both in their functioning *and* malfunctioning states.

Thus, the analysis of malfunctions also contributes to theoretical scientific understanding by discovering and testing hypotheses about a system's structure, its components and their relations, in short, its mechanism. Such development and testing of theoretical insight concerning normal functioning and its underlying processes and mechanisms based on the analysis of malfunctions can be identified in approaches used in a number of scientific fields, or disciplines. Still, such ideas have not been addressed in general or interdisciplinary methodological studies thus far.

On the other hand, a considerable body of literature exists in philosophy of science as well as in natural and social sciences discussing the general methods and strategies for identifying *mechanisms* responsible for a system's behaviour [Machamer, Darden *et al.* 2000], [Darden 2008], [Hedström & Ylikoski 2010]. The notion of using analysis of malfunctions as an instrument in aid of such strategies is at best mentioned in passing but has not been analyzed systematically yet.

Accordingly, the main questions addressed in this paper are: First, in what ways can the analysis of malfunctions, failures, deficits, errors, and related phenomena contribute to scientific research, or to understanding the world? Second, in how far are research strategies involving malfunctions, or inferences from malfunction to normal function, methodologically sound and yield reliable scientific conclusions? Third, how can the analysis of malfunctions help in identifying causal connections and unraveling complex causal mechanisms?

These questions arise because malfunctions are *de facto* being analyzed, either systematically, or on an *ad hoc* basis, in many different disciplines, from biology and psychology to linguistics and sociology. If coupled with methodological reflections at all, malfunctions have often been praised for of-

1. Throughout this article, the term “malfunction” will be used as a umbrella term for all of these related concepts. A definition for “malfunction” will be provided in section 3.

fering valuable and sometimes unique insight into the subject of investigation. Fromkin points out that

Speech error data do [...] provide us with a “window” into linguistic mental processes and provide, to some extent, the laboratory data needed in linguistics. [Fromkin 1973, 43]

However, concerns have been voiced as well, pointing out that inferences from malfunction to normal function are by no means straightforward: Gregory comes to the conclusion that “[...] arguments from malfunction to normal function are fraught with difficulties even for quite simple machines” [Gregory 1981, 85]. Moreover, general doubts as to the usefulness of the analysis of malfunctions have been expressed. It has even been claimed that inferences based on malfunctions might be outright misleading.

Since a general, systematic analysis of malfunction-based research strategies has not been undertaken yet, this paper shall try to provide some arguments to evaluate these competing claims concerning the role and value of malfunctions in order to close this gap in philosophy of science.

This paper will try to show, first, that the identification and analysis of malfunctions provide an opportunity for insights into the normal operation of the entities in which they occur; second, that the most typical features, aims, and results of malfunction-based research strategies lie in the decomposition of systems, the identification of sequences and chains of events, the localization of subsystems; culminating in the elucidation of complex mechanisms; third, that malfunction-based research strategies follow a certain common pattern, or logic; fourth, that this logic connects well with independently developed and commonly used strategies for the discovery of mechanisms and testing of mechanistic hypotheses; and, finally, that malfunction-based research strategies are indeed unique and valuable, in that they are neither self-reliantly guaranteeing correct analyses nor wholly misleading; instead they can and should be used, in particular in a heuristic role, but should be backed up, wherever possible, by complementing methods in order to obtain reliable conclusions.

The account offered here is based on description, comparative analysis and rational reconstruction of empirical research actually carried out in various scientific disciplines. The general patterns emerging from this comparison are in turn related to independently developed accounts of general strategies for analyzing mechanisms. Moreover, the general validity of inferences from malfunction to function is evaluated in order to determine to what extent these kinds of inference can indeed be rationally justified.

This paper's overall argument will be structured as follows: section 2 presents examples of malfunctions pointing out how their analysis contributes to theoretical understanding and scientific knowledge. Section 3 provides a more specific explication of the concept of malfunction. Section 4 explains typical features, aims, and results of malfunction-based research strategies. Section 5 discusses strategies for discovery and analysis of mechanisms and

relates them with malfunction-based approaches. Section 6 offers general theoretical conclusions and implications.

2 Malfunctions in use

Some instructive examples of how observation and analysis of malfunctions can advance research come from the field of behavioural studies:

If [...] the observer of captive animals sees the way in which a young wolf or dog carries a bone to behind the dining room drapes, lays it down there, scrapes violently for a while next to the bone, pushes the bone with his nose to the place where all the scraping was done and then, again with his nose and now squeaking along the surface of the parquetry flooring, shoves the nonexistent earth back into the hole that has not been dug and goes away satisfied, the observer knows quite a lot about the phylogenetic program of the behavior pattern. [Lorenz 1981, 48]

Thus, the researcher gains theoretical insight from the observation of a characteristic pattern of malfunction. This shows convincingly, in particular if the young wolf or dog has never had the opportunity of digging holes in the ground before, not only that innate components of behaviour exist, but also which particular patterns they assume without, or before being influenced by the environment.

Another discipline where malfunctions contribute to theoretical knowledge is genetics. Here, too, the observation of malfunctions and tracing these back to underlying processes, in particular, genetic mutations, plays an important role. One striking example is the *antennapedia* mutant of the fruit fly *Drosophila*. Here, normal individuals (the so-called “wild type”) have antennae on their forehead which, in the *antennapedia* mutant, are replaced by legs. This observation and the subsequent genetic analysis ultimately led to the discovery of the homeobox, a DNA sequence regulating patterns of anatomical development in animals, fungi and plants [McGinnis, Levine *et al.* 1984].

A third field where malfunctions have contributed strongly to theoretical insight is the psychology of perception. Here, optical illusions have played a major role in both demonstrating the imperfections of perception while at the same time assisting in tracing them back to the details of perceptive mechanisms. The Hermann grid illusion, e.g., misleads the eye into perceiving grey blobs at the intersections of a white grid on a black background, with the blobs disappearing when focused directly. This illusion can be explained by a process of lateral inhibition involving nerve cells in the retina: ganglion cells receive information from several photoreceptive cells, with stimuli from the centre of its “receptive field” exciting a ganglion cell and stimuli from its periphery inhibiting it. At an intersection of white lines, a ganglion cell receives

stronger stimuli from the periphery than its neighbouring cells, leading to a comparatively weaker excitation of the ganglion cell, and resulting in the perception of this spot as less bright than the surrounding white areas [Frisby & Stone 2010]. Thus, insights from the demonstrable malfunctioning of our visual sense lead to hypotheses about the structures and processes underlying normal vision.

3 Characterizing malfunction

After having introduced the phenomenon “malfunction” by a few examples, can we be more specific about how it should be characterized in the present context? Five considerations seem important:

First, the term *malfunction*, both in its everyday sense and when discussed as a tool for scientific research, is usually understood as describing any case of *something going wrong*. This, of course, requires the possibility of comparing such a case with a background of cases where something similar is “working well” or “working as expected”. Thus, the term “malfunction” can sensibly only be used in cases where a corresponding function is either known, or at least assumed hypothetically. Moreover, analyses of malfunctions typically require the comparison of one or several *different* variants of malfunction with a normal function. Therefore, the comparison of malfunction with function involves the comparison, within the same general type, of one or more “functioning” tokens with one or more “malfunctioning” tokens.

Second, malfunction may be defined in a roundabout fashion as any kind of absence of or deficiency in some function. This definition seems straightforward, but of course in turn requires an appropriate definition of “function”. This, however, is complicated by the fact that two different, contested definitions of “function” have been proposed. One such definition is based on an *etiological* view, looking to the history of an item to define its function as its contribution to the survival and reproduction, or in other words, the fitness of an organism it is part of [Millikan 1984]. Artifacts, in this view, may derive functions in case they contribute to the fitness of organisms. The competing definition is based on a *causal role* view, defining function as the causal contribution of some entity towards the capacity or operation of a larger system containing this entity [Cummins 1975].

The etiological definition of function seems narrower and more precise. However, it denies the status of function to many phenomena for which the term is commonly used. In particular, we would never be able to decide whether we can legitimately consider something to be a function unless the history of the item in question had been considered. Basing a definition of malfunction on an etiological definition would also prevent us from using “malfunction” in many cases where researchers themselves use this term, or closely related ones.

In what follows, therefore, the definition of “malfunction” will be based on the causal role view since this promises to provide a suitably inclusive concept of function. Accordingly, “malfunction” will be defined as any case in which a “function”, understood as a causal contribution of some entity towards the capacity or operation of a larger system containing this entity, and understood as being present in “normal” cases, or tokens of a type, is absent, weakened, or modified in a way that the original function is no longer present.

Third, it should be emphasized that labelling a phenomenon as a “malfunction”, in particular when humans are concerned, should never be understood in an evaluative or pejorative sense.

Fourth, these considerations clarify that the *scope* of any malfunction-based strategy is limited to fields where functions and malfunctions—at least in the wider *causal role* sense—can be associated with the target items. Thus, we may expect malfunction-based strategies to play a role in biology, psychology, linguistics, and social sciences as well as in sciences of artefacts, such as technology, but not, e.g., in physics.

Fifth, malfunction-based research strategies need to be sharply distinguished from all kinds of strategies based on “learning from error”, or “trial and error”, or falsification, where the main goal is identifying and rejecting erroneous hypotheses. In malfunction-based research strategies, the main focus is on the malfunctions of target items, not on the (potential) failure of hypotheses *about* these items. Of course, researchers search for potential errors in their hypotheses and theories, and sometimes analyses of such errors do indeed contribute to finding “truer” theories [Wimsatt 1987]. This, however is still different from employing malfunction-based research strategies to theories. This would amount to analyzing malfunctions of theories in order to achieve deeper understanding of how theories work in general, i.e., theory structure and theory dynamics. This question, though it appears to be a promising one, is not the subject of this article.

4 Inference from malfunction

Now, let us consider somewhat more in-depth which discoveries, or inferences can typically be achieved by using malfunction-based research strategies. Six of these stand out: first, discovery of the existence of hidden systems and sub-systems; second, decomposition of a system into modules; third, identification of sequences and chains of events in processes; fourth, physical localization of modules inside a system; and fifth, contributions to the testing of mechanistic hypotheses.

First, spontaneously occurring malfunctions can draw attention to the existence of systems, or substructures of systems that have hitherto escaped researchers’ notice. The existence of constancy mechanisms, e.g., might not be obvious: being used to a constant room temperature, only the failure of

a thermostat might draw attention to the fact that a complex control system is required to maintain a constant temperature. Moreover, the existence of substructures with an item performing a complex function might only become apparent when specific parts of the overall function are lost while others continue to work as usual. One example are the various specific impairments that can result from brain injuries, such as prosopagnosia, or face blindness, where the ability to recognize faces is severely reduced, while the ability to recognize other objects remains relatively intact. In this way, malfunctions may provide a unique insight into structures that are normally hidden, or transparent, and cannot be observed in or inferred from normal functioning systems alone.

Second, decomposition, or fractionation of a system into discernible functional or structural subsystems, or modules is often among the initial contributions an analysis of malfunctions can provide. This step moves beyond the mere recognition of the existence of a substructure to identifying and enumerating parts of substructures, usually by comparing different types of malfunctions.

One example from biology is the determination of the number of genes involved in a biological process: assuming that two different recessive mutations have been identified that both result in the same observable malfunction (e.g., some aberrant behaviour), then crossing experiments combining the two mutations in one organism may show these two mutants to either complement each other, resulting in normal, or “wild-type” behaviour, in which case the mutants may be inferred to be alleles of two different genes, or the two mutants might not complement each other, leading to aberrant behaviour, in which case they are inferred to be defective in the same gene. The first case, therefore, leads to the conclusion that at least two different genes are involved in the process, whereas only one gene can be discerned in the second case.

Another example is the identification of cognitive and neuronal modules in the human brain. Brain injuries, strokes, or other processes may lead to the loss of specific capabilities. One example is prosopagnosia, or face blindness, mentioned above. Such “dissociations” are used in (cognitive) neuropsychology to identify the number and the characteristics of functional modules.

Critics of this approach, however, have voiced concerns about the reliability of this method. Inferences from associations (with the loss of two functions commonly occurring together) or single dissociations (where one specific function is lost while others continue to function normally) are indeed unreliable. For instance, an associated loss of two capabilities might not happen due to functional commonalities but because of functionally irrelevant anatomical reasons, such as dependence on a common blood supply. So-called “double dissociations”, however, where one capability, *A*, can be lost while *B* continues to function normally, but where also *B* can disappear while *A* is still intact, are seen by a majority of cognitive neuropsychologists as allowing reliable inference to the existence of two different modules [Jones 1983; Dunn & Kirsner 2003].

An even more fundamental concern is that malfunction-based strategies might also break down in case the target system's structure turns out to be quite unlike a hierarchical mechanism, as for instance in a connectionist system [Bechtel & Richardson 1993]. To what extent, however, purely connectionist systems could exhibit, when damaged, for instance classical forms of double dissociations is not clear and should be investigated.

Third, malfunctions contribute to the determination of the *order of steps* in sequential processes. One example from biology is the identification of the sequence of steps involved in biochemical processes, e.g., the biosynthesis of the amino acid arginine. Here, a number of different mutants with defects in different genes are used, none of which can grow without arginine. Since mutations are blocking different enzymatic reactions in the metabolic pathway for the synthesis of arginine, different intermediate metabolic products synthesized *before* each defective step can accumulate in cells. Mutants with defects *earlier* in the process can be cross-fed by mutants with defects *later* in the pathway, since *later* mutants can provide *earlier* ones with substances needed for growth [Stanier, Ingraham *et al.* 1987]. Thus, analyzing the patterns of possible cross-feeding indicates the sequence of genes and biochemical steps involved.

A fourth type of inference allowed by an analysis of malfunctions is localization: by correlating data on malfunctions with information on sites of physical damage, *physical localization* of the corresponding function may be achieved. An example from neuropsychology illustrates this: based largely on lesion studies in monkeys, two main routes for processing visual information have been distinguished. One route analyzes "the physical properties of a visual object (such as its size, color, texture and shape)" and constitutes the so-called "what" system, whereas an anatomically different route is involved in "perception of spatial relations among objects, and not in their intrinsic qualities", called the "where" system [Mishkin, Ungerleider *et al.* 1983].

This type of localization in particular has been criticized for being too simplistic:

The effects of removing or modifying, say, the line scan time-base of a television receiver would be incomprehensible if we did not know the engineering principles involved. Further, it seems unlikely that we should discover the necessary principles from scratch simply by pulling bits out of television sets, or stimulating bits with various voltages and wave forms. [Gregory 1961, 320–322]

To others, this seems exaggerated:

From time to time it has been argued that one can learn nothing of interest about an intact system when it is broken. The argument often runs like this: given a radio, if you cut off the plug no sound will come out; you would consequently conclude that the plug was the source of the sound. Certainly the village idiot might draw such an inference [...],

however, actual scientific research can avoid such naivety:

No more than any other field of research is the domain of human deficit studies a haven for village idiots. [Kean 1984]

Finally, malfunctions can also offer additional and, in some cases, more severe ways of testing hypotheses compared with tests based on normal functioning alone. For example, a plausible demand is that models should not only show the same performance, but also commit the same malfunctions as the target system, since this would be an indication that the model does not only superficially simulate the target system's performance but brings about by the same means. In a similar vein, the "goal of speech-error research is [...] to identify, for particular issues of [...] theory, the particular errors or error classes which can provide relevant evidence" and "to find the perfect speech error" [Cutler 1988, 219].

5 Integrating mechanistic and malfunction-based strategies

Next, after pointing out characteristic contributions of, or inferences from malfunctions, it appears useful to relate these contributions in a systematic fashion to more general research strategies, in other words, to integrate malfunction-based strategies into more general research strategies.

Mechanistic research strategies, here, appear to be the ones malfunction-based strategies could be most easily related to and integrated into.

Thinking in terms of mechanisms is considered to be fruitful, and establishing mechanistic hypotheses as well as providing mechanistic explanations is seen as an important goal: "At least in biology, most scientists see their work as explaining types of phenomena by discovering mechanisms" [Wimsatt 1974]. The same holds for many other disciplines, such as psychology, social sciences, and technology, see [Bunge 1997; Mayntz 2004]. The following definition captures what seems close to a consensus:

A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism, manifested in patterns of change over time in properties of its parts and operations, is responsible for one or more phenomena. [Bechtel & Abrahamsen 2005, 423]

Thus, mechanistic research is concerned with functions, and while malfunctions are rarely being explicitly considered in this body of literature, the relation with malfunctions seems straightforward, if only because "function" is involved. However, the connections, as we will see, reach much further than that.

Strategies for the discovery of mechanisms have been discussed in much detail, see [Bechtel & Richardson 1993], [Machamer, Darden *et al.* 2000], [Darden 2008], often distinguishing construction, evaluation and revision of mechanist hypotheses [Darden 2006]. Here, we will focus first on strategies for the “construction” of causal mechanistic hypotheses, in particular “activation” and “modification” strategies [Craver 2002], many motifs of which correspond closely with methods used in the analysis of malfunctions.

Activation strategies for the discovery of mechanisms are based on modifying the environment of an item, hoping to evoke a reaction or a change in behaviour. Being able to provoke malfunctions in behaviour is one of the most significant outcomes of activation experiments since this shows that conditions necessary for normal functioning have been affected. Conversely, the observation of a malfunction often triggers a search for changes in the environment possibly responsible for this malfunction. Thus, malfunctions are often significant outcomes of either planned or natural activation experiments.

In *modification strategies* used in discovery of mechanisms, the internal setup of items is modified by the experimenter. They can be divided into interference strategies, also called subtractive strategies, and additive strategies.

In mechanistic *interference* or *subtractive strategies* some component of an item is diminished, retarded, eliminated, disabled, or destroyed, and the effect or effects downstream or on the system-level are registered [Craver 2002]. Examples of such active manipulations include deletion (“knockout”) of genes, ablation of organs, inactivation of brain regions by chemicals or by cooling, and many others. Quite frequently, observed malfunctions can be traced back to spontaneous or accidental modifications equivalent to such subtractive manipulations.

In mechanistic *additive strategies*, components are stimulated, augmented, intensified, or multiplied, and the effect or effects downstream or on the system-level registered [Craver 2002]. Examples are manifold: providing an excess supply of active biochemical substances, multiplying the number of copies of genes, or stimulating neurons. Again, perceived malfunctions can often be related to spontaneous or accidental modifications equivalent to such additive manipulations.

Thus, malfunctions quite often arise as a consequence of spontaneous or accidental analogues of subtractive or additive experimental interventions. Such situations amount to natural experiments in which subtraction or addition occur without interference by an experimenter.

With regard to spontaneous malfunctions as well as phenomena resulting from planned experiments, the exact modification or intervention causing the phenomenon may or may not be known, or obvious; in the latter case it needs to be identified or provisionally conjectured. Identifying the cause of an observed malfunction is of course not trivial. One reason is that malfunctions can occur on all hierarchy levels, and the underlying intervention, or spontaneously occurring modification may happen on the same, or a different hierarchy level.

As with all natural experiments, one major disadvantage of relying of either spontaneously occurring or randomly provoked malfunction is that the means of controlling both occurrence and the circumstances are missing or very limited. A unique advantage, however, is that malfunctions occur without the need for a researcher having good ideas of how and where to intervene by subtractive or additive manipulations. A second advantage malfunctions provide is that their spontaneous occurrence might allow an analysis even where direct manipulation is impossible (e.g., in classical genetics before molecular methods existed) or morally forbidden (e.g., harmful experiments on humans).

6 Conclusion

In conclusion, the role of malfunctions in strategies for analyzing mechanisms may be summarized as follows: malfunctions are phenomena that are often surprising, raise attention, and provide information about subsystems, or modules hidden inside other systems. Since many malfunctions are conspicuous or run contrary to researchers' expectations, they often provide both excellent starting points for and easily detectable markers in investigations. At the same time, many malfunctions are exactly the same kind of phenomena that are sought after, and that experimenters try to produce in common general setups for probing mechanisms, such as subtractive and additive strategies.

Typically, malfunction-based research strategies are involved in the generation of mechanistic hypothesis as well as in testing. They contribute by attracting attention to unknown systems or aspects of systems, or by revealing the hidden existence of a system by one of its malfunctions. They may also supply specific information on key features of a system, or provide access to data difficult or impossible to obtain by other means. Finally, they may enable reliable inference to a comprehensive model of the target system; in particular if complete profiles of all possible functions and malfunctions are available. For most of these features of malfunction-based analysis, the roles they play in mechanistic strategies could be identified. This helped to shed light on the exact kind of contribution malfunctions can provide for the discovery and testing of causal mechanistic strategies and how their analysis can complement the methods—mainly experimental—usually considered in mechanistic strategies so far.

Using malfunction-based research strategies alone usually cannot guarantee reliable results, not unlike many other strategies used in isolation. At least two important issues may arise when relying on data from malfunctions only: first, patterns of malfunctions may suggest the identification of subsystems, or modules, or the fractionation of a system. Erroneous fractionation, however, might result from secondary commonalities or differences between modules, irrelevant for the primary questions about structure and function. Second, malfunction-based strategies might break down when a system's structure is

radically unlike a modular mechanistic system. Thus, malfunction-based research strategies benefit from the integration of additional evidence—a recommendation in which they are not alone and which is apt for other, e.g., mechanistic strategies, too.

These recommendations include searching for separate and independent evidence concerning existence and features of hypothetically assumed entities, activities, or modules, trying to provide support for experimental data by paying attention to control, repeatability, and randomization, gaining additional support by using either techniques or results from multiple fields, and finally attempting integration of individual mechanisms into the larger matrix of scientific knowledge.

All arguments considered, malfunctions and malfunction-based strategies are indeed capable of contributing greatly to scientific research, in particular to the elucidation of mechanisms. The major contribution of malfunction-based strategies, due to their important role in discovery and their strengths in suggesting hypotheses, may be seen in the area of heuristics. However, unfolding their strengths in combination with other adequate strategies, in particular mechanistic ones, malfunction-based strategies do also contribute towards conclusively evaluating hypotheses, and successively refining them. Thus, malfunctions and their analysis should be seen as a very valuable, and sometimes even indispensable component of scientists' methodological toolkit.

Bibliography

- BECHTEL, William & ABRAHAMSEN, Adele [2005], Explanation: A mechanist alternative, *Studies in History and Philosophy of Science – Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421–441, doi:10.1016/j.shpsc.2005.03.010.
- BECHTEL, William & RICHARDSON, Robert C. [1993], *Discovering Complexity*, Princeton, NJ: Princeton University Press.
- BUNGE, Mario [1997], Mechanism and explanation, *Philosophy of the Social Sciences*, 27(4), 410–465, doi:10.1177/004839319702700402.
- CRAVER, Carl F. [2002], Interlevel experiments and multilevel mechanisms in the neuroscience of memory, *Philosophy of Science*, 69(S3), S83–S97.
- CUMMINS, Robert [1975], Functional analysis, *Journal of Philosophy*, 72, 741–765.
- CUTLER, Anne [1988], The perfect speech error, in: *Language, Speech and Mind: Studies in Honor of Victoria Fromkin*, edited by L. M. Hyman & C. N. Li, London; New York: Routledge, 209–223.

- DARDEN, Lindley [2006], *Reasoning in Biological Discoveries*, New York: Cambridge University Press.
- [2008], Thinking again about biological mechanisms, *Philosophy of Science*, 75(5), 958–969, doi:10.1086/594538.
- DUNN, John C. & KIRSNER, Kim [2003], What can we infer from double dissociations?, *Cortex*, 39(1), 1–7, doi:10.1016/S0010-9452(08)70070-4.
- FRISBY, John P. & STONE, James V. [2010], *Seeing: The Computational Approach to Biological Vision*, Cambridge, MA: MIT Press.
- FROMKIN, Victoria A. (ed.) [1973], *Speech Errors as Linguistic Evidence*, The Hague; Paris: Mouton.
- GREGORY, Richard L. [1961], The brain as an engineering problem, in: *Current Problems in Animal Behaviour*, edited by William H. Thorpe & Oliver L. Zangwill, Cambridge: Cambridge University Press, 307–330.
- [1981], *Mind in Science*, London: Weidenfeld & Nicolson.
- HEDSTRÖM, Peter & YLIKOSKI, Petri [2010], Causal mechanisms in the social sciences, *Annual Review of Sociology*, 36(1), 49–67, doi:10.1146/annurev.soc.012809.102632.
- JONES, Gregory V. [1983], On double dissociation of function, *Neuropsychologia*, 21, 397–400.
- KEAN, Marie-Louise [1984], Linguistic analysis of aphasic syndromes, in: *Biological Perspectives on Language*, edited by D. Caplan, A. R. Lecours, & A. Smith, Cambridge, MA; London: MIT Press, 130–140.
- LORENZ, Konrad [1981], *The Foundations of Ethology*, New York: Springer.
- MACHAMER, Peter, DARDEN, Lindley, & CRAVER, Carl F. [2000], Thinking about mechanisms, *Philosophy of Science*, 67(1), 1–25.
- MAYNTZ, Renate [2004], Mechanisms in the analysis of social macro-phenomena, *Philosophy of the Social Sciences*, 34, 237–259.
- MCGINNIS, W., LEVINE, M. S., HAFEN, E., KUROIWA, A., & GEHRING, W. J. [1984], A conserved DNA sequence in homoeotic genes of the *Drosophila* Antennapedia and bithorax complexes, *Nature*, 308(5958), 428–433, doi: 10.1038/308428a0.
- MILLIKAN, Ruth Garrett [1984], *Language, Thought, and Other Biological Categories*, Cambridge, MA: Bradford Books.

- MISHKIN, Mortimer, UNGERLEIDER, Leslie G., & MACKO, Kathleen A. [1983], Object vision and spatial vision: Two cortical pathways, *Trends in Neurosciences*, 6, 414–417.
- STANIER, Roger Y., INGRAHAM, John L., WHEELIS, Mark L., & PAINTER, Page R. [1987], *General Microbiology*, London: Macmillan, 5th edn.
- WIMSATT, William C. [1974], Complexity and organization, in: *PSA 1972: Proceedings of the Philosophy of Science Association*, edited by K. F. Schaffner & R. S. Cohen, Dordrecht: Reidel, 67–86.
- [1987], False models as means to truer theories, in: *Neutral Models in Biology*, edited by Matthew H. Nitecki & Antoni Hoffman, New York; Oxford: Oxford University Press, 21–55.