Reflex theory, cautionary tale

Synthese
https://doi.org/10.1007/s11229-021-03351-w

ORIGINAL RESEARCH

Reflex theory, cautionary tale: misleading simplicity in early neuroscience

M. Chirimuuta[1]

Received: 28 September 2020 / Accepted: 3 August 2021

Abstract
This paper takes an integrated history and philosophy of science approach to the topic of "simplicity out of complexity". The reflex theory was a framework within early twentieth century psychology and neuroscience which aimed to decompose complex behaviours and neural responses into simple reflexes. It was controversial in its time, and did not live up to its own theoretical and empirical ambitions. Examination of this episode poses important questions about the limitations of simplifying strategies, and the relationship between simplification and the engineering approach to biology.

Keywords History of neuroscience · History of psychology · Philosophy of neuroscience · Philosophy of psychology · Behaviourism · Instrumentalism · Experimentation

″no concept more problematic and open to question than the concept of simplicity" Kurt Goldstein (1934/1939, p. 2)
"Do this dog's eyes not sparkle with joy? Why not investigate the phenomenon of joy in the dog; here it is much more elementary and therefore accessible." Ivan Pavlov (quoted in Todes, 2014, p. 295)

1 Introduction

Most recent philosophical work on simplifying strategies, such as abstraction, in science examines cases of prima facie successful, or at least useful science, and offers accounts of how it is that the science works as well as it does, given that its theories

This article belongs to the topical collection "Simplicity out of Complexity? Physics and the Aims of Science", edited by Florian J. Boge, Martin King, Paul Grünke and Miguel Ángel Carretero Sahuquillo.

✉ M. Chirimuuta
m.chirimuuta@ed.ac.uk

[1] Department of Philosophy, University of Edinburgh, Edinburgh, Scotland

Published online: 23 August 2021

and models involve such drastic simplifications. In this paper my subject is a now obsolete branch of neuro- and behavioural science, which is not, with hindsight, considered to be very successful. As much as a historian of science might blanche at this 'negative-Whig' manoeuvre—the study of past science not as the glorious precursor of the present, but as a source of cautionary tales instructive today—I hope to show that this episode provides some clues about how to judge when such strategies in neuroscience are *not* helpful, when simplifications are rather too *simplistic*.

My topic is "reflexology," the now outmoded term for the branch of neurophysiology and psychology[1] centred around the simplifying assumption that *complex* patterns of behaviour are concatenations of *simple* reflex responses, exemplified by the sensory-motor reflex arc discovered in the 1830's.[2] In his history of the reflex theory, the psychologist Franklin Fearing writes that,

> For those sciences which are primarily devoted to the study of the integrated responses of living organisms, the concept of reflex action has played, in the 19th and first twenty-five years of the 20th century, a dominating role, comparable, perhaps, to the influence of the Newtonian hypotheses in physics. (1930, p. 4)[3]

Unlike Newtonian mechanics, which is now deemed usable and valuable, even if strictly speaking false, 'reflex mechanics' is no longer employed by behavioural and neuroscientists today. It does not feature in textbooks and has not been part of the training in theoretical neuroscience for decades. This retrospective lack of success makes reflexology worth considering as a case of simplification-gone-wrong. We can ask whether the in-aptness of the simplifying assumptions contributed to the failure of this programme. In addition, while the reflex theory was the mainstream framework in the period we are considering, the first three decades of the twentieth century, it did have its vocal critics. It is valuable to consider the skeptical voices—the reasons given for rejecting the simplifications presumed by proponents of the reflex theory. We might then ask if any of these criticisms appear applicable to the abstractions deployed in neuroscience today. We can wonder if the theories of contemporary neuroscience will, with hindsight, look as over-simplified as the reflex theory seems to us today.

Contemporary discussions of simplification in science centre on the use of abstraction and idealisation in modelling and theorisation. In my account I stick to the more general term, simplification, except where the term "abstraction" is employed by historical actors. This is because the reflex theory is not easily accommodated by contemporary treatments (e.g. Levy, 2018 and references therein). Rather than involving

[1] This combination sciences used to be referred to as "physiological psychology" (Smith, 1973). The concept "reflex action", accordingly, can refer either to a pattern of nerve activation, or to an involuntary kind of behavioural response (Fearing, 1930, p. 6).

[2] See Canguilhem (1994) and Clarke and Jacyna (1987) on the early discoveries in reflex physiology.

[3] Compare the following remark from B.F. Skinner:
"Somewhere, possibly in a series of articles in the *Dial* in the late 20's, Russell pointed out that the concept of the reflex in physiology had the same status as the concept of force in physics." (Skinner, 1931/1961, p. 319, from the preface to the reprint of the early publication).

deliberate omission of detail (abstraction), or introduction of falsehood (idealisation), within a representation such as a mathematical model, the reflex theory employed the simplifying strategy of decomposition of a complex system (the nervous system) or phenomenon (behaviour of an animal) into putative elementary parts—the reflex arc or the reflex response. These reflexes were metaphorical "elements" or "atoms"—parts whose properties were posited to stay the same regardless of varying conditions around them, or observable changes in the animal's pattern of movements. This simplifying strategy offers a kind of reductive explanation, as argued by Bechtel and Richardson (2010, p. xxxvii), and to be further discussed in Sect. 4. One of the proponents of the reflex theory, Jacques Loeb, stated the fundamental intuition quite succinctly:

> it is better for the progress of science to derive the more complex phenomena from simpler components than to do the contrary. (1912, p. 58)
> The progress of natural science depends upon the discovery of rationalistic elements or simple natural laws. (1912, p. 59)

An important question, raised by the work of Loeb and others, is whether the belief in the fundamental simplicity of nature that these statements seem to express, has anything more than a pragmatic status.

In the next section of this paper I give a brief exposition of the reflex theory of the early twentieth century, followed in Sect. 3 by an account of some major criticisms that were raised against it by various scientists and philosophers. These attacks homed in on the shakiness of the atomistic assumption. There were repeated empirical observations of the context sensitivity of the supposedly most elementary reflex; the simple reflex was, arguably, no more than an artefact of experimental origin. However, as I will argue in Sect. 4, the criticisms have less bite against the ontologically non-committal version of the reflex theory, defended by B.F. Skinner, under the influence of anti-realist philosophies of science authored by Mach and Bridgman. A reflex theory, construed instrumentally, was more defensible. Such a construal brings our attention to the engineering aims of the reflex theory and the behaviourist psychology which grew out of it. This programme was reliant on the artificial production of simplicity via constrained experimental conditions, in order to produce specific effects. On this view, the real shortcoming of the reflex theory, and likely reason for its obsolescence, is its failure to meet its own instrumental goals.

2 Atoms of the nervous system, elements of behaviour

In my exposition of the reflex theory I will focus initially on two influential scientists, active in the early decades of the twentieth century—Jacques Loeb and Ivan Pavlov.[4] Loeb's research in this area was primarily in the physiology of

[4] See Pauly (1987) for an account of Loeb's career and methodology and Todes (2014) on Pavlov; also Smith (1992) on Pavlov and his predecessors. For ease of exposition I have smoothed over much of the variety of opinion within reflexology. To get a better sense of the range of views within this programme, see Fearing's (1930) discussion of many now forgotten figures. Note also that in my account Pavlov's

invertebrates, whereas Pavlov is still renowned for his experiments on learning in dogs. Both envisaged that the concept of the reflex could be the basis of an integrated and complete explanation of brain, nervous system, and behaviour. Another shared characteristic of their outlook was that they sought to model biology and psychology on the physical sciences which, in particular, meant applying the analytical methods of mechanics to the animal. That is, they aimed for a decomposition of the nervous system, and movements of animals, into the components (i.e. reflexes), whose occurrence could, when taken together, account for the activity of the whole nervous system or whole animal.[5] The science of Loeb and Pavlov was self-consciously, and literally, mechanistic.

The very first words of the introduction to Loeb's *Comparative Physiology of the Brain* are as follows:

> The understanding of complicated phenomena depends upon an analysis by which they are resolved into their simple elementary components. If we ask what the elementary components are in the physiology of the central nervous system, our attention is directed to a class of processes which are called reflexes. (Loeb, 1900, p. 1)

Loeb goes on to give two examples of reflexes—the eyelid closing on the advance of a foreign body, or the narrowing of the pupil in response to light. He then defines the reflex by describing what is common to these examples:

> In each of these cases, changes in the sensory nerve-endings are produced which bring about a change of condition in the nerves. This change travels to the central nervous system, passes from there to the motor nerves, and terminates in the muscle-fibres, producing there a contraction. This passage from the stimulated part to the central nervous system, and back again to the peripheral muscles, is called a reflex. (Loeb, 1900, pp. 1–2)

And he adds that, "[t]here has been a growing tendency in physiology to make reflexes the basis of the analysis of the functions of the central nervous system" (Loeb, 1900: p. 2).

In the first lecture of Pavlov's *Conditioned Reflexes* we find a comparable definition.[6] In addition, he emphasises the "necessity" of the connection between stimulus and response, a characteristic which ensures that the reflex is a "genuine scientific conception" (Pavlov, 1927/1960, p. 7). Like Loeb, Pavlov describes reflexes as

Footnote 4 (continued)
classification of *unconditioned* and *conditioned* reflexes is treated as equivalent to the distinction between *simple* and *complex* reflexes by Sherrington and others. Both terminologies capture the notion of elementary versus non-fundamental reflexes but this papers over the differences between these sets of concepts, and the roles they play in their respective theories.

[5] See Falkenburg (2019, pp. 76–9) on these methods of decomposition and explanatory reconstitution (AKA "analysis and synthesis"), and their relation to mechanistic science.

[6] "An external or internal stimulus falls on some one or other nervous receptor and gives rise to a nervous impulse; this nervous impulse is transmitted along nerve fibres to the central nervous system, and here, on account of existing nervous connections, it gives rise to a fresh impulse which passes along outgoing nerve fibres to the active organ, where it excites a special activity of the cellular structures." (Pavlov 1927/1960, p. 7).

having an elementary status—he refers to them as "the elemental units in the mechanism of perpetual equilibration"—the mechanism through which an animal adapts its behaviour to its surroundings (Pavlov, 1927/1960, p. 8). A distinctive feature of Pavlov's theory is the division between *unconditioned* and *conditioned* reflexes.[7] The former type being those that exist from the birth of an animal and which persist after the removal of the cerebral cortex.[8] The second kind are the central topic of Pavlov's research. Conditioned reflexes are said to be learned through the creation of an association between an arbitrary stimulus (e.g. the ticking of a metronome at a particular frequency) and a stimulus that innately causes a reflexive response—such as food, the stimulus for the unconditioned reflex of salivation. The cerebral cortex is said to be the neural substrate of conditioned reflexes, since these responses do not persist after surgical removal of this structure. Pavlov posited no additional, "higher" mental structures which exert control over reflexive behaviour—all action is said to be determined by these two kinds of reflexes.

This radical conclusion did find adherents, such as Hull and Baernstein (1929, p. 14) who write,

> It is believed by increasing numbers of students of human and other mammalian behaviour that the conditioned reflex, with its power of substituting one stimulus for another, is the basic mechanism not only of ordinary habits but of the entire mental life.

The proposal of these authors is that the conditioned reflex should be replicable in a man-made device. Their idea was to build a reflex machine, and assess its capacities for learning and intelligence, as a test of this generalised theory of the mind. Still, the thesis that the reflex is the basis of "mental life" in its entirety was never endorsed by the British physiologist Charles Sherrington who himself was a major contributor to the field. Sherrington's own way of characterising the elementary status of the reflex was very influential:

> The reflex-arc is the unit mechanism of the nervous system when that system is regarded in its integrative function. The unit reaction in nervous integration is the reflex, because every reflex is an integrative reaction and no nervous action short of a reflex is a complete act of integration. (Sherrington, 1906, p. 7)[9]

His view was that these fundamental units of activity, the reflexes, were compounded together during the evolution and development of the nervous system to

[7] Todes (2014, p. 1) points out that this is a mistranslation of Russian terms better conveyed as "unconditional" and "conditional". However, I keep to the standard translation as the difference in meaning is not pertinent to my discussion.

[8] Pavlov describes the unconditioned sort as the, "numerous machine-like, inevitable reactions of the organism—reflexes existing from the very birth of the animal, and due therefore to the inherent organization of the nervous system." (1927/1960, p. 8).

[9] Casper (2014) gives a useful overview of Sherrington's career, and the notion of "integration".

generate the basis for more complex behaviour. The motions of running, walking and leaping are the outcome of reflexes, concatenated together.[10]

I have highlighted how proponents of the reflex theory took reflexes to have an "elementary" status. Relatedly, Kurt Goldstein, a major critic of reflexology (see Sect. 3), refers to its methodology as "atomistic". To understand the claim that the reflex approach is "atomistic", we must take "atom" in its original meaning, where "atoms" are the ultimate simples—being indivisible and unchanging, they are the basic constituents of more complex wholes. Just as there can be metaphorical atoms, outside of physics, we have metaphorical elements outside of chemistry—by extending the idea that there are fundamental, simple kinds of substances out of which more complex compounds are made. To treat simple reflexes as the atoms or elements of nervous system and behaviour is to assert that these processes, or patterns of response, are stable (at least after established, in the case of conditioned reflexes) and that they underlie the apparent complexity of behaviour that is manifestly varying. The atomistic methodology is therefore a simplifying strategy—it is an attempt to explain complex appearances in terms of simpler fundamental components.

Fearing writes of reflexologists such as Pavlov that:

> It is characteristic of this point of view that the 'simple' reflex is described as it appears in the lower animals or in the spinal animal,[11] and there is a tacit assumption that these characters are the same for the more complicated types of nervous action, e.g., those involving the cerebrum. (1930, p. 296)

The idea is, firstly, that the scientist can discover the atoms of the nervous system, or the elements of behaviour, by examining experimental subjects such as invertebrates ("lower animals"), or brainless vertebrates, incapable of any complexity of action; it is presumed that those actions reveal the fundamental components of behaviour in all their raw simplicity. For instance, it is assumed that invertebrates, lacking a cerebral cortex, will only demonstrate responses attributable to unconditioned reflexes. Likewise, vertebrates like dogs and cats who have been experimentally prepared by decapitation or removal of the cortex, will only manifest those simple reflexes. Secondly, the assumption is that in animals who can and do demonstrate more complex behaviour (e.g. dogs and cats with their brains unharmed), the simple reflexes are still there, with the same physiological characteristics as in the "prepared" dog or cat, but somewhat masked by the overlay of complex, conditioned reflexes, as well as top down inhibition of simple reflexes from the brain (Fearing, 1930, p. 296). So while Pavlov emphasises the *necessity* of stimulus–response cause and effect connections, which for him makes the reflex into a properly scientific concept, he does not propose that these connections will always be *observable*. Even conditioned reflexes, he reports, are subject to countless influences, "disturbing factors", which interfere with their manifestation (Pavlov, 1927/1960, p. 20). For this reason, he conducted his experiments on the conditioned reflex in dogs in a special isolation chamber in which the animal had no contact with its fellows, or even

[10] See Graham Brown (1914, pp. 19–20) for discussion.

[11] I.e. one whose entire brain has been removed (Sherrington, 1909).

with the experimenter.[12] Pavlov's reliance on such scenarios is significant. We will now turn to criticisms of the reflex theory, and see that much of the scepticism is focussed on the assumption of the existence–*outside of controlled laboratory conditions*–of any of these simple and elementary reflexive responses.

3 The criticisms of reflexology

The supremacy of the reflex theory did not go uncontested. In this section I summarise a number of the objections that were levelled at the ontological and methodological assumptions of reflexology, including the interventions of two philosophers, John Dewey and Maurice Merleau-Ponty. My summary is not exhaustive. For example, I do not include the case against the reflex theory presented by physiologist Thomas Graham Brown precisely because he does not take issue with the simplifying assumption of decomposition into fundamental units, but instead proposes an alternative kind of element, the "half centre" (e.g. Graham Brown, 1914).

Six strands of criticism can be listed as follows. I will discuss each in turn.

1. Empirical findings of the lack of stability of simple reflexes and conditioned reflexes.
2. The ad-hocness of the postulates added to the reflex theory in order to achieve consistency with those empirical findings.
3. Questioning of the reductionist methodology which seeks explanation of behavioural wholes in terms of simple parts.
4. Dubiousness of the extrapolation from the neurophysiology of the periphery and spine, assumed by the reflex theory, to anatomically unknown structures within the brain.
5. The lack of ecological validity of physiological experiments performed on surgically altered animals, and of behavioural experiments performed in highly artificial laboratory conditions.
6. Dubiousness of the notion of the simple reflex, even when construed as an abstraction.

[12] It is worth reading Pavlov's own justification for the selection of such unnatural conditions for his experiments on conditioning:

"It was evident that the experimental conditions had to be simplified, and that this simplification must consist in eliminating as far as possible any stimuli outside our control which might fall upon the animal, admitting only such stimuli as could be entirely controlled by the experimenter. …. The environment of the animal, even when shut up by itself in a room, is perpetually changing. Footfalls of a passer-by, chance conversations in neighbouring rooms, slamming of a door or vibration from a passing van, street-cries, even shadows cast through the windows into the room, any of these casual uncontrolled stimuli falling upon the receptors of the dog set up a disturbance to the cerebral hemispheres and vitiate the experiments. To get over all these disturbing factors a special laboratory was built at the Institute of Experimental Medicine in Petrograd…" (Pavlov 1927/1960, p. 20).

The purpose of this section is to give a presentation of these criticisms, without evaluation or endorsement. Assessment of the cogency of the criticism is deferred to Sect. 4.

One case against the reflex theory was put forward by the German neurologist, Kurt Goldstein. His first point of attack is that experimental reports of the putative simple reflexes do not show the stability, the constancy of response, postulated by the theory (Goldstein, 1934/1939, 69ff.). Summarising the results of various researchers, including some proponents of the reflex theory, it appears that the simple responses, like the patellar reflex, are alterable with bodily posture and attentional state. As shown by Sherrington (1906), the "receptive field" of the scratch reflex in dogs—the area on the skin in which a stimulus can elicit the scratching movement of the leg—varies from day to day, in terms both of its location and the kind of response elicited. Regarding the conditioned reflex of Pavlov, Merleau-Ponty (1942/1967, p. 58), following Buytendijk and Plessner (1936), argues that it is too unstable to do the theoretical work required of it. A striking case of instability comes from Pavlov's reports of the behaviour of two dogs who had been subjected to repeated conditioning experiments. They appear to fall into a "hypnotic" state and fail to give the expected reactions either to the conditioned or unconditioned stimulus.

This brings us to the next allegation, that the reflex theory is full of ad-hoc modifications—unprincipled use of terms such as "excitation, inhibition and disinhibition"—brought in to mask the disagreement between theory and observation (Merleau-Ponty, 1942/1967, 58ff.; pp. 19–20; cf. Buytendijk & Plessner, 1936). In particular, when the usual stimulus fails to elicit the expected reflex, it is posited that a process of inhibition has been activated, preventing the response; but independent evidence for the inhibitory mechanism is not established. Goldstein contends that with the proliferation of hypotheses accounting for "modifications" of normal, simple reflexes[13] the theory loses its justification for drawing a distinction between the normal reflex and variants of it. This lack of justification goes unnoticed because researchers automatically classify the responses produced in certain kinds of artificial experiments as the normal ones (Goldstein, 1934/1939, pp. 80–1).

Given this, Goldstein is sceptical of the classification made of the "normal" reflex, versus its variants, of the simple reflex versus the complex patterns of behaviour that they are said to comprise. Put together with the above-noted observation of lack of stability of the supposed elementary responses, Goldstein calls into question the reductionist methodology that attempts to explain a complex behavioural whole in terms of simpler parts. The following passage is worth quoting at length:

> The customary method attempts to reduce variable to constant reactions, seeing, in the latter, the basic ones, and regarding the former as modifications. This tendency is understandable as a very natural desire to deal with

[13] Goldstein lists these hypothetical factors: "inhibition, facilitation, neural switching or shunting of different kinds, influence through peripheral factors, such as the state of tension of the muscles, position, enforcement or diminution through other reflexes, 'central' factors, and amongst these, particularly, psychic factors" (1934/1938, p. 80).

> constant factors. The supposedly greater simplicity of constant reactions lends itself as a starting point for a theory, in that the variable responses can then be understood as complexes derived from the more simple and constant ones. However, there is no question but that the so-called variable processes are, in reality, no less constant, if one takes into consideration all their causal conditions. Concerning the question of simplicity and complexity, and whether the complex can be deduced from the simple, we shall see, in our later discussion, that the converse view is probably nearer the truth. (Goldstein, 1934/1939, p. 80)

What this suggests is that Loeb's assertion that it is, "better for the progress of science to derive the more complex phenomena from simpler components than to do the contrary" has met an obstacle: in the science of the nervous system and behaviour, Goldstein argues, no foundation can be found in a substrate of components or processes, both simple and stable.

Along with the attempt to derive the complex from the simple, the reflex theory proposes to infer facts about brain processes from observations of the more accessible processes in the spine and peripheral nervous system. This extrapolation is challenged by Theodore Hough in an address to the American Association for the Advancement of Science:

> we miss entirely the satisfaction of seeing the cerebral functions clearly pictured in terms of neurone structure. We trace the 'way in' and the 'way out'; we see that the connection between the afferent and efferent nerve fibers is in the cortex; but what takes place in the cortex? Is it objectively nothing more than our typical reflex raised to the *n*th power of complexity? Perhaps it is; but does any one feel reasonably sure of it? For one, I confess I do not. (Hough, 1915, p. 408, quoted in Fearing, 1930, p. 287)

It is significant here that the modelling of cortical neurophysiology, as merely a more complicated kind of reflex, is reported as if it has been taken as a matter of faith. At this time, the neuroanatomy and physiology of the brain was uncharted, in comparison with that of the nervous system below the neck. Given their ignorance, reflexologists made the parsimonious assumption that there was nothing radically different going on in the brain. But research later in the twentieth century showed that the parsimony was in this case misleading.

A core challenge to reflexology turns on its deficiency in what we would now call *ecological validity*—the lack of applicability of experimentally generated phenomena to the explanation of the intact nervous system or unconstrained animal behaviour.[14] Critics of reflexology go so far as classifying the central phenomena of the research programme to be experimental artefacts. Having rejected the notion that reflexes are the elementary components of nervous system and behaviour, Goldstein

[14] Note that a version of ecological validity was a core "methodological postulate" for Goldstein: "no phenomenon should be considered without reference to the organism concerned, and to the situation in which it appears." (1934/1939, p. 25).

concludes that they are no more than an "expression of experimentally produced injury" (1934/1939, p. 157), especially because of the manifest difference between reflex reactions of the legs and the normal flow of movements of an animal walking over its accustomed ground (1934/1939, pp. 169–170). The failure to notice the lack of similarity between ordinary movements and reflexive ones stems, Goldstein argues, from the fact that many physiologists never dealt with intact animals in their research (1934/1939, p. 90).[15]

Although Pavlov performed his conditioning experiments on animals whose nervous systems were unharmed, he still faced the criticism that the artificiality of his experimental set up—such as the confinement and isolation of the dogs—placed limitations on what could be inferred from his results about learning and behaviour in general. Goldstein (1934/1939, p. 175) argues that the precise, repeated paring of unconditioned and conditioned stimuli does not occur in the lives of animals away from human control. Thus, they do not help to explain animals' learning in the wild, but do shed light on the processes in play during human training of animals (p. 178). A comparable point about difficulties arising with use of artificial stimuli had been made by Herbert Spencer Jennings, a zoologist and former student of Dewey, in response to Loeb's attempt to make *galvanotropisms* (reflex responses to electric currents) fundamental to the explanation of movement (see Loeb, 1900, chap. XI). Not only did the movements elicited in those experiments appear highly unnatural, but the electrical stimulus was one that simply did not occur in the environment of the organism—how could it then form an explanatory basis for the account of ordinary locomotion? (Jennings, 1906, chap. XIV; and see Pauly, 1987, chap. 6). Dwelling on Pavlov's report of the two dogs for whom repeated conditioning experiments led to their entering into a hypnotic stupor, Buytendijk and Plessner (1936) conclude that his research on conditioning is far more informative about the genesis of neurosis than any non-pathological phenomena.

We now move to the final point, on the lack of utility of the reflex concept, even as an abstraction. Dewey (1896) was one early critic of reflex psychology. In his 1896 article, He argues that the assumption foundational to the theory, of a clear distinction between stimulus and response, sensory and motor operations, is an artificial, misleading abstraction, masking the concrete fact of the interdependence of sensation and movement. In *The Integrative Action of the Nervous System,* Sherrington concedes that the notion of the simple reflex is an abstraction, but makes the claim that it is at least a "convenient fiction":

> [A] simple reflex is probably a purely abstract conception, because all parts of the nervous system are connected together and no part of it is probably ever

[15] Canguilhem (1965/2008, p. 113) gives an interesting commentary on Goldstein's view on experimental conditions:

"The situation of a living being commanded from the outside by the milieu is what Goldstein considers the archetype of a catastrophic situation. And that is the situation of the living in a laboratory. The relations between the living and the milieu as they are studied experimentally, objectively, are, among all possible relations, those that make the least sense biologically; they are pathological relations."

Elsewhere in the essay (pp. 107–111), the changing fortunes of reflexology are discussed.

> capable of reaction without affecting and being affected by various other parts, and it is a system certainly never absolutely at rest. But the simple reflex is a convenient, if not a probable, fiction. Reflexes are of various degrees of complexity, and it is helpful in analyzing complex reflexes to separate from them reflex components which we may consider apart and therefore treat as though they were simple reflexes. (Sherrington, 1906, p. 8; Cf. 114)

This is one way to deal with the objection that stable and constant reflexes are never actually observed (Point 1 above). Given their ubiquity in other branches of science, one may rightly ask what is wrong with "abstractions" or "fictions", so long as they are recognised as such.

Goldstein rejects even the "fictional" concept of the simple reflex because (for reasons discussed above) he does not think that it delivers the requisite understanding of the intact organism.[16] Merleau-Ponty delivers an involved response to Sherrington's deployment of the notion of abstraction. He counters Sherrington's claim that the reflex is an abstraction by asserting that it is actually a concrete phenomenon, albeit one which is contrived experimentally, and lacking more widespread significance:

> But neither is the reflex an abstraction, and in this respect Sherrington is mistaken: the reflex exists; it represents a very special case of behavior, observable under certain determined conditions. But it is not the principal object of physiology; *it is not by means of it that the remainder can be understood.* (Merleau-Ponty, 1942/1967, p. 46)

Merleau-Ponty criticises Sherrington for his deployment of the abstract idea of the reflex in order to preserve an ontology of animal machines in the face of countervailing evidence (Moinat, 2012, pp. 95–97). Yet these appeals to the reflex are not adequate to account for Sherrington's key discoveries of "integration"—the coordination of movement required for adaptive behaviour. As Merleau Ponty puts it,

> It is paradoxical to conserve the notion of the reflex arc theoretically without being able to apply it anywhere in fact. As in all the particular questions which we have mentioned, in his general conception of nerve functioning Sherrington seeks to save the principles of classical physiology. His categories are not made for the phenomena which he himself has brought to light. (1942/1967: p. 33)

And here we are brought back, in a roundabout way, to the beginning of our list of criticisms: the failure of the reflex theory to properly meet empirical facts.

There is not space in this paper to tell the story of the decline of the reflex. Without further historical research, I cannot comment on the impact of the criticisms listed above. Eventually the reflex theory was eclipsed when computation came to provide an alternative simplifying framework for neuroscience and cognitive science

[16] See Chirimuuta (2020) for a more detailed discussion of Goldstein's views on abstraction.

in the mid twentieth century.[17] In terms of its core tenet—that nervous system and behaviour can be explained via decomposition into elementary reflexes—it is a theory without retrospective success.[18] Arguably, though, there is continuity in a mechanistic ethos, stretching from reflexology to cybernetics, and from there to cognitive science and computational neuroscience, programmes which also seek the duplication of cognition in machines, as we saw with the proposal Hull and Baernstein (1929).[19] And certainly there is continuity between the methodologies of the reflexologists and neuro- and behavioural science today, such as the division between stimulus and response criticised by Dewey, the practice of reductionism, and the use of artificial experimental conditions. Thus we can pose the question of whether this methodological continuity leaves contemporary science open to the kinds of criticisms outlined in this section. But before that, we must do some evaluation of the merits of these attacks. We will find that some of them miss their mark because of a more fundamental disagreement, between critics and proponents of the reflex theory, concerning the aims of scientific research.

4 Simple, simplistic, simplified

From a vantage point built up of one hundred years of further research on the central nervous system, it might seem incredible that scientists of stature ever believed that the reflex arc would be the one key to demystify brain and behaviour. To our retrospective view, the reflex theory now appears obviously too simple to account for the phenomena it was supposed to—it looks *simplistic*. We might ask what it was that made the reflex theory so appealing. There are some suggestions that its value was precisely in its being so simple—an attractive over-simplification. Hough, in his critical piece on the theory, writes of how its "diagrammatic clearness" has shaped researchers' "mental approach" to their problems, and how it naturally aligns with textbook expositions that begin with peripheral neuroanatomy and end with the physiology of the brain (1915: p. 408). Fearing observes that, "The reflex arc is easily diagrammed in the textbook" but warns, "such a diagram readily forms the basis for a discussion of simple stimulus–response relationships, which is misleading even in connection with the simpler animal responses, and positively inapplicable to the more complex organic responses" (1930, p. 288). Karl Lashley (a onetime student of the behaviourist Watson) relates that the passing down of the textbook picture, across generations, has given it an entrenched, unquestioned, status:

> In the course of time there has been built up a *simple,* traditional, textbook account of the mechanism of reaction, prepared for students' consump-

[17] Elsewhere I discuss the rise of computationalism, arguing that its appeal rested not least in the simplifications that it offered to neurophysiologists (Chirimuuta, 2021).

[18] See e.g. Todes (2014, pp. 300–2) on the failed ambitions of Pavlov's project.

[19] It is interesting that Barack and Krakauer (2021) use the term "Sherringtonian" to refer to classic, single neuron based theories of computation in the brain, alluding to a connection with the earlier reflex theory.

> tion. Repeated copying from one text to another has crystallized it, and early instruction has given us faith in it. The original sources have been almost lost to view and with them the appreciation of the difficulties, the uncertainties, the many unsubstantiated assumptions which underlie every assertion of the classical account. (Lashley, 1931, p. 16, emphasis original)

But what we learn from the cautionary tale of the reflex theory is that the scientists' instinctive tendency to head in the direction of simplicity—their *simpletropism*, we might name it—can sometimes send them in the wrong direction. But how, for any current piece of un-settled science can one tell if it is over-simple, or rather if its simplicity is due its partaking of the triad of beauty, truth, and parsimony?

4.1 Evaluation and arbitration

One way to answer this question would be to see if any of the criticisms listed in Sect. 3 can rightly be levelled at research today. But that is to assume that those criticisms were appropriate and correctly diagnosed the flaws of over-simplification within the reflex theory. So we must first turn to evaluation of the criticisms. One way to summarise Goldstein's complaint against reflexology is that it is a reductionism gone rogue, due to the lure of parsimonious explanation. The reflexologist employs a reductionist methodology—opting to study parts (simple reflexes) in isolation, with the aim of seeing how their operation together will yield an explanation of the whole nervous system.[20] Furthermore, the reflexologist tends to assume a reductionist ontology, supposing that reflexes are elementary components, which when aggregated together comprise the whole nervous system; he takes it that the organism is "a bundle of isolable mechanisms which are constant in structure, and which respond, in a constant way, to events in the environment (stimuli)" (Goldstein, 1934/1939, p. 67). Which is to say, the possibility of context dependency for these responses—of parts behaving differently when situated in their wholes—is not considered by the reflexologist. Much of the content of Goldstein's magnum opus, *The Organism,* is a statement of the importance of context dependency in biology.

According to Todes (2014), the picture of the organism as a mere aggregate of physico-chemical mechanisms, the reflex being the relevant one for the nervous system, was indeed foundational for Pavlov. He was a reductionist in both the methodological and ontological senses. The failure, described in Sect. 3, of detailed experimental work to provide support for the existence of stable, elementary reflexes therefore stands as a challenge to Pavlov's picture. However, not all practioners of reflexology went along with the ontological interpretation. The psychologist B.F. Skinner is an important contrast case. Under the influence of Mach and Bridgman,[21] he asserted an operationalist philosophy of science in which it was unnecessary and

[20] This formulation of reductionism is very much in line with the one presented by Bechtel and Richardson (2010): reduction as decomposition of the living system into component mechanisms.

[21] See Moore (2005) on Skinner's philosophical influences. The connections between Mach, Loeb, and Skinner's teacher, Crozier, are especially interesting.

misguided to entertain the question of whether a simple reflex *really* exists, and whether any experiment has been adequate to reveal it.

> Is a reflex a unitary mechanism? Is behavior a sum of such mechanisms? Then, if by reflex we mean a hypothetical entity which exists apart from our observations but which our observations are assumed to approach, the questions are academic and need not detain us; if, on the other hand, we define a reflex as a given observed correlation or as a statistical treatment of observed correlations, the questions are meaningless, for they ignore the process of analysis implied in the definition. A reflex, that is to say, has no scientific meaning apart from its definition in terms of such experimental operations as we have examined, and, so defined, it cannot be the subject of questions of this sort. (Skinner, 1931/1961, p. 341)

Thus it becomes clear that the first line of criticism, that simple, stable reflexes do not exist, and a fortoriori, nervous system and behaviour are not compounded from them, can only be levelled at the ontologically committed version of the reflex theory, but not the operationalist version put forward by Skinner.

However, this anti-realist move does not by itself deflect concerns about methodological reductionism within reflexology, and the related criticism of the lack of ecological validity. These are concerns about the appropriateness of studying physiological responses, and behaviours in experimentally prepared conditions of isolation, when the ultimate goal of research is to understand the physiology and behaviour of the whole organism. In a rather harsh review of Fearing's (1930) book, Skinner and his teacher Crozier do in fact respond to this complaint. Their basic point is that any charges levelled against the concept of the reflex, and experimental practice derived from it, would generalise to successful scientific practices in physics, chemistry, and elsewhere in biology. For, they argue, those sciences employ concepts or posits that do not have empirical standing other than through technical operationalisations, and whose generalisation to phenomena outside of the laboratory is moot.

> The reflex is exactly comparable to any other well-established scientific concept, from the electron to the gene. It is not observed in isolation but the notion at it is arrived at by a certain set of operations. It is the conceptual expression of a correlation between certain observed events (called, in this case, stimulus and response), and has no validity beyond this correlation. (Skinner & Crozier, 1931, p. 126)

Thus, we see that their rebuttal of the criticisms is bound up with a generalised operationalist stance towards all the other branches of science.

With this stance there comes a shift away from the justification of reductionism in terms of *simplicity in nature* and towards an emphasis on *simplification*—the production of simple phenomena and cause and effect relationships that are not claimed to be present in nature, independently of laboratory procedures, but nonetheless serve some practical purposes. This position is consistent with the ideas of Mach and Bridgman. According to Mach, the "task of science is to provide the fully developed human individual with as perfect a means of orienting himself as possible"

(1886/1914, p. 37; quoted in Smith, 1995, p. 42). That is, the aim of science is not to supply disinterested knowledge of nature, but to furnish the agent with tools for effective practice. As such, Mach's doctrine of science as "economy of thought"—where science is to provide "the concisest and simplest possible knowledge of a given province of natural phenomena" (1883/1919, pp. 6–7)—cannot be taken as the claim that science must reveal order and simplicity in nature, but that simplicity is strived for because of its utility.[22] Similarly, Bridgman (1927, pp. 51–2; cf. p. 204ff.) remarks that any scientist's conviction in the fundamental simplicity of nature—owing, for example, to belief in there being only a small number of elements—has no more than a pragmatic status. Bridgman observes also that methodological reductionism has practical appeal because of its ease of application, but then it, "will appear to be of disproportionate importance" (1927: pp. 221–2).

To tie together these threads, we see that the charges against reflexology regarding methodological reductionism, lack of ecological validity, and even the non-existence of the simple reflex, are not devastating against an operationalist and instrumentalist construal of the reflex theory. If simplifications are justified for the part played in a quasi-engineering project, the production of specific responses and behaviours, then these criticisms are not pertinent. The reflexologist-as-technologist no longer sees himself as a disinterested researcher seeking out the truths of nature, but more as an investigator whose goal is to take command of natural processes.[23]

Not coincidentally, the behaviourists in America were enthusiastic about the second of these two job descriptions. At the head of John Watson's behaviorist manifesto it is declared that psychology's "theoretical goal is the prediction and control of behavior" (1913, p. 158).[24] As Pauly (1987, p. 174) reports, Loeb's model of biology as a discipline aiming at control of nature was a direct influence on his student, Watson:

> Watson's central innovation was to place the control of behavior at the foundation of psychology as a science. By arguing that control was knowledge, he broke down the barriers between the aims of pure psychology and those of behavioural technology. In this sense behaviourism was a model Loebian science, organized around the desire 'to get life phenomena under our control.' In both its positivistic methodology and its radical social claims it was

[22] See Smith (1995, p. 45) where it is argued that Skinner follows Mach in this science-as-economy view.

[23] One complication is that Skinner himself does sometimes talk like a realist regarding simplicity ("order") in nature: "I never face a Problem which was more than the eternal problem of finding order….. Of course, I was working on a basic Assumption—that there was order in behavior if I could only discover it." (Skinner, 1972, p. 112; quoted in Moore, 2005, p. 100).

Note that this was written long after Skinner's reflexology research of the 1930's. However, it is consistent with one account of Skinner as a follower of Francis Bacon in his philosophy of science, and not therefore a thorough-going anti-realist (Smith, 1996, p. 65). The Baconian scientist, guided by the maxim that *Nature, in order to be commanded, must be obeyed*, does have a stake in attempts to reveal the underlying properties and causal structures within natural systems.

[24] And indeed, Titchener's (1914, p. 14) rebuttal to Watson's attack on the structuralist (introspectionist) psychology was to say that behaviourism is technology, whereas structuralist psychology is an actual science.

> the direct descendant of the ideas developed by Loeb in the early 1890s. For Watson himself, the engineering standpoint represented independence and excitement—from the level of laboratory innovation to that of power for social change. He saw himself in opposition to the received wisdom of his field; like Loeb he would cut through complexity with continuous experimental activity.

Watson's uptake of the idea of the simple reflex as a means to analyse behaviour was directly influenced by Loeb (Pauly, 1987, p. 175). The key idea is that complexity is reduced ("cut through") and simplicity generated—not discovered—through experimental activity. Another important point is that the "natural" state of things—how organisms are independently of experimental manipulations—is not privileged in the Loebian view.[25] Against this stance, the criticisms of reflexology centred on the artificiality of the experimental preparations, and their lack of ecological validity, do not have force. They merely avert to a bigger dispute between the proponents and critics of reflexology—a disagreement about what biological science fundamentally is, and how it should be conducted—whether its task is the control of life or the understanding of organism and its behaviour, in natural circumstances.[26]

We now see that some of the tenets of reflexology which appeared, to its critics, as simplistic and misguided can be more charitably regarded as simplifications subject to justification, not by their closeness to nature but by their utility within certain practical projects. This is not the attitude taken by all reflexologists towards their posits—Pavlov has already been mentioned as an exception here—but it is attributable to Loeb and the behaviourists influenced by him. By taking up this charitable interpretative stance, we are now in a position to judge reflexology by a standard held by some of its own practitioners—the standard of instrumental success. Yet even by this benchmark the reflex theory cannot be judged a success. Skinner is quite notorious for his claims made for the potential of behaviourism to bring about a utopian world through social engineering (Smith, 1996). But these goals were, alas, not met. Indeed, other than the flourishing of the industries of marketing and advertising (Buckley, 1989, chap 8), reflexology does not have a hallmark success comparable with those of other areas of research associated with Loeb, such as—Pincus's invention of the contraceptive pill (Pauly, 1987, p. 194).

An illustrative case in point comes from the work of Keller and Marian Breland, two former assistants on Skinner's "pigeon project". As Skinner (1947/1961, p. 227) wrote on the true agenda of experimental psychology, "the basic engineering problem is to acquire control. …. It is not a matter of bringing the world into the

[25] Pauly (1987, p. 199) summarises, "the original organization and normal processes of organisms no longer seemed scientifically privileged; nature was merely one state among an indefinite number of possibilities, and a state that could be scientifically boring.".

[26] As mentioned in Sect. 3, Goldstein was willing to grant that research on the conditioned reflex afforded insights into human interventions on animal behaviour (processes of training and drill). On his own conception of biology, however, its goal is the understanding of the "natures of organisms" (1934/1939). Still, his practice as a neurologist did include the aim of therapeutic intervention on brain damaged patients. This was different from the Loebian instrumentalism because he did not think therapeutic success could be achieved through reductionist methods (Harrington, 1996), and he took it to be dependent on the self actualisation of the organism rather than on the imposition of external conditions.

laboratory, but of extending the practices of an experimental science to the world at large." The Brelands took up this challenge, setting out to mass-produce novelty displays of conditioned behaviours, in a variety of animal species, for commercial purposes. Yet, the theoretical predictions derived from laboratory experiments were overwhelmed by "animal misbehavior"—the failure of animals to learn simple, reinforced actions because of the interruption of instincts (Breland & Breland, 1961). Ramsden (2021) summarises, in a detailed study of this episode:

> As the Brelands took operant conditioning beyond the confines of the laboratory, Skinner's tidy system began to fracture, and the 'nature' of the organism began to override the machine-like predictability of conditioned behavior.

To Skinner's consternation, the Brelands reached the conclusion that ethology—the study of the behaviours of animals within their environmental niches—*is* indispensable in the study of animal psychology. What this attempted application of reflexology shows is that in the end, ecological *in*validity bit back: when reflexology was extended beyond its narrow experimental conditions, the over-simplified poverty of this conceptual framework undermined its ambitions. A well crafted conditioning schedule could never bring about happiness if, as it happens, there is too little in common between the joys of life in the wild and the sparkle observed in the eye of Pavlov's dog.

4.2 Lessons?

This study of the reflex theory was sold to the reader as a "cautionary tale", whose contents would have some lesson about how to tell of a current theory that it has taken the bad turn from elegant parsimony to over-simplification. It was hoped that indications were to be learned from the critics of reflexology, but as we saw in Sect. 4.1 at the root of their charges was failure to appreciate that reflexology was in many instances an *engineering approach* to biology, one that aimed at control of behaviour rather than an understanding of organisms and their actions for its own sake. With that goal in view, the *simplistic* was more charitably regarded as *simplified.* What we then found is that the research programme would stand and fall with its technological achievements– which, as it happened, fell below expectations.

This granted, we may still draw some instruction from the past critics of reflexology. In terms of warning signals that a theoretical framework is employing misleading simplification, one indication to be gleaned from our study is that simplification has gone wrong when it is achieved with ad hoc manoeuvres and verbal tricks. This was the second line of criticism mentioned in Sect. 3. Fearing observed this in relation to Pavlov's (1927/1960, pp. 11–12) terminology:

> The organism is regarded as an aggregate of reflexes which may be released by inter-changeable stimuli.
> By the application of this principle*, complex behavior is made verbally simple*, *at least.* Restless movements of the dog when confined are regarded as a

> manifestation of the 'Freedom reflex;' inquisitive behavior is denominated the 'What-is-it?' reflex, etc." (Fearing, 1930, p. 284 and footnote; emphasis added)

This is valid and instructive, in so far as it goes, but one may be hard pressed to find any equivalents in current neuroscientific practice. A surer point of connection is in the assessment of simplifying strategies favoured by reflexologists and still employed in contemporary brain and behavioural sciences. In particular, the division of stimulus and response, the use of highly artificial experimental conditions, and the raft of methodologies that go under the heading of reductionist—such as the fine grained analysis of parts of the system in isolation from the context within cell, organ or organism, and the attribution to parts of the system of a restricted number of elementary functions, whose combined operation is supposed to explain the complex whole.

John Dupré (2012) is one contemporary philosopher who argues, like Goldstein, that context-dependency is under-appreciated in biology, and that reductionism is only plausible because of the failure to recognise that the behaviour of biological parts is highly sensitive to the state of the whole to which they belong. In contrast, Burnston (2020) holds that contemporary reductionistic and mechanistic approaches in neuroscience do not assume atomism—the total context-independence of the behaviour of parts—and that scientists' acknowledgement of some context-dependence of neural functions puts the decompositional approach on a surer footing. The anti-reductionist could still argue that Burntston's suggested attribution of a small number of context-dependent functions to such neural structures massively underestimates the instability of structure–function relationships, and that these restricted attributions only seem plausible because of the selection of experimental conditions.

However, we saw in Sect. 4.1 that concerns about the artificiality of the relationships discovered under laboratory conditions, and of the lack of evidence for a parsimonious cognitive and neural ontology, can be alleviated if one takes an instrumentalist stance to such practices. The simplifying methodology justifies itself if its assumptions and techniques make feasible the production of experimental phenomena, and explanations of them, that would otherwise be unmanageable because of the unstable complexity of brain and behaviour left unconstrained. A simplifying framework can lead to the discovery of powerful interventions within controlled conditions. But here we find that there is an important lesson to take from the failure of reflexology, even in its instrumentalist guise. Skinner's methods failed to achieve the hoped for instrumental success beyond the walls of the laboratory. What this suggests is that the weakness of simplifying methods will show itself when the attempt is made to control and explain natural processes in the world at large. As such, it would be particularly relevant to examine cases of translational neuroscience to find out if this pattern of disappointed expectations for the wider effectiveness of lab-developed interventions is to be found. Anecdotally, this does seem to be the case (Nutt & Need, 2014). To conclude, there are two lessons of this study. The first is that wariness is needed regarding the ontological posits of neuroscience employing radical simplifications. A "cognitive ontology" generated using constrained laboratory procedures could be as much a mirror of the scientist's abstraction-scheme as a map of the mind itself. The second is to grant that a simplified picture of the

mind and brain can be justified instrumentally, but that instrumental success may turn out to be quite dependent on the holding of artificial background conditions, and not translate into effectiveness beyond the lab. Experimental neurophysiology is currently undergoing a minor revolution towards the use of "ethological" behavioural paradigms (Musall et al., 2019, Nastase et al., 2020). This is in part because of the failure of models of neuronal responses to highly controlled stimuli to generalise with accuracy. It may well turn out that in order to achieve widespread instrumental control, neuroscience will have to give up on the long standing goal of seeking simplicity in nature.

Acknowledgements I thank Michael Dietrich and James Lennox for sharing their thoughts on this paper, and two anonymous reviewers for their many constructive comments.

References

Barack, D. L., & Krakauer, J. W. (2021). Two views on the cognitive brain. *Nature Reviews Neuroscience, 22*(6), 359–371.

Bechtel, W., & Richardson, R. C. (2010). *Discovering Complexity.* (2nd ed.). MIT Press.

Breland, K., & Breland, M. (1961). The Misbehavior of organisms. *American Psychologist, 16*, 681–684.

Bridgman, P. W. (1927). *The logic of modern physics*. New York: MacMillan.

Brown, T. G. (1914). On the nature of the fundamental activity of the nervous centres; together with an analysis of the conditioning of rhythmic activity in progression, and a theory of the evolution of function in the nervous system. *Journal of Physiology, 48*, 18–46.

Buckley, K. W. (1989). *Mechanical man: John B Watson and the beginnings of behaviorism*. New York: The Guilford Press.

Burnston, D. (2020). Getting over atomism: functional decomposition in complex neural systems. *British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axz039

Buytendijk, F. J. J., & Plessner, H. (1936). Die physiologische erklärung des verhaltens: Eine kritik an der theorie pawlows. *Acta Biotheoretica, 1*, 151–172.

Canguilhem, G. (1965/2008). The living and its milieu. In: P. Marrati & T. Meyers (Eds.), *Knowledge of life* (pp. 98–120). New York: Fordham University Press.

Canguilhem, G. (1994). The concept of reflex. In F. Delaporte (Ed.), *A Vital Rationalist: Selected Writings from Georges Canguilhem.* New York: Zone Books.

Casper, S. T. (2014). History and neuroscience: An integrative legacy. *Isis, 105*, 123–132.

Chirimuuta, M. (2021). Your brain is like a computer: Function, analogy, simplification. In F. Calzavarini & M. Viola (Eds.), *Neural Mechanisms: New Challenges in the Philosophy of Neuroscience.* Berlin: Springer.

Chirimuuta, M. (2020). The reflex machine and the cybernetic brain: The critique of abstraction and its application to computationalism. *Perspectives on Science, 28*(3), 421–457.

Clarke, E., & Jacyna, L. S. (1987). *Nineteenth-Century Origins of Neuroscientific Concepts*. Berkeley, CA: University of California Press.

Dewey, J. (1896). The reflex arc concept in psychology. *The Psychological Review, 3*(4), 357–370.

Dupré, J. (2012). It is not possible to reduce biological explanations to explanations in chemistry and/or physics. *Processes of Life: Essays in the Philosophy of Biology.* Oxford: Oxford University Press.
Falkenburg, B. (2019). Mechanistic explanations generalized: How far can we go. In B. Falkenburg & G. Schiemann (Eds.), *Mechanistic Explanations in Physics and Beyond* (pp. 65–90). Berlin: Springer.
Fearing, F. (1930). *Reflex Action: A Study in the History of Physiological Psychology*. New York: Hafner Publishing Company.
Goldstein, K. (1934/1939). *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*. New York: American Book Company.
Harrington, A. (1996). *Reenchanted science: Holism in German Culture from Wilhelm II to Hitler*. Princeton, NJ: Princeton University Press.
Hough, T. (1915). The classification of nervous reactions. *Science, 41*(1055), 407–418.
Hull, C., & Baernstein, H. D. (1929). A Mechanical parallel to the conditioned reflex. *Science, 70*(1801), 14–15.
Jennings, H. S. (1906). *Behavior of the Lower Organisms*. New York: Columbia University Press.
Lashley, K. S. (1931). Cerebral control versus reflexology: A reply to professor hunter. *Journal of General Psychology., 5*, 3–19.
Levy, A. (2018). Idealization and abstraction: Refining the distinction. *Synthese*. https://doi.org/10.1007/s11229-018-1721-z
Loeb, J. (1900). *Comparative Physiology of the Brain and Comparative Psychology*. New York: J. P. Putnam's sons.
Loeb, J. (1912). *The Mechanistic Conception of Life*. Chicago, IL: University of Chicago Press.
Mach, E. (1883/1919). *The Science of Mechanics* (T. J. McCormack, Trans.). Chicago, IL: Open Court.
Mach, E. (1886/1914). *The Analysis of Sensations, and the Relation of the Physical to the Psychical* (C. M. Williams, Trans.). Chicago, IL: Open Court.
Merleau-Ponty, M. (1942/1967). *The Structure of Behaviour* (A. L. Fisher, Trans.). Boston: Beacon Press.
Moinat, F. (2012). *Le vivant et sa naturalisation: Le problème du naturalisme en biologie chez Husserl et le jeune Merleau-Ponty*. Berlin: Springer.
Moore, J. (2005). Some historical and conceptual background to the development of B.F. Skinner's 'Radical Behaviorism'—Part 2. *Journal of Mind and Behavior, 26*(1–2), 95–124.
Musall, S., Urai, A. E., Sussillo, D., & Churchland, A. K. (2019). Harnessing behavioral diversity to understand neural computations for cognition. *Current Opinion in Neurobiology, 58*, 229–238.
Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage, 222*, 117254.
Nutt, D., & Need, A. (2014). Where now for schizophrenia research? *European Neuropsychopharmacology, 24*, 1181–1187.
Pauly, P. (1987). *Controlling Life: Jacques Loeb & the Engineering Ideal in Biology*. Oxford: Oxford University Press.
Pavlov, I. (1927/1960). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. New York: Dover Publications Inc.
Ramsden, E. (2021). Behavioral engineering and the problems of animal misbehavior. In L. Campos, M. R. Dietrich, T. Saraiva, & C. C. Young (Eds.), *Nature Remade: Engineering Life, Envisioning Worlds.* Chicago, IL: Chicago University Press.
Sherrington, C. S. (1906). *The Integrative Action of the Nervous System*. New York: Charles Scribner's Sons.
Sherrington, C. S. (1909). A mammalian spinal preparation. *Journal of Physiology, 38*(5), 375–383.
Skinner, B. F. (1931/1961). The Concept of the Reflex in the Description of Behavior. In B. F. Skinner (Ed.), *Cumulative Record*. New York: Appleton-Century-Crofts.
Skinner, B. F. (1947/1961). Current Trends in Experimental Psychology. In: B. F. Skinner (Ed.), *Cumulative Record*. New York: Appleton-Century-Crofts.
Skinner, B. F. (1961). *Cumulative Record*. New York: Appleton-Century-Crofts.
Skinner, B. F., & Crozier, W. J. (1931). Review of reflex action, a study in the history of physiological psychology. *Journal of General Psychology, 5*, 125–129.
Smith, L. D. (1995). Inquiry nearer the source: Bacon, Mach, and the behavior of organisms. In J. T. Todd & E. K. Morris (Eds.), *Modern Perspectives on B. F. Skinner and Contemporary Behaviorism* (pp. 39–50). Westport, CT: Greenwood Press.
Smith, L. D. (1996). Knowledge as power: The baconian roots of skinner's social meliorism. In L. D. Smith & W. R. Woodward (Eds.), *B. F. Skinner and Behaviorism in American Culture.* Bethlehem: Lehigh University Press.

Smith, R. (1973). The background of physiological psychology in natural philosophy. *History of Science, 11*, 75–123.
Smith, R. (1992). *Inhibition: History and Meaning in the Sciences of Mind and Brain*. Berkeley, CA: University of California Press.
Todes, D. P. (2014). *Ivan Pavlov : A Russian Life in Science*. Oxford: Oxford University Press.
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158–177.