

This is a PRE-PRINT, with actual pagination added in [square brackets]. The published version has some formatting differences. Full reference information is as follows:

Klein, Alexander (2017). “The Curious Case of the Decapitated Frog: On Experiment and Philosophy.” *British Journal for the History of Philosophy*, 26 (4): 1 – 28.

DOI: 10.1080/09608788.2017.1378866

[p. 1]

Abstract

Physiologists have long known that some vertebrates can survive for months without a brain. This phenomenon attracted limited attention until the 19th century when a series of experiments on living, decapitated frogs ignited a controversy about consciousness. Pflüger demonstrated that such creatures do not just exhibit *reflexes*; they also perform *purposive* behaviours. Suppose one thinks, along with Pflüger’s ally Lewes, that purposive behaviour is a mark of consciousness. Then one must count a decapitated frog as conscious. If one *rejects* this mark, one can avoid saying peculiar things about decapitated animals. But as Huxley showed, this position leads quickly to epiphenomenalism. The dispute long remained stalemated because it rested on conflicting sets of intuitions that were each compatible with the growing body of experiments. What eventually resolved it is that one set of intuitions supported a research program in physiology that came to seem more fruitful *on the whole*. So my case study suggests an alternative model for experimental philosophy as compared with more recent practice. Rather than using experiment to bolster our philosophical intuitions directly, we should explore how our philosophical intuitions might bolster (or block) fruitful experimental inquiry in science.

Keywords

Epiphenomenalism, Experimental Philosophy, Automata, T. H. Huxley, G. H. Lewes

As World War II drew to a close, an unassuming chicken became an American celebrity. He had survived his own decapitation, and for 18 months the headless bird could be found strutting around sideshows across the western United States. The chicken's nickname—richly earned, it seems—was “Miracle” Mike, and by the time *Life Magazine* ran a national feature on him he had become an authentic cultural icon (Stokel-Walker 2015).

Some suspected a hoax, apparently. But since the time of Aristotle, students of physiology have understood that some vertebrates can survive for [*p. 2*] months without a brain (Fearing 1930/1964, 10 – 11). Miracle Mike pulled through his botched decapitation merely by chance; but physiologists have long employed a purposeful procedure, called “pithing,” to prepare brainless animals for experimentation. In a laboratory the procedure is typically performed with a blunt needle, and one can find instructions in many physiology textbooks right through the early 20th-century.¹

The historian Franklin Fearing has said that experimentation on pithed animals “occupied the attention of almost all physiologists who lived during the second half of the 19th century” (Fearing 1930/1964, 161).² Experimentalists pithed fish, birds, and even

¹ Here is a typical example that explains how to pith a frog at the medulla oblongata:

Pithing a Frog. — Wrap a frog in a cloth and hold him head upward in the left hand. Press the front of the head downward with the left index finger, thus making a bend at the occipito-vertebral junction. Push the point of a pithing needle directly down into the depression between the skull and the first vertebra, and move the point quickly from side to side, thus cutting across the medulla. Push the needle forward into the skull cavity to destroy the brain.... (Cannon 1911/1913, 4)

Skilled experimentalists can adapt this procedure for making the cut at the location of other structures in the brain or brain stem, instead. See figure 3 for a summary of some specific incapacities that result from destroying various structures in frogs.

² Fearing's almost 90-year old work remains the most relevant and detailed secondary source on this pithing research by far. For despite the prominence of pithing experiments in 19th-century physiology, they have been almost totally ignored in recent historical work. For example, Clarke and Jacyna's otherwise thorough history mentions such experiments only in passing (Clarke and Jacyna 1987, 132). Stanley Finger's detailed work on the history of neuroscience neglects the 19th century pithing experiments (e.g., Finger 1993, 2000). Even the 952-page (Finger, Boller, and Tyler 2010) does not cover the experiments I will be discussing. There are two somewhat more recent, article-length works that investigate the conscious automaton theory (Gray 1968, Daston 1978; for more on automatism, see below, section 5), but neither discusses Pflüger and Lewes's pithing experiments in any detail. Thus as far as I can tell, Fearing's history of reflex action has yet to be superseded.

dogs. But the most frequently pithed animal seems to have been the common frog (Fearing 1930/1964, 166).

Why were so many 19th-century physiologists preoccupied with the pithed frog? The story begins with Eduard Pflüger's 1853 experiments showing that some decapitated vertebrates exhibit behaviour it is tempting to call *purposive*. The results were controversial because purposive behaviour had long been regarded as a mark³ of consciousness. Those who continued to think it *was* such a mark had to count a pithed frog—and presumably, a chicken running around with its [*p. 3*] head cut off—as conscious. Those who said purposive behaviour was *not* a mark of consciousness could avoid saying peculiar things about decapitated animals. But this view opened the way for epiphenomenalism: just as pithed frogs *seem* to act with purpose even though their behaviour is not really guided by phenomenal consciousness, so intact human behaviours may seem purposive without really being guided by phenomenal consciousness.

The dispute pitted defenders and critics of Pflüger's like (respectively) G. H. Lewes and T. H. Huxley against one another. It spurred extensive experimentation, and in the heart of this paper I will evaluate whether the experiments actually have probative value for the philosophical theses at issue in the debate. I will contend that at least as our 19th-century figures framed it, the choice of whether to accept that phenomenal consciousness influences behaviour is *not* one that can be supported directly by experimental results, despite what they thought. That having been said, the story offers some suggestions about the role experiment might legitimately play in philosophy, even today. Let me explain.

³ By “mark” I do not mean either a necessary or a sufficient condition. To borrow an example from Russell, a card's being in the catalogue may be a *mark* of the book's being on the library shelf in the sense in which I am using the term. But the card's being in the catalogue is neither necessary nor sufficient for the book's being on the shelf. Nobody I shall be considering thinks that choosing behaviour is either necessary or sufficient for consciousness. *A fortiori*, they do not think choosing behaviour *constitutes* consciousness. The card can be an extremely useful guide to telling us whether a book is on the shelf, even if cards are sometimes mistakenly left in the catalogue after the book has been checked out. Similarly, choosing behaviour might be a reliable indicator of consciousness even if there are rare circumstances where such behaviour exists in the absence of consciousness. The exceptions would have to be rare and, ideally, explainable.

The dispute remained stalemated through late in the 19th century (James 1890/1981, 138). I will contend that this is because it rested on two mutually exclusive sets of intuitions that both remained consistent with the growing body of experimental results. What finally resolved the dispute was that the competing intuitions each issued different research programs in science, and as one of these programs proved more fruitful on the whole, the corresponding set of intuitions was thereby vindicated. But the vindication did not come from any particular experiment. Indeed, it *could* not have come from any experiment, I will contend.

So my case study suggests an alternative model for what experimental philosophy might look like, today. Some philosophers have lately been using experiment to try to bolster their intuitions directly, and I will explore some of that work in the final section. But if the kinds of philosophical intuitions at play in our 19th-century story are at all common, then there is another form of experimental philosophy we might think is worth pursuing. Instead of trying to prove our intuitions through experiment, we should explore how our philosophical intuitions might bolster (or block) fruitful experimental inquiry in science.

Before turning to my case study, I will begin with some historical background.

1. Some Background: Mechanism and Animism

Descartes famously contended that living animals might be like machines in the sense of being non-conscious organisms all of whose behaviours are produced strictly mechanistically. Those in the 17th and 18th century who adopted a broadly Cartesian approach to animal physiology [p. 4] are often called “mechanists,” and their approach is typically contrasted with so-called “animists.” What separated the two groups was the issue of whether and to what extent the mechanical principles of Newton and Boyle could account for the functioning of living organisms (King 1964, 120).

Mechanists like Friedrich Hoffmann (1660 – 1742), Herman Boerhaave (1668 – 1738), and Julien Offray de La Mettrie (1709 – 1751) sought as far as possible to expand the purview of Cartesian, strictly mechanistic animal physiology to include even human physiology.⁴ Animists like Jan Baptista van Helmont (1580 – 1644) and Georg Ernst Stahl (1660 – 1734) gave the soul a causal role in the production of bodily motion, including in (what came to be called) “reflex action.”⁵ (A reflex action is an involuntary behaviour like flinching or blinking in response to an object moving towards your eye.)⁶

Even for those more inclined towards mechanism, though, animistic tendencies still underlay much physiological thinking throughout the early modern period. For instance, Giovanni Borelli (1608 – 1679) had developed a mechanistic account of how the heart pumps blood. But even Borelli gave the soul a small but important role in this motion. As Hubert Steinke has pointed out, Borelli contended that “the unpleasant accumulation of blood in the heart of the preformed embryo would be perceived by the ‘sentient faculty’ (*facultas sensitiva*) of the soul through the nerves, which would then prompt the ventricle to contract” (Steinke 2005, 22). Only *after* the process was thus initiated would the circulation continue mechanistically, as a kind of physical, acquired habit. But the ultimate cause of this motion was the soul.

By the 19th century, physiologists like Marshall Hall (1790 – 1857) began pulling more cleanly away from animism, particularly in accounting for non-voluntary motion. Inspired by Robert Whytt (1714 – 1766) among others, Hall developed an early,

⁴ One might think that expanding mechanistic physiology to humans would be a decidedly *non*-Cartesian strategy. But that is not how things seemed to some mechanists, particularly in the 19th century. For example, almost half of Huxley’s “On the Hypothesis that Animals are Automata” was devoted to arguing that Descartes was the forefather of modern physiology, the “fundamental proposition” of which is that “the living body is a mechanism.” Descartes laid the blueprint for this approach, according to Huxley, and even if Descartes demurred in applying it in the human case, he should be credited with having “opened up that road to the mechanical theory of these processes, which has been followed by all his successors” (Huxley 1874/1894, 199 – 201).

⁵ The division between mechanists and animists came out particularly sharply in an 18th-century dispute between Hoffmann and Stahl (King 1964). On La Mettrie’s relationship to Descartes, *see* (Fearing 1930/1964, 87 – 88). And for general background on Hoffmann, Boerhaave, van Helmont, and Stahl, *see* (Haigh 1984, King 1964, 1967, Fearing 1930/1964, chs. 3 – 6).

⁶ The term “reflex” was first used by the French doctor Jean Astruc in 1736; *see* (Boring 1929/1950, 35).

mechanistic account of reflex action according to which physical stimuli cause nerve signals to pass through the spinal [*p. 5*] cord and back out to produce muscular contraction directly, with no intervention from the soul.⁷ Such anti-animistic theories established the idea of a so-called *reflex arc*, which would come to play a central role in physiology and psychology. As William Clifford later put it, reflex arc theorists hold that “it is not you who wink your eye, but your body that does it” (Clifford 1874/1886, 251).

2. *Pflüger (1853)*

The central controversy I want to discuss begins with a direct challenge to Whytt and Hall’s mechanistic account of reflex action. The challenge dates to an 1853 book by the German physiologist Eduard Pflüger (1829 – 1910; Pflüger 1853). Whytt and Hall had both discussed reflex action in pithed animals at some length; Hall in particular contended that such behaviour could be accounted for purely mechanistically.⁸ But Pflüger did not agree, and to get a grip on his worries, it helps to consider some of his experiments.

Many 19th-century texts on physiology offered instructions for performing Pflüger’s most famous experiment, and they went something like this (consult figure 1 for an illustration):

⁷ Whytt’s relationship to animism is complicated. On one hand, Huxley treats Whytt as the “father” of modern animism because of the latter’s opposition to Albrecht von Haller (Huxley 1870, 4 – 5). What is more, Fearing sees a direct influence from Stahl in Whytt’s notion that a “sentient principle” controls or directs involuntary motions (Fearing 1930/1964, 78). But others have contended that since Whytt does not make the nervous system “subservient” to consciousness, he is *not* properly grouped with Stahlian animism (Rocca 2007, 88 – 89). Still, even Rocca acknowledges that Whytt generally preserves a causal role for this non-material “sentient principle” (which Fearing calls nothing but “a convenient verbal substitute for an all-pervasive soul”) in physiological processes (Fearing 1930/1964, 78), and in that sense is very much an animist. Also see footnote 8, below. In any case, Whytt’s positive influence on Hall is not in doubt; and for a discussion of Hall’s theoretical approach, *see* (Fearing 1930/1964, ch. 9).

⁸ See (Fearing 1930/1964, 162), though “purely” is probably too strong for Whytt. The latter insisted on a background role for a “sentient principle” even in reflex action, as at (Whytt 1751, 117 – 118). But despite this, Whytt is typically taken to have laid the groundwork for Hall’s more fully mechanistic account of reflex action, as at (Boring 1929/1950, 35 – 39), so much so that Hall was publically accused of failing properly to acknowledge his own reliance on the work of Whytt and other pioneers; *see* (Fearing 1930/1964, 136 – 139).

- STEP 1: Pith a frog and suspend it from a hook.⁹ Dip some filter paper in acetic acid, and touch the frog's side.
- STEP 2: Observe which foot the frog uses to wipe away the acid.
- STEP 3: Dip the frog in water to wash off the acid. Wait five minutes.
- STEP 4: Amputate the acid-wiping foot. Apply the acidulated paper to the same spot on the side again, and observe the frog's reaction. You should find that the frog will *choose a different means* to achieve the [p. 6] same end—it will choose a different foot to try to wipe away the irritant, typically, or it will rub the irritated skin against a foreign surface, if available.¹⁰

⁹ Pflüger himself apparently performed his experiments on fully decapitated (rather than merely pithed) frogs. See below, fn. 12. But later physiologists more typically worked on pithed frogs.

¹⁰ Pflüger's original discussion can be found at (Pflüger 1853, 16 – 17, 124 – 126). Just a few examples of related descriptions can be found at (Lewes 1859a, II.245 – 247, 1877, 429 – 430, Haldane 1884, 40, James 1890/1981, 22).

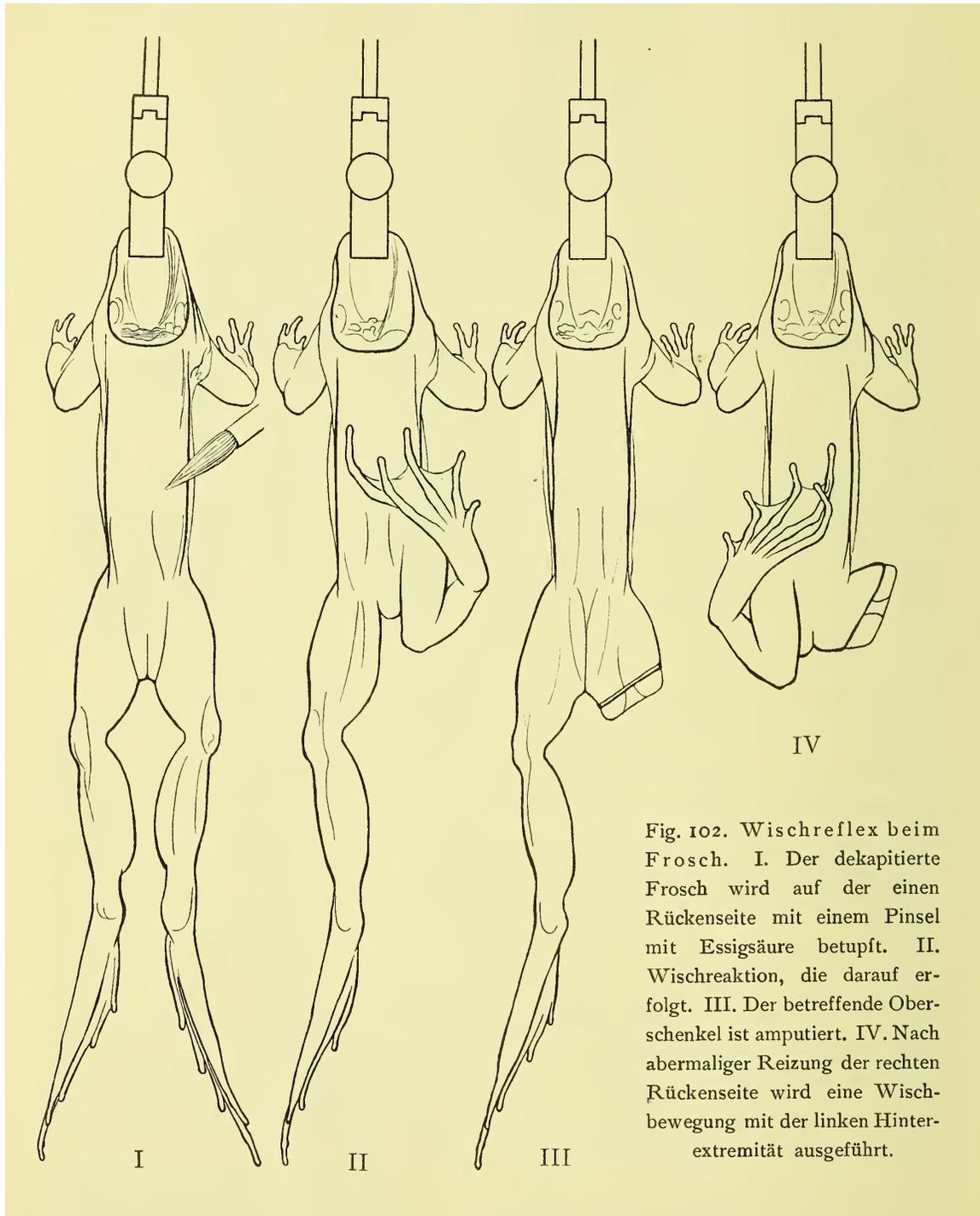


Figure 1. Pflüger's experiment, as depicted in (Verworn 1912, 198).

One might assume that the initial wiping behaviour (in step two) is merely a reflex—that is, a fully mechanistic causal chain that proceeds from stimulus through to the response without any intervention from a wilful, conscious [p. 7] mind. Indeed, for

Hall, that the frog can wipe acid in the manner of step two constituted evidence that reflex action is purely mechanical, and not modulated by a conscious mind (Fearing 1930/1964, 162). But what should we say about the frog's behaviour in step four, when its favoured foot is amputated? The frog does not reflexively wave around its stump, as one would expect if the reaction in step two had been produced purely mechanically. Instead, the frog *chooses* a novel means to achieve the same end, and it is tempting to count such choosing-behaviour as *purposive*. Now, suppose one accepts purposive behaviour as a mark of consciousness (or sensation, or volition, or all of these). Then one arrives at a surprising result indeed—that the brainless frog, properly prepared, remains a conscious agent.

Of course, there is a lot riding on just what is meant by “consciousness,” “sensation,” and “volition.” Pflüger himself often wrote about the decapitated frog's supposed “consciousness” (*Bewusstsein*), but was rather loose and poetic in spelling out what that term was to mean.¹¹ Still, his general thesis was clear enough: that in addition to the brain, the spinal cord is also an organ that independently produces consciousness. One controversial implication is that consciousness itself may be divisible (and so literally extended; see Huxley 1870, 5 – 6)—it may exist in various *parts* of the nervous system, even in a part of the spinal cord that has been divided from the brain (Fearing 1930/1964, 162 – 163).

Pflüger's provocative discussion sparked a battery of experimental investigation from others in the coming decades. As it turns out, the pithed frog is capable of remarkably complex behaviour, even more so than what we have just seen. For example, a “brainless” frog will swim if dropped in water (Lewes 1877, 190). If *completely* submerged, it will swim to the surface. And not only that; if one impedes the emerging,

¹¹ E.g., “Consciousness [*Bewusstsein*] is life and becoming. Life is motion. The causes that underlie this life are determined, and yet could not be fundamental. They are a something that at one and the same time is done and is not done. This motion, what we call consciousness, is a part of the whole, this life a part of the whole life of the world” (Pflüger 1853, x – xi, my translation).

pithed frog by putting an inverted jar in its path, the frog will not easily be trapped. It will actually *re-descend* until it can swim out of the jar, and *then* will swim up to the surface (see figure 2; Goltz 1869, 70). This is an astonishing sequence of behaviours for an animal that lacks a brain.

To give another memorable example, a frog whose cerebral hemispheres have been destroyed might be expected to have no sense of vision. And yet Goltz reported that such a frog will hop towards light from a distant window if its hind legs are irritated, even *hopping around a barrier* placed between the frog and the light. What is more, if the procedure is repeated with the barrier moved to block the original path, Goltz found that the frog [p. 8]

Fig. 5.

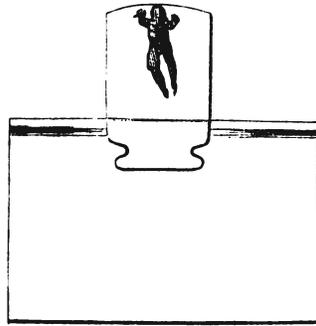


Figure 2: Goltz's pithed frog can escape underwater entrapment; from (Goltz 1869, 70).

will again simply manoeuvre around the barrier, finding another path with little apparent trouble (Goltz 1869, 65).

Now, opposition to Pflüger's original analysis had come swiftly, first of all from the philosopher, physician, and psychologist Hermann Lotze. Lotze claimed that the amputated frog's behaviour could arise as a result of the nervous system having been in some way trained or educated before the animal was pithed (Fearing 1930/1964, 164, Danziger 1980, 99 – 100). Lotze's thought was that these behaviours *seem* purposive only because they are complex. If we allow that the nervous system can *acquire* complex, reflexive actions through bodily learning, then we can maintain that these behaviours are

mechanically determined, and not guided or accompanied by any phenomenal consciousness.

The difficulty with this response is that pithed frogs find ways to solve physical challenges they cannot be supposed to have faced before being pithed. For instance, suppose one places a pithed frog on its back, holds one leg straight up, perpendicular to the body, and irritates the leg with acid. The pithed frog will then raise *the other* leg to the same, odd position so as to be able to wipe away the irritant (Huxley 1870, 3). [p. 9] Few if any frogs can have experienced such an unusual leg position before being pithed and probed in this manner in the laboratory.

In any case, I now want to turn to the way this debate would play out in the coming decades, particularly in Great Britain. One can usefully divide that ensuing controversy into two categories. First, there was considerable discussion about the facts themselves, a discussion that did resolve itself through further experimentation. But it is instructive to notice that there was another, more philosophical debate over how to *interpret* the facts, a debate that by its very nature *resisted* a purely experimental resolution. I will take the debate about the facts first.

3. *Getting Clear on Some Facts*

It should be unsurprising to find that just what behaviours the pithed frog is capable of depends on precisely what parts of the frog's brain are disabled during pithing. In the decades following Pflüger's original work, one does see a consensus on these details, but it forms slowly. For instance, one of Pflüger's key allies George Henry Lewes (1817 – 1878) began writing about these issues in 1859. But in his early contributions to this literature, Lewes reported the behaviours he and others had observed in what he simply

characterized as “decapitated” frogs (e.g., throughout Lewes 1859a II, 1859b, and 1873)¹²—he does not refer to the more careful process called “pithing.”

But later contributions to this literature took more care in specifying which structures had been disabled in various experiments, so that by the 1870s one finds converging accounts of precisely which behaviours are knocked out by destroying which specific brain structures (the *fin-de-siècle* diagram in figure 3 [p. 10])

¹² In one place Lewes writes: “my own experiments ... show that animals manifest volition after removal of the whole Brain, including the Medulla [Oblongata]” (Lewes 1859a, II.140). So, although when he reports the details of those experiments he only tells us he “cut off” the creature’s “head” (Lewes 1859a, II.176; similar language is used at 244), he apparently believed that his procedure was precise enough reliably to destroy the brain and the entire brain stem. Pflüger also frequently discusses the “beheaded” (*enthauptet*) or “decapitated” (*geköpfte*) frog (Pflüger 1853, e.g. x, 14, 55, 128). He says he decapitated his frogs using scissors, writing that “it is quite indifferent where one beheads the frog, so long as one cuts below the atlas, which is to say the medulla oblongata” (Pflüger 1853, 123). Later on, Lewes would become much more careful to identify brain structures, resorting to more precise decerebration procedures than full decapitation (Lewes 1877, II.429). In the English-speaking world, Lewes was the chief defender of Pflüger’s position, and it may be that the general vagueness in the early English-language literature about which brain structures were being destroyed in “decapitated” animals is a symptom of Lewes’s influence on how this debate played out in the UK (e.g., see Waters 1859, 690, Duncan 1869, 33 – 41). In German speaking lands, more precise pithing techniques were soon being employed by people like Goltz, who used a “sharp lance” (Goltz 1869, 56) in combination with other tools to destroy more precise areas of the brain and brain stem.

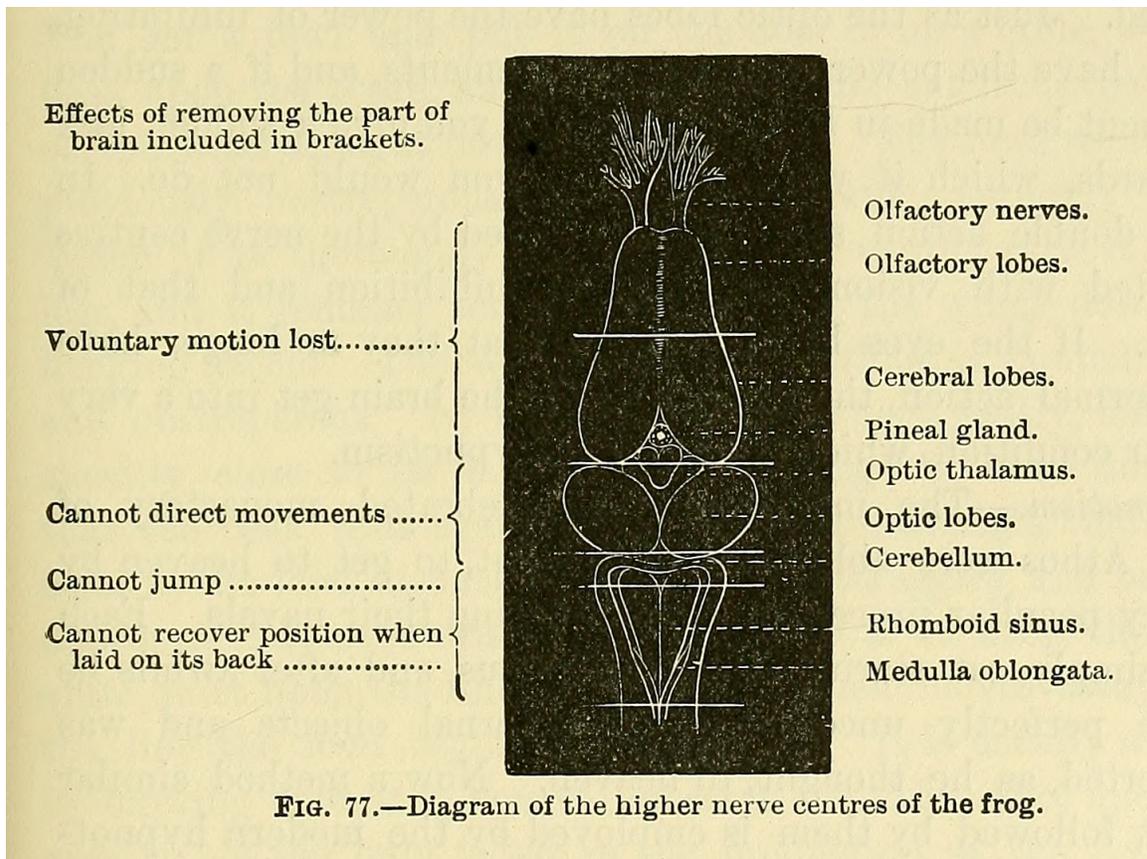


Figure 3" Frog brain structures, along with summary of pithing results (Brunton 1898, 227).

summarizes some of these results). For example, T. H. Huxley reports that a frog that is pithed *above* the medulla oblongata (but below the cerebellum) loses the ability to jump, even though the frog with the brain stem and cerebellum both in tact is able to perform this action, at least in response to irritation. A frog pithed just below the cerebrum “can see, swallow, jump, and swim,” though still will typically move only if prompted by an outer stimulus (Huxley 1870, 3 – 4).

One can see why researchers wanted to get a grip on coordinated behaviours like jumping or swimming. The most clearly purposive behaviours, after all, involve a coordination of disparate movements (think of the frog escaping the inverted jar).

Although Huxley does not emphasize this point, some of Pflüger’s careful defenders would show that frogs clearly *are* capable of complex behaviour even with everything *above* the medulla oblongata destroyed (but with the medulla in tact). Perhaps such a frog

cannot jump, as Huxley reported. But it can regain its normal posture if placed on its back (Lewes 1877, 164 – 165). True, there was already a consensus that severing *below* the medulla renders the frog seriously incapacitated—*then* it can no longer flip itself over, and it ceases even to breathe on its own. But surprisingly, *Pflüger's classic experiment* works on frogs whose spinal cord is [*p. 11*] severed below the medulla.¹³ What is more, such a frog can even perform the task I described above, where an acid irritant is wiped from a leg that is artificially extended perpendicular to the body.¹⁴ Surely that is a complex behaviour.

What I want to emphasize is that as soon as replicable experiments established what abilities the frog loses when specific brain structures are damaged, consensus seems to have formed easily. One does not have to believe that there are *experimenta crucis* to notice that there were few background assumptions at play that might have been dubious enough to be worth debating. Thus I would like to call these questions straightforwardly empirical. Answers to these questions come relatively directly from experiment.

4. *Getting Clear on Interpretation? Mechanist's Dilemma*

But what about the big question that motivated this debate to begin with—namely, whether the spinal cord itself is an organ of consciousness (or sensation or volition)? I will argue that this is *not* a straightforwardly empirical question because any answer must rest heavily on brute and inherently untestable intuitions.

To see this, it helps to tease out some conceptually prior questions that arise in the work of Pflüger's most able ally, the English literary critic, philosopher, and physiologist G. H. Lewes. In this section, I will tease out two of these in particular.

¹³ See above, fn. 12.

¹⁴ Inexplicably, Huxley himself acknowledged that the medulla-free frog performs this unusual behaviour, and in the very same article where he had denied that frogs without a medulla can perform “complex” actions; *see* (Huxley 1870, 3).

Lewes had repeated Pflüger's procedures and had created some new experiments of his own. But he was sophisticated enough to understand that the experiments do not speak for themselves about the existence of sensation and volition in the pithed frog. (Lewes tended to focus on "sensation" and "volition" rather than on "consciousness," the latter of which he regarded as an ambiguous term to be used only with special care; see Lewes 1877, 353 – 355.) He saw that one must first settle an important issue before it is possible to interpret these experiments. He wrote,

we have no *proof*, rigorously speaking, that any animal feels; none that any human being feels; we *conclude* that men feel, from certain external manifestations, which resemble our own, under feeling; and we conclude that animals feel—on similar grounds.

Now, inasmuch as the actions of animals furnish us with our sole evidence for the belief in their feeling, and this evidence is universally considered as scientifically valid, it is clear that similar actions in decapitated animals will be equally valid; and when I speak of *proof*, it is in this sense. *Spontaneity* and *choice* are two [*p. 12*] signs which we all accept as conclusive of sensation and volition. (Lewes 1859a, II.237 – 238, italics original)

Does Pflüger's experiment prove that there is sensation or volition in the pithed frog? We cannot tell, Lewes suggests, until we first settle on some third-person-accessible mark of sensation and volition. And the marks Lewes proposes are spontaneity and choice. "Our task," he continues, "is to ascertain whether animals, after decapitation, manifest these palpable signs" (Lewes 1859a, II.239). Obviously, we cannot occupy an animal's consciousness ourselves to check for sensation and volition from a *first-person* perspective. In this sense, the question of what behavioural mark of subjective experience we should accept is conceptually prior to the question of whether the cord is in some

sense an “organ” of subjective experience.¹⁵ I will have more to say about this issue in section five, below.

Now what does Lewes mean by “sensation” and “volition”? Let us take “sensation” first, which Lewes clearly thinks connotes a *subjective* phenomenon.

Do what we will, we cannot altogether divest Sensibility of its psychological connotations, cannot help interpreting it in terms of Consciousness; so that even when treating of sensitive phenomena observed in molluscs and insects, we always imagine these more or less suffused with Feeling, as this is known in our own conscious states. (Lewes 1877, 188 – 189)

Here and elsewhere Lewes uses “sensibility” to mark out a subjective experience that contributes to the phenomenal character of an occurrent mental state. Sensations can sometimes go “unperceived,” but they still contribute to the overall “stream of Consciousness” (a term Lewes invented; see Rylance 2000, 10 – 11) just as stars at midday contribute to the ambient light without themselves being discriminable (Lewes 1859a, II.63, 65 – 68).

Idiosyncratically, Lewes counts *every* action of a living organism as an act of “sensibility.” This is because he thinks that every living organism “has Sensibility as its vital property” (Lewes 1877, 374). So it is not just that *some* indiscriminable or “unperceived” sensations influence the tone of the “stream of Consciousness.” For Lewes, *every* physiological change is in some sense sensory, and every physiological change thereby influences the “stream of Consciousness,” however slightly.

It follows that even “involuntary” actions must in some sense be accompanied by sensibility (Lewes 1859a, II.224, 1877, 374), and indeed this is precisely why he thinks the distinction between “voluntary” and “involuntary” actions is one [*p. 13*] of degree and not of kind (Lewes 1877, 367). It is a distinction we draw “for convenience,” not

¹⁵ Pflüger also emphasizes the need for settling on a criterion prior to performing these experiments (Pflüger 1853, 123).

because there is some fundamental difference between two types of muscular responses (Lewes 1877, 378).

What is Lewes's basis for this "useful" distinction (Lewes 1877, 374)? A "voluntary" action is one that has *actually* been performed with a relatively high "degree" of conscious "control" (Lewes 1877, 365, 367 – 373, 375, also see Lewes 1859a, II.197 – 199, 201 – 204). Lewes presents his approach as importantly new, since traditionally, actions merely had to be *controllable* to count as voluntary (Lewes 1859a, II.201). His central objection to the traditional way of drawing the distinction between the voluntary and involuntary is that *every* physiological action can supposedly come under conscious control given appropriate training.¹⁶ In short, Lewes uses "volition" to connote a phenomenally conscious sensation that *in fact* has a causal impact on behaviour.

If this is right, then the question of whether there is a mark of consciousness really splits into two. We have the question of whether there is a behavioural mark that indicates the mere presence of phenomenally rich, subjective *sensation*. And we have the question of whether there is a mark that indicates that a given behaviour is in some way *controlled* or *impacted* by a subjective sensation—in other words, whether there is a mark of what Lewes calls "*volition*."

We have seen that Lewes regards *spontaneity* as a mark of sensation and *choosing behaviour* as a mark of volition. With that standard in place, he offers an initial argument in support of the spinal cord as an organ of sensation and volition (Lewes 1859a, II.239 ff., 1877, 429 – 430). We can reconstruct the argument for volition as follows. If one finds choosing behaviour in an animal, then one must attribute volition to the animal. Pflüger's classic amputating experiment provides a clear example of choosing behaviour in pithed frogs. So we must attribute volition to these pithed frogs. Since such frogs

¹⁶ As evidence he cites the role of training in controlling supposedly involuntary actions like "breathing," "the rhythmic movements of the heart and the contractions of the iris," "certain movements of the toes," moving the ears, and "winking" (Lewes 1877, 369 – 372).

exhibit choosing behaviour even when the cord has been severed from the brain stem below the medulla oblongata, volition cannot be a capacity afforded exclusively by brain or brain-stem structures. There is no remaining organ other than the spinal cord that could plausibly cause volition in pithed frogs. Hence the spinal cord plays a causal role in producing volition in pithed frogs.

The reasoning here is valid, but what is controversial is the truth of that crucial *if ... then* premise at the start. Why, we might now wonder, should anyone accept that choosing behaviour *is* a mark of volition? [p. 14]

Lewes had an interesting strategy for bolstering this claim. Rather than directly arguing that he had identified the “right” behavioural criteria for conscious control of behaviour—i.e. that choice in fact goes along with volition invariably—he instead offered a modified version of the above argument that only required a generalized version of that *if ... then* premise.

The generalized argument took the form of a dilemma for opponents of spinal sensibility. “Experiment leads decisively to this alternative,” Lewes wrote. “[E]ither [intact] animals are unconscious machines, or decapitated animals manifest sensibility and will” (Lewes 1859b, 136). This is a dilemma Lewes returns to frequently in his writing on spinal sensibility and volition (e.g. Lewes 1877, 431 – 432, 1859a, II.178 – 179).

We can reconstruct the dilemma as follows. One can either accept or reject behavioural marks of sensation and volition.

Horn A: If one accepts *any* behavioural mark of sensation and volition, then the experimental data will force us to attribute sensation and volition to *both* decapitated and intact vertebrates alike.

Horn B: If one rejects the existence of a behavioural mark, then one has no grounds for ascribing sensation or volition to either decapitated *or* intact vertebrates.

I shall call this “the mechanist’s dilemma.”

This line is more general than the initial, Pflüger-style reasoning I just sketched in the sense that the mechanist's dilemma does not rely on any claim about what *particular* feature of behaviour might be a reliable mark of sensation or volition. Instead, Lewes suggests that accepting *any* behavioural mark will force one to attribute sensation and volition to the pithed frog.

One might suspect that Lewes has merely pushed the problem under the rug—for what evidence does he have, now, for the *if ... then* claim in Horn A? What underwrites his confidence that *any* behaviour worthy of being treated as a mark of sensation or volition is a behaviour he can elicit in the pithed frog? I take it this *if ... then* assertion is not offered as an a priori, conceptual truth. And it is not quite meant as an empirical truth, either. Rather, step two amounts to a challenge. Give me *any* reasonable mark of sensation and volition, Lewes says, playing the role of confident experimentalist; I will then produce the relevant behaviour in the pithed frog.

He was constantly pressing this dilemma, and sought to illustrate it through a battery of experiments. An early example involves an otherwise intact frog whose spinal cord has been cut about halfway down the back, between the fifth and sixth cervical vertebrae. Once the frog has recovered from the surgery, a casual observer would simply say that paralysis has set in below the cut. The hind legs seem totally incapacitated. If one irritates [p. 15] the frog's anterior, it will crawl away using its front legs only, dragging the seemingly lifeless hind legs along. But if the tail or *hind* legs are irritated with acid, the entire *posterior* section begins to move, attempting to initiate a crawling action. The attempt does not succeed, since in this case the *front* legs lie motionless, and the back legs are not strong enough to overcome the resistance.

What is interesting is not the result itself—physiologists had long observed what they considered reflex action in paralyzed limbs (e.g., human paraplegics withdraw their feet if their soles are tickled). What is interesting is the lesson Lewes drew from the case. Suppose we propose complex action (rather than purposiveness) as a mark of volition.

We cannot say that it follows *from* the fact that the frog exhibits no complex action in its hind legs when the anterior part of the frog is irritated *that* those hind legs have lost all volition. For that same reasoning would then have to be applied to the front legs, which lie motionless when the hind legs are attempting to crawl. The experimentalist must *either* attribute conscious control to *both* the front and back legs (that is Pflüger and Lewes's position), or she can attribute conscious control to *neither* (Lewes 1859b, 137).¹⁷

The principle at play is that if we pick a behavioural mark of conscious control, *we must apply it evenly*. If we do, Lewes is confident that he can elicit that mark in the behaviour of a pithed frog. This would force us to attribute sensation and volition to both intact and pithed frogs alike.

We have been considering Horn A of the mechanist's dilemma. Lewes clearly intended Horn B to be the pointy one. If one *denies* that there is an observable, behavioural mark of sensation and volition, then one seemingly loses any grounds for asserting the existence of sensation and volition even in healthy, intact vertebrates (including not just frogs, but even our fellow human beings). Lewes perhaps assumed that mechanistically-inclined physiologists would find the wholesale denial of sensation and volition even to intact vertebrates to be intolerable. But by the early 1870s, some mechanists were quite prepared to choose the pointy horn of the dilemma, in a sense. They offered a completely different argument for the existence of sensation in vertebrates; and they were happy to *reject* the existence of volition (in Lewes's sense) altogether. Thus was born the *conscious automaton* theory.

5. *Getting Clear on Interpretation? Spinal Consciousness*

T. H. Huxley offered the most influential and provocative version of the conscious automaton theory in an 1874 address in Belfast. According to this view, "consciousness"

¹⁷ Lewes uses a different experiment for similar rhetorical purposes at (Lewes 1877, 428).

(Huxley's word—his usage is roughly synonymous [*p. 16*] with Lewes's "sensation") accompanies the body without acting on it, just as "the steam-whistle which accompanies the work of a locomotive engine is without influence upon its machinery." Conscious states are continually being caused by brain states from moment to moment, on this view, but are themselves causally inert. In other words, although Huxley accepted the existence of sensation, he rejected the existence of "volition" (as Lewes had used that word).¹⁸ This is an early form of epiphenomenalism.¹⁹

Pflüger and Lewes had indeed established the existence of purposive behaviour in pithed frogs, Huxley readily conceded (Huxley 1874/1894, 223). But since it is *absurd* (according to Huxley) to think the behaviour of brainless frogs is under conscious control, the correct lesson to draw from Pflüger and Lewes's results was that purposive actions are *not* sufficient to establish volition. In fact, Huxley evidently was unwilling to accept the existence of *any* behavioural mark of either sensation or volition.

So if Huxley denied the existence of any behavioural mark, what grounds did he offer for ascribing *sensation* to vertebrates? He held that the existence of sensation is obvious in his own case. And he suggested that what he called "the doctrine of continuity" in evolution forbade him to posit that a trait as "complex" as consciousness could appear suddenly on the phylogenetic tree (Huxley 1874/1894, 236). So we should think that phenomenal consciousness is likely to exist at least in other humans and in non-human vertebrates, Huxley argued. This ascription of phenomenal consciousness does not depend on recognizing any behavioural mark at all.

¹⁸ Huxley himself used the word "volition" differently than Lewes. For Huxley, "volition" connotes "an emotion indicative of physical changes, not a cause of such changes" (Huxley 1874/1894, 240). Huxley accepted the existence of this epiphenomenal variety of volition, but rejected the existence of volition in Lewes's sense. I will continue to use "volition" in Lewes's sense.

¹⁹ Perhaps the original epiphenomenalist during this era was Shadworth Hodgson, who adopted the view in (Hodgson 1870, I.416 ff.). William Clifford was sympathetic with Huxley's epiphenomenalism, but was not himself an epiphenomenalist. For Clifford, there is a "parallelism" between mental and physiological causal chains, "but there is no interference of one with the other" (Clifford 1874/1886, 262).

There are two interpretive issues between Huxley on one side, and Pflüger and Lewes on the other, that need to be teased apart. When it comes to the issue of whether any observable properties of behaviours (such as choosing) are marks of either sensation or volition, Huxley has effectively chosen Horn B of the mechanist's dilemma. He was happy to say that there are *no* marks of consciously controlled behaviour, and then to accept that no behaviour is consciously controlled.

But the bodily *seat* of sensation (phenomenal consciousness) is a separate issue. *Prima facie*, Huxley could coherently have taken Horn B of the mechanist's dilemma, but still agreed with Pflüger and Lewes that the spinal cord produces sensation. But that was not what happened. [p. 17]

Huxley insisted that the brain is the sole organ of phenomenal consciousness. He pointed out that in humans, total loss of consciousness results from serious injury to “the anterior division of the brain” (Huxley 1874/1894, 220; I assume he meant the frontal lobe of the cerebrum, though he may have meant the entire cerebrum). On that basis, he characterized the “brain” as the “seat” of human consciousness (Huxley 1874/1894, 205, 1872, 14). He then insisted that for broadly evolutionary reasons, we should expect that what is true of humans is true of other vertebrates (Huxley 1874/1894, 221, 236). The upshot (*contra* Pflüger and Lewes) is that we should assume that “a frog's spinal cord is not likely to be conscious” since “a man's is not” (Huxley 1874/1894, 222).

The ball was back in Lewes's court, at least when it came to the question of a so-called “spinal soul.” Three years later, he responded by rejecting the notion that an organ can be found to be the “seat” of some vital process if the process ceases whenever the organ stops working. Such a view would require us to say something absurd, Lewes pointed out: “that the medulla oblongata was the ‘organ’ of Respiration, because Respiration ceases when this centre is destroyed” (Lewes 1877, 162). Again, the medulla oblongata is in the brain stem. Successful respiration in most vertebrates requires that this structure functions properly, but surely it requires that lungs, blood, and other bodily

structures function properly as well. What is more, Huxley's claim that a human becomes "unconscious" when her cerebrum is damaged does not even attempt to address Pflüger and Lewes' suggestion that the spinal cord might produce some separate consciousness *of its own*, a consciousness that (for all we can tell) might persist even when the *brain's* consciousness lapses.

Huxley in fact anticipated the impasse, and had actually acknowledged that Pflüger and Lewes's position on spinal consciousness could not be defeated by any rational argument. But their position should obviously, Huxley contended, be regarded as absurd nevertheless:

It must indeed be admitted, that, if any one think fit to maintain that the spinal cord below the injury is conscious, but that it is cut off from any means of making its consciousness known to the other consciousness in the brain, there is no means of driving him from his position by logic. But assuredly there is no way of proving it, and in the matter of consciousness, if in anything, we may hold by the rule, '*De non apparentibus et de non existentibus eadem est ratio*' ['what does not appear and what does not exist have the same evidence'].²⁰ (Huxley 1874/1894, 220) [*p. 18*]

So even while seeming to offer experimental evidence that the brain alone is the organ of consciousness, in the next breath Huxley conceded that experiment *could not* settle this debate. Huxley simply pronounced it absurd to say that the spinal cord has its own consciousness, claiming that such a consciousness could never be *observed*.²¹

But notice something interesting. Huxley's pronouncement piggybacks on the position he took in the mechanist's dilemma. His claim that spinal consciousness cannot be *observed* amounts to the claim that such a consciousness cannot be observed *first-*

²⁰ One commonly finds the Latin phrase used in legal contexts, where it is standardly translated somewhat loosely as "what is not juridically presented cannot be judicially decided."

²¹ William Carpenter offered similar reasoning that same year; *see* (Carpenter 1874, 69 – 70).

personally. But that is the crux of the mechanist's dilemma. Remember that Huxley could not accept Pflüger and Lewes's reasons for thinking the spinal frog had either sensation or volition precisely because he (Huxley) was unwilling to accept a *third-personally* observable, behavioural mark of sensation or volition, such as purposive action. Thus the argument Huxley was deploying against spinal consciousness in fact depended on his response to the mechanist's dilemma. And if that is right, then the basic disagreement between the two parties is actually over whether there is a behavioural mark of consciousness.

The question now becomes whether there is a way to establish the existence of a behavioural mark of consciousness by way of experiment. The answer is a clear no, I think, and for reasons that have cropped up again in more recent consciousness research. Here is David Chalmers:

Consciousness just is not the sort of thing that can be measured directly. What, then, do we do without a consciousness meter? ... How does all this experimental research proceed? I think the answer is this: we get there with principles of *interpretation*, by which we interpret physical systems to judge the presence of consciousness. We might call these *pre-experimental bridging principles*. ... We cannot reach in directly and grab those experiences, so we rely on external criteria instead. That is a perfectly reasonable thing to do. But something interesting is going on. These principles of interpretation are not themselves experimentally determined or experimentally tested. In a sense they are pre-experimental assumptions. (Chalmers 2010, 91 – 92, my underline, original italics)

I have just argued that it was precisely because Huxley rejected the existence of this sort of behavioural mark that he rejected the existence of spinal consciousness. But no experiment could establish whether there *is* a behavioural mark of consciousness because, as Chalmers puts it, we have no “consciousness meter.” We cannot test which behaviours are correlated with conscious experiences because we have no independent way to know

when phenomenal consciousness is actually present (in cases other than our own). So no experiment could force us to accept or reject the existence of a general behavioural mark in the first place. [p. 19]

The dispute over whether the spinal cord produces consciousness comes down, therefore, to two brute stipulations that are mutually incompatible, yet separately consistent with any related experiment one could hope to run, apparently. On one side, Pflüger and Lewes stipulate that spontaneity and purposive behaviour are marks of sensation and volition. And once they accept this standard, then (and only then) can they rationally interpret their own experiments as supporting the existence of sensation and volition in pithed frogs. On the other side, Huxley stipulates that there is *no* behavioural mark of sensation or volition. And once he has rejected the existence of any such standard, there is no behaviour one *could* ever elicit in a laboratory that would count as evidence of either sensation or volition in a pithed frog.

In short, unless we stipulate a behavioural mark, nothing (i.e., neither the presence nor absence of any behaviour) can count as experimental evidence either for or against the existence of consciousness (i.e., of sensation or volition). We can therefore say that the acceptance of a behavioural mark is a *presupposition* of there being meaningful, experimental evidence one way or the other concerning sensation or volition.²²

It should not be surprising that the debate between the two sides thus came to seem intractable.²³ If one has the intuition that purposive behaviour (say) is a reliable mark of phenomenal consciousness, *then* these experiments exert evidentiary pressure to ascribe consciousness not only to the intact frog but to the pithed frog as well. Instead, if one has the intuition that there is *no* reliable, third-person mark of consciousness, then one can easily avoid claiming that the pithed frog is conscious. But in that case one has just as

²² For more on the sense of “presupposition” I have in mind, *see* (Friedman 2001, 74).

²³ I have been concerned with purely theoretical reasons for the apparent stalemate. Readers interested in the broader, cultural implications of the kinds of physiological disputes Lewes and Huxley were engaged in do well to consult (Rylance 2000, esp. chs. 3 and 7).

little reason to think the apparently purposive behaviour of the *intact* frog is any more controlled by consciousness than is its behaviour after decapitation. This is what an exasperated William James meant in 1890 when he wrote, “it is clear that such arguments as these can eat each other up to all eternity” (James 1890/1981, 137 – 138).

6. *A De Facto Resolution*

Critics might worry that I am overreaching when I claim that *no* experiment could ever show us whether there is a mark of either sensation or volition, and if so what the “right” mark is that we should look for. After all, it is obviously hard to know what would happen if we were to inquire indefinitely. But in the case of our historical dispute over pithed frogs, we at least have a record of how the dispute was *in fact* resolved. And despite a half century of experimentation, it was not an experiment that finally resolved the dispute. [p. 20]

Recall that Pflüger published his original work in 1853, and we have traced the way the debate developed through Lewes’s response to Huxley in 1877. By that time, the stakes had grown considerably. If one goes back to the days of Marshall Hall, the question was how to account narrowly for reflex action. But Hall had conceived of reflexes as only one of four basic types of muscular action (the other three he called “voluntary,” “respiratory,” and “involuntary”; Hall 1833, 638). Pflüger had initially attacked Hall’s mechanistic account of reflex action, specifically.

But by the time of the Lewes contribution from 1877, the question was no longer whether this one *subset* of muscular action could be accounted for purely mechanistically. Now, the question had become whether the mechanistic approach to reflex action might be expanded to cover all muscular action. Lewes wrote that the “Reflex Theory” had become a strategy where one attempted to specify “the elementary

parts involved” in *every* physiological function without ever appealing to “Sensation and Volition” (Lewes 1877, 354).²⁴

A kerfuffle over beheaded frogs had thus turned into a controversy over the very nature of physiological explanation. That controversy was, however, eventually resolved. “That the majority of physiological opinion by the close of the century was in favour of the position of Pflüger’s opponents seems certain,” Fearing writes. “Mechanistic physiology and psychology was firmly seated in the saddle” (Fearing 1930/1964, 185).

The concept of a mechanistic reflex arc came to dominate not just physiology, but psychology too. The behaviourist B. F. Skinner, for example, wrote his 1930 doctoral dissertation on how to expand the account of reflex action to cover all behaviour, even the behaviour of healthy organisms.²⁵ Through the innovations of people like Skinner and, before him, Pavlov, behaviourism would establish itself as the dominant research paradigm. I want to suggest that what drove physiologists to Huxley’s side in this old debate were broad, programmatic concerns. Research in physiology and psychology that sought mechanistic, reflex-arc explanations just *flourished*.

We get a hint of the programmatic nature of this drift towards Huxley from a 1911 textbook by the Harvard physiologist W. B. Cannon. Cannon [*p. 21*] had just given directions on how students can perform Pflüger’s old experiment on the pithed frog. But unlike the 19th century discussions we have talked about, Cannon’s treatment was not followed by any tortured rumination on whether there is a spinal soul, or on whether purposiveness is a mark of consciousness. Instead, we get this simple pronouncement:

²⁴ Lewes’s suggestion that the reflex theory had expansionist ambitions was not at all unique. For example, James had characterized Huxley’s epiphenomenalism as “an inevitable consequence of the extension of the notion of reflex action to the higher nerve centres” (James 1879, 2).

²⁵ Skinner once wrote that Bertrand Russell had “pointed out that the concept of the reflex in physiology had the same status as the concept of force in physics” (quoted at Smith 1986, 264 – 265), and the comment is apt. The reflex arc, mechanistically understood, had become a presupposition of empirical research in both physiology and psychology in roughly the way purposiveness had been a presupposition of more animistic approaches to the field (see above, p. 19). Elsewhere I have explored “presuppositions” in empirical psychology; see (Klein 2008). On the way in which the concept *force* functions as an a priori presupposition in Newtonian physics, see (Friedman 2001, 35 – 37).

Purposive movements are not necessarily intended movements. It is probable that reaction directed with apparent purposefulness is in reality an automatic repetition of movements developed for certain effects in the previous experience of the intact animal. (Cannon 1911/1913, 38)²⁶

That's it. Cannon gave no real argument for *why* students should not regard purposive movement as a mark of genuine volition (beyond a quick gesture at Lotze's long-discredited retort to Pflüger). Without citing any actual experiments, Cannon simply reported, as settled scientific fact, that purposiveness does not entail intended action.

It was not established *fact*, as I have tried to show, but rather an established convention. By discarding the burden of a behavioural mark of conscious control, physiologists freed themselves up to model *every* bodily motion as a fully mechanistic transaction between stimulus and response, without any super-physical input from phenomenal consciousness. But this was a move motivated by practical convenience, not by any particular experiment.

7. *On Experiment and Philosophy, Then*

I began by characterizing the debate about decapitated frogs as involving incompatible *philosophical* intuitions. But where, one might wonder, is the *philosophy* in all of this?

We can distinguish two philosophical issues that are intertwined in the pithed frog debate. On the one hand, there is the question of whether non-human vertebrates have what we now call “phenomenal consciousness.”²⁷ Here the issue is whether there is “something it is like” for, say, a frog to swim out from under an enclosure, or instead whether animals are unconscious automata, as Descartes thought. On the other hand, there are questions about the nature of what is now called the *cognitive control of*

²⁶ Similar interpretations crop up even earlier, for instance in (Flint 1888, 600).

²⁷ The canonical account of what phenomenal consciousness amounts to is (Block 1995).

behaviour, or “the ability to coordinate thoughts and actions in relation with internal goals” (Koechlin, Ody, and Kouneiher 2003).²⁸ Everyone [p. 22] involved in our 19th-century debate accepts that vertebrates have phenomenally conscious mental states (I have been calling these states “sensations,” following Lewes’s usage); and they all accept that vertebrates are typically capable of cognitive control. The question is whether phenomenal qualities *themselves* play a causal role in the control of behaviour (that is, whether vertebrates have what Lewes calls “volition”).

Contemporary psychologists may well have abandoned this hybrid question, but it is still alive today in philosophical debates about epiphenomenalism. As a leading proponent of this latter doctrine puts it, epiphenomenalism is the view that “[p]henomenal consciousness is inefficacious” (Robinson 2004, 159), and not just for the specific purposes of cognitive control—inefficacious for producing any bodily changes of any kind. This is the central philosophical issue I take to have been up for debate in our 19th-century controversy.

I have argued that Huxley’s epiphenomenalism did not ultimately rest on the experimental results he had adduced in favor of his view. Rather, his view rested on a brute intuition. If one has the intuition that choosing behaviour (for example) is a reliable mark of conscious control, then these experiments exert evidentiary pressure to ascribe conscious control to intact and pithed frogs alike, as well as to healthy human beings. But then one has effectively adopted interactionism. Instead, if one has the intuition that there is *no* reliable, third-person mark of conscious control, then one has no evidence that the pithed frog’s behaviour is consciously controlled. But in that case one has just as little evidence for thinking the behaviour of *intact* frogs or even intact humans are controlled by consciousness. And then one has effectively adopted epiphenomenalism.

²⁸ I thank Luca Barlassina for pushing me to distinguish these two issues.

Epiphenomenalism and interactionism can “eat each other up to all eternity,” at least as Huxley and Lewes framed the issue, because even as physiologists amassed voluminous experimental results about pithed vertebrates, those results remained perfectly consistent with either set of intuitions. But it now seems that those intuitions are closer to methodological stipulations than to any independently contentful statement about how the world is. So is there a fact of the matter about whether epiphenomenalism or interactionism is actually *right*?

If my analysis is correct, then these positions come down to incompatible methodological stipulations that only masquerade as empirical claims. And such stipulations are not themselves truth-apt. They are conventions more appropriately judged on the basis of pragmatic considerations concerning the wider research programs they support. Indeed, it looks as though Huxleyan methodological mechanism in fact issued in a theoretical paradigm that was (for a good while) more empirically adequate with respect to the wider body of experimental results in physiology and psychology, on the whole, and perhaps was more fruitful, useful, and [*p.* 23] so on.²⁹ The attention-grabbing *philosophical* claim—epiphenomenalism—got carried along as a kind of free rider, and then only to the very limited extent that behaviourists wished to talk about philosophical questions about the mind at all.

Behaviourism, with its mechanistic paradigm, flourished until the latter part of the 20th-century. And then, as we have just seen quite explicitly in Chalmers, one finds a different set of methodological stipulations (he calls them pre-experimental assumptions) supporting a new, *interactionist* science of consciousness.³⁰ Thus epiphenomenalism and interactionism are both by-products of a package of procedural assumptions needed to get

²⁹ On the sort of pragmatic considerations about theory choice I have in mind, see (Kuhn 1977, Friedman 2001, 2002).

³⁰ Of course there are many other currents in late 20th-century psychology that departed from behaviourism. I am not implying that consciousness science is the most important, or even that behaviourism was the only viable research program in the middle of the century. For a criticism of simplistic yet common divisions of the history of psychology into three or four canonical periods, see (Costall 2006).

competing experimental inquiries off the ground. It is misleading to see them as making contentful claims that could be directly tested on their own.

8. *On Experiment and Philosophy, Now*

Finally, the role of experiment in philosophy has been much discussed lately because of the rise of so-called experimental philosophy, or x-phi. So it is fitting to close by contrasting these newer uses of experiment with those we find at play in our 19th century debate.

X-phi is a 21st-century movement united not just by practitioners' use of empirical data to untangle philosophical knots, but by their use of experiment to study intuitions in particular (Sytsma 2014b, 1). The rationale for this recent experimental work has a lot to do with the foundational place of intuitions in more traditional analytic philosophy. When we construct a philosophical argument, we must be willing to grant that *some* premises do not themselves need to be supported by further argument. That is, it seems we must accept that some premises are simply “intuitive” or (as I am using the word) obvious. Much of the x-phi literature today amounts to an empirical investigation into whether the intuitions of professional philosophers are actually shared by non-professionals.³¹

To take one example relevant to this paper, consider a recent, x-phi response to the familiar “hard problem” of consciousness. As Chalmers originally conceived it, the hard problem takes off from the claim that every conscious state has its own *phenomenal feel*—again, that there is “something it is [*p. 24*] like” to see red, or hear a clarinet play a *b^b*, or feel hunger pangs, or whatever. Even if we could explain how the brain achieves each cognitive function—how it gathers colour or auditory or bodily information, for example—there would still remain a “hard” question: “Why should physical processing give rise to a rich inner life at all?” as Chalmers puts it (Chalmers 1995, 201).

³¹ One can consult numerous collections for an overview of this movement: for example, (Knobe and Nichols 2008, 2014). And (Sytsma 2014a) is a kind of textbook introduction to the field.

Critics have asked why we should think conscious states all *have* a phenomenal feel. Chalmers acknowledges that the issue comes down to a brute intuition (Chalmers 1996, xiii). He says that phenomenal feels are “the most central and manifest aspect of our mental lives” (Chalmers 1995, 206). In other words, he thinks it is just obvious that our conscious states all *feel like something*.

But some recent work in x-phi challenges the alleged obviousness of this claim. Sytsma and Machery administered a series of surveys to determine whether “ordinary people (viz. people without training in philosophy or in consciousness studies)” (Sytsma and Machery 2010, 299) agree with most philosophers that conscious experiences all have a phenomenal feel in the relevant sense. And they found some striking evidence of divergence—ordinary people are apparently *less* inclined than philosophers to say that there must be some phenomenal feel to every mental state.³²

Whether or not one finds their line persuasive, notice the stark difference between the ways x-phi practitioners and our 19th-century figures use experiment. Contemporary philosophers are testing what we might call the *intuitiveness* of our intuitions. They want to establish (or more typically, refute) claims about the obviousness of some philosophical proposition by examining what people outside the bubble of professional western philosophy might say about that proposition. In contrast, we have seen that the 19th-century figures in our story were not testing how widespread their own intuitions were. They were trying to produce direct, experimental evidence for claims they previously did not think *required* such evidence.

Unfortunately, the intuitions at issue in our 19th-century debate ultimately resisted this kind of verification, as I have argued. But their *attempted* verification should teach us a lesson about how philosophy might make fruitful contact with experimental results nevertheless.

³² For a series of responses and rejoinders to this work, see (Talbot 2012b, Sytsma and Machery 2012, Talbot 2012a).

Here is a suggestion for an alternative form that x-phi might take, today. Whatever one thinks of the role of intuitions in analytic philosophy, intuitions also play an important role in *science* in that they are often bound up with our choice of methodological stipulations (again, see Chalmers' "pre-experimental [p. 25] assumptions"). I have been looking at consciousness studies; but many other varieties of experimental research no doubt rely on pre-experimental assumptions about how to define central theoretical terms and tests as well. Philosophers interested in both experiment and intuition might go into business, so to speak, *evaluating* these kinds of pre-experimental assumptions we see at play in science.

The pressing questions would not be how widespread this or that pre-experimental assumption shows itself to be in a survey. If I am right that these assumptions are not directly verifiable, we should be concerned with two other questions instead. First, does this pre-experimental assumption or that one support or inhibit some promising avenue of empirical inquiry? And second, if scientists were to choose some particular set of pre-experimental assumptions, what would be the wider conceptual implications? Behaviourists did not care to look very closely at the philosophical implications of their work, but perhaps that oversight eventually contributed to their eclipse.

Huxley's "On the Hypothesis that Animals are Automata" provides a remarkable example of the sort of alternative x-phi work I am proposing. Huxley saw clearly that physiologists would be led to epiphenomenalism if they responded to the Pflüger-Lewes experiments by abandoning any behavioural mark of consciousness. It took almost a century of behaviourist research in psychology for a creeping discomfort with this implied epiphenomenalism to break through and produce a new generation of researchers willing to take interactionism seriously again. But we should be heading into this new world of so-called "consciousness science," I suggest, with a clear awareness of the conceptual implications of our *new* pre-experimental assumptions. Philosophers are well

placed to help us draw out those implications, but only if we are willing to attend to the concrete scientific research the assumptions stand to support.

Acknowledgements

Thank you to Joshua Knobe, Kevin Tobia, and Hagop Sarkissian for organizing the Yale/NYU History of Experimental Philosophy Workshop, where I delivered a first draft of this paper. Thanks also to David Chalmers, who pointed me in helpful directions in subsequent correspondence, and to other participants and audience members at the Workshop. I also thank co-panellists and audiences at HOPOS in Minneapolis, the University of Sheffield, and Gresham College, London, where I delivered various versions of this paper. Finally, I revised an earlier draft of this paper in response to perceptive feedback from two referees.

Funding

I would like to thank the US-UK Fulbright Commission for financially supporting this research, and the University of Sheffield Philosophy Department for hosting me during my Fulbright year. [p. 26]

Bibliography

- Block, Ned Joel. 1995. "On a Confusion About a Function of Consciousness." *Behavioral and Brain Sciences* 18 (2):227-287.
- Boring, Edwin Garrigues. 1929/1950. *A History of Experimental Psychology*. 2nd ed. New York: Appleton-Century-Crofts.
- Brunton, T. Lauder. 1898. *Lectures on the Action of Medicines: Being the Course of Lectures on Pharmacology and Therapeutics Delivered at St. Bartholomew's Hospital During the Summer Session of 1896*. London: Macmillan.
- Cannon, Walter Bradford. 1911/1913. *A Laboratory Course in Physiology*. 2nd ed. Cambridge: Harvard University Press.
- Carpenter, William Benjamin. 1874. *Principles of Mental Physiology: With Their Applications to the Training and Discipline of the Mind, and the Study of Its Morbid Conditions*. New York: Appleton.
- Chalmers, David John. 1995. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (3):200-219.
- Chalmers, David John. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chalmers, David John. 2010. *The Character of Consciousness*. Oxford: Oxford University Press.
- Clarke, Edwin, and L. S. Jacyna. 1987. *Nineteenth-Century Origins of Neuroscientific Concepts*. Berkeley: University of California Press.
- Clifford, William Kingdon. 1874/1886. "Body and Mind." In *Lectures and Essays, by the Late William Kingdon Clifford*, edited by Leslie Stephen and Frederick Pollock, 244-273. London: Macmillan and co.
- Costall, Alan. 2006. "'Introspectionism' and the Mythical Origins of Scientific Psychology." *Consciousness and Cognition: An International Journal* 15 (4):634-654.
- Danziger, Kurt. 1980. "The Unknown Wundt: Drive, Apperception, and Volition." In *Wilhelm Wundt and the Making of a Scientific Psychology*, edited by R. W. Rieber and David K. Robinson, 95-120. New York: Plenum Press.
- Daston, Lorraine J. 1978. "British Responses to Psycho-Physiology, 1860-1900." *Isis* 69 (2):192-208.
- Duncan, George. 1869. *The Various Theories of the Relation between Mind and Brain Reviewed*. London: Trübner and Company.
- Fearing, Franklin. 1930/1964. *Reflex Action: A Study in the History of Physiological Psychology*. New York: Hafner.
- Finger, Stanley. 1993. *Origins of Neuroscience: A History of Explorations into Brain Function*. New York: Oxford University Press.
- Finger, Stanley. 2000. *Minds Behind the Brain: A History of the Pioneers and Their Discoveries*. Oxford: Oxford University Press.
- Finger, Stanley, François Boller, and Kenneth L. Tyler. 2010. *History of Neurology*. Edinburgh: Elsevier.
- Flint, Austin. 1888. *A Text-Book of Human Physiology*. Fourth ed. New York: D. Appleton and company.

- Friedman, Michael. 2001. *Dynamics of Reason*. Stanford, CA: CSLI Publications.
- Friedman, Michael. 2002. "Kant, Kuhn, and the Rationality of Science." *Philosophy of Science* 69:171-190.
- Goltz, Friedrich Leopold. 1869. *Beiträge zur Lehre von den Functionen der Nervencentren des Frosches*. Berlin: A. Hirschwald.
- Gray, Philip Howard. 1968. "Prerequisite to an Analysis of Behaviorism: The Conscious Automaton Theory from Spalding to William James." *Journal of the History of the Behavioral Sciences* 4 (4):365-376.
- [p. 27]
- Haigh, Elizabeth. 1984. "Animism, Vitalism, and the Medical University of Montpellier." *Medical History Supplement* (4):15-46.
- Haldane, J. S. 1884. "Life and Mechanism." *Mind* 9 (33):27-47.
- Hall, Marshall. 1833. "On the Reflex Function of the Medulla Oblongata and Medulla Spinalis." *Philosophical Transactions of the Royal Society of London* 123:635-665.
- Hodgson, Shadworth Hollway. 1870. *The Theory of Practice: An Ethical Enquiry*. 2 vols. London: Longmans, Green, Reader, and Dyer.
- Huxley, Thomas Henry. 1870. "Has a Frog a Soul? And of What Nature Is That Soul, Supposing It to Exist?" *Papers Read before the Metaphysical Society* (Privately Published):1-7.
- Huxley, Thomas Henry. 1872. *Lessons in Elementary Physiology*. 6th ed. London: Macmillan and Co.
- Huxley, Thomas Henry. 1874/1894. "On the Hypothesis That Animals Are Automata, and Its History." In *Collected Essays: Method and Results*, 199-250. New York: Appleton.
- James, William. 1879. "Are We Automata?" *Mind* 4 (13):1-22.
- James, William. 1890/1981. *The Principles of Psychology*. Edited by Frederick H. Burkhardt, Fredson Bowers and Ignas K. Skrupskelis. Cambridge: Harvard University Press.
- King, Lester S. 1964. "Stahl and Hoffmann: A Study in Eighteenth Century Animism." *Journal of the History of Medicine and Allied Sciences* 19 (2):118-130.
- King, Lester S. 1967. "Basic Concepts of Early 18th-Century Animism." *American Journal of Psychiatry* 124 (6):797-802.
- Klein, Alexander. 2008. "Divide Et Impera! William James's Pragmatist Tradition in the Philosophy of Science." *Philosophical Topics* 36 (1):129-166.
- Knobe, Joshua, and Shaun Nichols. 2008. *Experimental Philosophy*. Oxford: Oxford University Press.
- Knobe, Joshua, and Shaun Nichols. 2014. *Experimental Philosophy*. Vol. 2. New York: Oxford University Press.
- Koechlin, Etienne, Chrystèle Ody, and Frédérique Kouneiher. 2003. "The Architecture of Cognitive Control in the Human Prefrontal Cortex." *Science* 302 (5648):1181-1185.
- Kuhn, Thomas S. 1977. "Objectivity, Value Judgment, and Theory Choice." In *The Essential Tension: Selected Studies of Scientific Tradition and Change*, 320-339. Chicago: University of Chicago Press.

- Lewes, George Henry. 1859a. *The Physiology of Common Life*. 2 vols. Edinburgh: W. Blackwood.
- Lewes, George Henry. 1859b. "The Spinal Chord a Sensational and Volitional Centre." In *Report of the Twenty-Eight Meeting of the British Association for the Advancement of Science; Held at Leeds in September of 1858*, 135-138. London: John Murray.
- Lewes, George Henry. 1873. "Sensation in the Spinal Cord." *Nature* 9:83-84.
- Lewes, George Henry. 1877. *Problems of Life and Mind, Second Series: The Physical Basis of Mind*. London: Trübner & Co.
- Pflüger, Eduard. 1853. *Die sensorischen Functionen des Rückenmarks der Wirbelthiere: nebst einer neuen Lehre über die Leitungsgesetze der Reflexionen*. Berlin: Hirschwald.
- Robinson, William S. 2004. *Understanding Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Rocca, Julius. 2007. "William Cullen (1710-1790) and Robert Whytt (1714-1766) on the Nervous System." In *Brain, Mind and Medicine: Essays in Eighteenth-Century Neuroscience*, edited by Harry A. Whitaker, C. U. M. Smith and Stanley Finger, 85-98. New York: Springer.
- [p. 28]
- Rylance, Rick. 2000. *Victorian Psychology and British Culture: 1850-1880*. New York: Oxford University Press.
- Smith, Laurence D. 1986. *Behaviorism and Logical Positivism: A Reassessment of the Alliance*. Stanford: Stanford University Press.
- Steinke, Hubert. 2005. *Irritating Experiments: Haller's Concept and the European Controversy on Irritability and Sensibility, 1750-90*. Amsterdam: Rodopi.
- Stokel-Walker, Chris. 2015. "The Chicken That Lived for 18 Months without a Head." *BBC News Magazine*. Accessed February 9, 2016. doi:<http://www.bbc.com/news/magazine-34198390>.
- Sytsma, Justin. 2014a. *Advances in Experimental Philosophy of Mind*. New York: Bloomsbury Academic.
- Sytsma, Justin. 2014b. "Introduction." In *Advances in Experimental Philosophy of Mind*, edited by Justin Sytsma, 1-10. New York: Bloomsbury Academic.
- Sytsma, Justin, and Edouard Machery. 2010. "Two Conceptions of Subjective Experience." *Philosophical Studies* 151 (2):299-327.
- Sytsma, Justin, and Edouard Machery. 2012. "On the Relevance of Folk Intuitions: A Commentary on Talbot." *Consciousness and Cognition* 21 (2):654-60.
- Talbot, Brian. 2012a. "The Irrelevance of Dispositions and Difficulty to Intuitions About the "Hard Problem" of Consciousness: A Response to Sytsma, Machery, and Huebner." *Consciousness and Cognition* 21 (2):661-666.
- Talbot, Brian. 2012b. "The Irrelevance of Folk Intuitions to the "Hard Problem" of Consciousness." *Consciousness and Cognition* 21 (2):644-650.
- Verworn, Max. 1912. *Physiologisches Praktikum für Mediziner*. 2nd ed. Jena: G. Fischer.
- Waters, A. T. H. 1859. "Address in Physiology." *The British Medical Journal* 2 (139):685-690.
- Whytt, Robert. 1751. *An Essay on the Vital and Other Involuntary Motions of Animals*. Edinburgh: Hamilton, Balfour, and Neill.

