

ARTICLE

# Abduction and Composition

Kenneth Aizawa<sup>1\*</sup>  and Drew B. Headley<sup>2</sup>

<sup>1</sup>Department of Philosophy, Rutgers University, Newark, NJ, US and <sup>2</sup>Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ, US

\*Corresponding author: Email: [ken.aizawa@gmail.com](mailto:ken.aizawa@gmail.com)

(Received 07 March 2020; revised 07 October 2020; accepted 22 January 2021; first published online 10 February 2022)

## Abstract

Some New Mechanists have proposed that claims of compositional relations are justified by combining the results of top-down and bottom-up interlevel interventions. But what do scientists do when they can perform, say, a cellular intervention, but not a subcellular detection? In such cases, paired interlevel interventions are unavailable. We propose that scientists use abduction and we illustrate its use through a case study of the ionic theory of resting and action potentials.

Recent work by some New Mechanists has proposed that scientists test hypotheses about what composes what by pairing the results of “top-down” and “bottom-up” “interlevel experiments” (e.g., Craver 2007; Craver and Darden 2013)<sup>1</sup> In a “top-down interlevel experiment,” one manipulates a whole, then detects some feature of one or more of its constituent parts. As an example, one might have a subject view an image, then try to detect changes in blood oxygenation levels in particular regions of the brain. In a “bottom-up interlevel experiment,” one manipulates a part, then detects some feature of the whole. As an example, one might perform a brain lesion on an animal, then detect what changes there might be in the animal’s performance on some task.

Notice that such interlevel experiments require two technical capabilities. There must be an ability to manipulate some experimental preparation and an ability to detect something at a distinct level, either higher or lower, in that preparation. What, however, do scientists do when they lack one of these capabilities? Consider a particularly salient kind of case. Biologists might develop a technique for manipulating cells, but then lack the ability to detect subcellular responses. What do they do when they lack top-down interlevel experimental results to pair with bottom-up interlevel experimental results?<sup>2</sup>

---

<sup>1</sup> This account obviously presumes that there is a legitimate notion of levels in the special sciences, so this presumption will be honored, if only for the sake of argument, in this paper.

<sup>2</sup> Glennan (2005, 450) recognizes this methodological challenge but does not articulate the solution to it that we present. One reviewer proposed that Gebharter (2017) and Harbecke (2010, 2015) provide confirmation methods that differ from interlevel experiments. Whether or not that is true, their methods

An apt instance of this type of situation is found in the history of accounts of the action potential. In broad outline, over the course of the first half of the 20th century, physiologists developed techniques that enabled them to make precise measurements of neuronal action potentials, but they did not have techniques to detect the contemporaneous lower-level features of the ions or their movements in those neurons. Thus, top-down interlevel experiments were not available to them. How did the physiologists of that day grapple with this methodological challenge?<sup>3</sup>

Looking closely at the relevant primary literature, one finds that scientists made ample use of abductive reasoning. By abductive reasoning, we mean inferring that the world is arranged thus and so because its being so arranged has some explanatory value. So, as an illustration, consider some marks in the snow. One postulates that a deer walked through the area, since that would explain the marks in the snow. Implicit in this explanation is the presupposition that the deer walking through the snow caused the marks. Similarly, Alan Hodgkin and Andrew Huxley (1952), for example, postulated lower-level fluxes of sodium and potassium ions because such fluxes would explain the higher-level neuronal action potentials. In this case, the implicit assumption of the explanation is that the ion fluxes compose (or, as the New Mechanists often put it, “mechanistically constitute”) the action potential.<sup>4</sup> What the case suggests is that New Mechanists should accept that there is a broader range of methods by which scientists justify compositional hypotheses.<sup>5</sup>

To make the foregoing case, section 1 will provide some theoretical background regarding what we mean by “abduction” and how it contributes to the justification of hypotheses about what is compositionally related to what in ways that interlevel experiments cannot. Sections 2 and 3 will provide the historical case study. Section 2 will briefly review the scientific understanding of the compositional basis of the resting and action potentials prior to 1952, whereas section 3 will detail the experiments and reasoning from the first experimental paper in Hodgkin and Huxley’s seminal series of reports from 1952. In brief outline, we find that pre-1952 work had established qualitatively the particular roles of sodium and potassium ions in the

---

apparently share the limitations of interlevel experiments in requiring data on both the upper and the lower level in a single preparation, hence will suffer the same limitation just described for interlevel experiments. In personal correspondence from May 2020, both Gebharter and Harbecke read the foregoing portion of the footnote and confirmed the analysis.

<sup>3</sup> An important secondary feature of the example is that some New Mechanists have written about it in detail (e.g., Craver 2007, 2008; Levy 2014). So, the example falls neatly within what at least some New Mechanists take to be the purview of the New Mechanist approach.

<sup>4</sup> For a theory of composition, as opposed to causation, see Gillett (2016, chapter 2). For present purposes, we need not suppose that all abductive inferences involve postulating causes or “composers”; it is merely that the relevant species of abductive inference are apparently like this. Perhaps a noncausal, non-“composer” case is the following. Suppose one has a law that whatever contains sugar is sweet and a law that all pineapples are sweet. One might explain why all pineapples are sweet by the hypothesis that all pineapples contain sugar. The law that all pineapples are sweet might be taken not to be a cause or a “composer.” See Schurz (2008, 212), from which the example, though perhaps not the moral, is drawn.

<sup>5</sup> Baumgartner and Casini (2017) propose that scientists use a kind of abductive reasoning to explain an unbreakable connection between certain higher-level processes and their lower-level bases. In the version of abduction we describe, what is abductively explained is simply the higher-level process, such as the firing of an action potential. Helpfully for us, Baumgartner and Casini close their paper by drawing attention to the need for case studies of how scientists confirm compositional hypotheses.

compositional basis of the resting and action potentials, whereas Hodgkin and Huxley's papers of 1952 provided measurements of membrane currents that were consistent with those ions fluxing during the action potential, both confirming and refining the qualitative pre-1952 account.

### I. Philosophical preliminaries on abduction and interlevel experiments

We have proposed that early 20th century physiologists used abductive inferences to justify hypotheses about what composes what.<sup>6</sup> Given such an observation, one philosophical project would be to develop a theory of what abductive inferences are. Thagard (1978), for example, reviews a number of episodes in the history of science that *prima facie* illustrate abduction as a prelude to the development of a theory of abduction.<sup>7</sup> Our project here is not meant to rival Thagard's. Nor will our project weigh in on any number of issues concerning how to understand abduction, such as the extent to which abduction is a matter of inference to the best explanation, what features of explanations might make one explanation better than another, and which possible explanations are up for comparison. Instead, our project is meant to provide a (detailed) example of abductive inferences used to justify claims about compositional bases, thereby establishing that there is a scientific practice for which one should provide a theory.<sup>8</sup> Having set out this caveat—that we will not attempt a full-blown theory of abduction—we can nevertheless offer some clarificatory remarks on what we mean by “abduction” and one context in which it is likely to have a prominent role to play in scientific reasoning.

It is common to distinguish deductive, inductive, and abductive inferences. Deductive inferences are those in which, if the premises are true, then the conclusion must be true, whereas with inductive and abductive inferences, the premises may be true, while the conclusion is nevertheless false. Inductive and abductive inferences are often said to be ampliative in the sense that they go beyond what is asserted in the premises, whereas deductive inferences are nonampliative. While “inductive inference” is sometimes used as a blanket term that includes abductive inference, it is also sometimes used more narrowly to cover instances in which, say, a property or regularity of some observed individuals is applied to some unobserved individuals. Abductive inferences are also intended to warrant claims about the unobserved; however, they do this in a more general way by providing an explanation. In an abductive

---

<sup>6</sup> In observing that scientists sometimes use abductive methods to justify hypotheses, we do not mean to endorse this use as genuinely justificatory. As far as our present concerns go, it could turn out that scientists take abductive inferences to be justificatory, but that they are not.

<sup>7</sup> Both Thagard (1978) and Lipton (1991, 2004) early on provide putative scientific examples of inference to the best explanation as showing the need for a theory of this sort of reasoning. This is relevant to “abduction” insofar as this term is often used interchangeably with “inference to the best explanation.” See Schurz (2008) and Douven (2017) for endorsement of this usage; see Campos (2011) for reasons against this usage.

<sup>8</sup> We take it that it is perfectly sensible to document the existence of a practice before developing a theory of that practice. Indeed, this seems to be exactly what Craver did with interlevel experiments and the formalism of mutual manipulability. Craver (2002) provides a taxonomy of interlevel experimental strategies, namely activation strategies, interference strategies, and additive strategies. This was largely a description of scientific practices without a theory of them. It did not include the mutual manipulability theory of those practices that appeared in Craver (2007).

inference, some entities are postulated because the existence of those entities would provide an explanation of something. In other words, the fact that some entities explain something provides some justification for postulating, accepting, or believing in those entities (e.g., Schurz 2008, 202; Douven 2017, 5).<sup>9</sup>

There are two features of abductive inference that are noteworthy with respect to theories of what composes what. Abductive inference can involve the postulation of entities that are 1) qualitatively distinct from what is cited in the observations that form the premises of the abductive inference and 2) not directly empirically detected. Consider these points in turn. In an inductive inference, one might have as evidence that copper sample 1 has a melting point of 1085°C, copper sample 2 has a melting point of 1085°C, and copper sample 3 has a melting point of 1085°C, from which one might infer that all copper has a melting point of 1085°C. In this case, the copper and the melting point of the conclusion are not qualitatively distinct from the copper and the melting point of the observational premises. By contrast, with abduction, when one observes certain marks in the snow then infers that a deer must have walked through that snow, the inference postulates an entity, a deer, that is qualitatively distinct from the marks in the snow. When one measures the peak voltage of an action potential, then infers that certain ion fluxes must have given rise to it, the inference postulates individuals and activities that are qualitatively distinct from the action potential. In the abductive cases, the inferences are ampliative as they invoke qualitatively distinct entities.<sup>10</sup>

While there is certainly room to question what is meant by “qualitative distinctness,” the idea seems to be implicit throughout much of the New Mechanism.<sup>11</sup> In this literature, one often reads that an individual  $S$ ’s engaging in a process of  $\psi$ -ing is to be explained in terms of a collection of individual  $x_i$ ’s engaging in processes of  $\phi_i$ -ing. The supposition, suggested by the notation, is that  $S$  (e.g., a neuron) is qualitatively distinct from the  $x_i$ ’s (e.g., ions) and that the process  $\psi$  (e.g., firing an action potential) is qualitatively distinct from the processes  $\phi_i$  (e.g., crossing the membrane). Moreover, such descriptions appear to be implicitly motivated by the idea that mechanistic explanations are supposed to be better than, and distinct from, “homuncular” explanations.<sup>12</sup> In a “homuncular” explanation, one might explain  $S$ ’s  $\psi$ -ing by appeal to some component  $x_i$   $\psi$ -ing, as in explaining that a normal human being  $S$  sees because there is an inner homunculus  $x_i$  that sees. That such an explanation is, in some sense, wanting motivates the idea that  $\psi$ -ing should be distinct from  $\phi$ -ing.

<sup>9</sup> It has been maintained the abductive inferences serve only as guides for future research. One reason to doubt this view is that the scientific community appears to have originally accepted the relationship between action potentials and ion fluxes based only upon abductive inferences. This provides at least prima facie reason to believe that the scientific community at least treats abductive inference as justificatory.

<sup>10</sup> Schurz (2008, 216) refers to “micro-part abductions,” in which “one abduces a hypothesis about the microscopic composition of observable objects in terms of micro-parts which obey the same laws as the observable macroscopic objects, in order to explain various observed empirical phenomena.” Our pre-supposition here, apparently shared by many New Mechanists, is that there is an abundance of what might be aptly described as “micro-part” abductions that do *not* fit the model Schurz describes.

<sup>11</sup> For some non-New Mechanist discussion of “qualitative distinctness,” see Gillett (2010).

<sup>12</sup> See Craver and Darden (2013, 88–9) for comments on this score.

A second important feature of abduction is that one postulates entities, not on the basis of any direct measurement or detection of them, but because they are explanatory. In other words, through abduction, a compositional hypothesis regarding some entity is supposed to be warranted, even though one has not observed or detected that entity. In the case of the marks in the snow, the deer is unobserved, since it was present in the past. In the case of the action potential, the total ionic current could be measured, but the fluxes of individual ions were not observed, since early 20th century physiologists did not have the technology to isolate them.

This second feature of abduction serves to distinguish it from so-called “interlevel experiments” in the way mentioned in the introduction. A “top-down” interlevel experiment requires the ability to manipulate entities, then detect, simultaneously and directly, something at a lower level in that very preparation.<sup>13</sup> But in some cases, scientific technology may be limited so that one cannot directly detect at the lower level. In such cases, top-down interlevel experiments will be technologically impossible, but one might be able to justify hypotheses through abductive inferences. Looking ahead to subsequent sections of the paper, mid-20th century physiologists had technology, such as the voltage clamp, that enabled them to make precise measurements of current flowing across the neuronal membrane when it was held at a particular voltage, but did not have technology to enable them to directly and simultaneously measure the action potential and its contemporaneous ion fluxes. Instead, the voltage and time dependence of the membrane conductance were measured in different ionic solutions, and from that separate ion fluxes were postulated that could explain the measurement of the action potentials.

The foregoing observations provide what seem to us to be basic features of at least some important cases of abduction that will help the reader begin to appreciate what we mean by “abduction” and why scientists would sometimes use abduction to justify compositional claims. There are, of course, many philosophical issues we have not touched on, but we shall return to some of them in the concluding remarks.

## 2. The pre-1952 understanding of the resting and action potentials

At the time Hodgkin and Huxley published their series of papers in 1952, it was already established that the resting and action potentials associated with nerves were related to the concentration of specific ions inside and outside the membrane. In 1902, Julius Bernstein put forward the theory that neuronal membranes exhibited a resting potential due to a selective permeability to potassium. This “membrane theory” was

---

<sup>13</sup> In principle, one might claim that there are “interlevel experiments” wherein scientists in one laboratory perform an experiment on one level, whereas scientists in another laboratory perform an experiment on the lower level. Thereafter, the results are somehow combined. This, in fact, happens in many cases, but it is implicitly not what is involved with interlevel experiments as they are commonly understood in the New Mechanist literature. Notice that, as the New Mechanists typically understand them, the manipulations and the detections are supposed to occur in one subject or preparation. The method does not propose having a subject view an image, then measure changes in the brain of another subject. The method does not propose giving one animal a brain lesion, then measuring the behavior of another animal. Moreover, the idea that the higher- and lower-level interventions are supposed to be in one experimental preparation is implicit in much of the discussion of mutual manipulability, insofar as this is the source of the inability to perform ideal interlevel interventions. See Baumgartner and Gebharder (2016) for discussion of this problem.

predicated on earlier work by Walther Nernst (1889) showing that electrical potentials could be established by the concentration gradients of ions, and the theory of Wilhelm Ostwald (1890) that the voltage across a membrane arises from its selective permeability to particular ions. The Nernst equation describes how the potential across a membrane is related to the relative concentration of a particular ion, denoted *ion*, inside and outside the membrane.

$$E_{\text{ion}} = \frac{RT}{zF} \ln \frac{[\text{ion}]_i}{[\text{ion}]_o} \quad (1)$$

In the equation, *E* is the potential across the membrane for a particular ionic species (e.g.,  $K^+$  or  $Na^+$ ), *R* is the gas constant, *T* is temperature in Kelvin, *z* is the charge of the ion, *F* is Faraday's constant, and the square brackets signify the ionic concentration with an *o* subscript for the outside (extracellular) concentration and an *i* subscript for the inside (intracellular) concentration.<sup>14</sup> If a membrane is at the equilibrium potential, then the concentration- and voltage-dependent ionic fluxes are balanced across the membrane; that is, they cancel each other out and so there is no net ionic current.

Bernstein's experimental preparation consisted of a frog nerve suspended in an oil-filled jar. One end of the nerve was cut, saline-soaked clay electrodes were positioned at the cut and uncut ends of the nerve, and the voltage difference between them was measured (known as an injury potential). By recording from the cut end of the nerve, Bernstein gained crude electrical access to the intracellular solution of the axons that were bundled into the nerve. The difference in electrical potential between this cut portion and the intact segment of nerve allowed him to measure the average voltage difference across the axonal membranes that comprised the nerve. Bernstein noted that this potential was around 50 mV, which is approximately what one would expect from the Nernst equation if the membrane was permeable to potassium ions, and given the already known concentrations of potassium inside and outside nervous tissue (Bernstein 1971, 1912). Crucially, this experiment does not constitute an interlevel experiment. Only the potential of the nerve was measured, and the manipulation of cutting the nerve was performed merely to give access to the intracellular milieu. Later in the experiment, Bernstein had suspended his oil-filled jar in a water bath and was able to change its temperature from 36° to -2° Celsius. He found that the potential changed its value linearly with temperature, in accordance with the Nernst equation. This too does not constitute an interlevel experiment since both the manipulation, a change in temperature, and the measurement, the injury potential, were performed at the level of the whole nerve. Since the Nernst potential for potassium best fit his results, he abductively concluded that the resting potential of the neuron arose from the charges on the potassium ions, their relative concentrations across the membrane, and the membrane's partial permeability to them.

Curtis and Cole (1942) further confirmed the fact that the resting potential matched that predicted by the Nernst equation for potassium by varying its concentration outside the nerve. Using the giant axon of the squid, they were able to record

<sup>14</sup> It is worth noting that modern forms of the Nernst equation have the outside concentration as the numerator and inside concentration as the denominator; however, to maintain fidelity with the reviewed literature, we will adopt the older form.

the resting potential from inside the axon by inserting a fine glass micropipette down its length. Since the recording electrode was placed inside the axon, they were able to bathe the nerve in sea water, which would have shorted out the externally placed electrodes used by Bernstein (hence why he immersed his nerves in oil). This allowed them to vary the concentration of potassium outside the axon and observe its effect on the resting potential. As with Bernstein's experiments, Curtis and Cole performed a manipulation at the level of the entire preparation by varying the concentration of potassium outside the axon. They found that increasing the external concentration of potassium, bringing it closer to the concentration found inside the axon, drove the resting potential toward zero, agreeing with the Nernst equation.<sup>15</sup> Again, Curtis and Cole used an abductive inference, rather than an interlevel experiment to reach their conclusion; the Nernst equation for potassium provided the best agreement with the data.

Altogether this body of work suggested that the resting potential arose from the membrane's slight permeability to potassium ions, the charge on the potassium ions, and the relative intracellular and extracellular concentrations of potassium ions. Subsequently, the Nernst equation was modified to incorporate the small permeability of the membrane to sodium and chloride ions ( $P_{Na}$ ,  $P_{Cl}$ ), to allow for a more accurate estimate of the resting membrane potential (Hodgkin and Katz 1949). Known as the Goldman, Hodgkin, and Katz equation,

$$E_r = \frac{RT}{F} \ln \frac{(P_{Na}[Na]_i + P_K[K]_i + P_{Cl}[Cl]_o)}{(P_{Na}[Na]_o + P_K[K]_o + P_{Cl}[Cl]_i)} \quad (2)$$

it gives the resting membrane potential as a function of the external and internal concentrations of sodium, potassium, and chloride ions, along with the membrane's permeability to each of these ionic species ( $P_{Na}$ ,  $P_K$ ,  $P_{Cl}$ ). In their proposal for this explanation of the resting potential, Hodgkin and Katz (1949) are agnostic as to the exact basis for the membrane permeability, but they do adopt a formalism from diffusion that is grounded in the mobility of each ionic species in the membrane.

In the same 1949 paper, Hodgkin and Katz implicate sodium as the principal ion responsible for the action potential using an approach similar to Curtis and Cole (1942). Originally, as part of his membrane theory, Bernstein had suggested that the action potential arose from a "breakdown" in the membrane to potassium, which would reduce toward zero the potential difference between the inside and outside of the axon. This was based on measurements he had made using extracellular electrodes, in which he found that the potential was near 0 mV during action potential generation.

With the advent of improved recording techniques, in particular the ability to insert electrodes inside the cell (Hodgkin and Huxley 1939), it was found that the membrane potential did not go to 0 mV. Instead, it overshoot, typically reaching around -50 mV. This contradicted the theory of the action potential put forward by Bernstein. Instead, Hodgkin and Katz (in collaboration with Huxley) put forward the theory that the action potential reflected an increase in membrane permeability to sodium (Hodgkin and Katz 1949). This was based in part on early experiments by

<sup>15</sup> While Curtis and Cole make no mention of Nernst in their 1942 publication, Cole reproduces the figures from that paper in his 1968 book and ties them directly to the Nernst equation (Cole 1968, 148).

Overton (1902) showing that reducing the sodium concentration in the solution bathing a nerve decreased its excitability. However, an argument based on the properties of the ionic solutions is the reason stated by Hodgkin and Katz (1949):

The hypothesis is based upon a comparison of the ionic composition of the axoplasm of a squid nerve with that of the sea water in which experimental preparations are normally immersed. . . . the sodium and chloride ions may be present in [intracellular] concentrations which are less than one tenth of those in sea water. . . . The reversed potential difference which could be obtained by a mechanism of this kind might be as great as 60 mV. (Hodgkin and Katz 1949, 37–38)

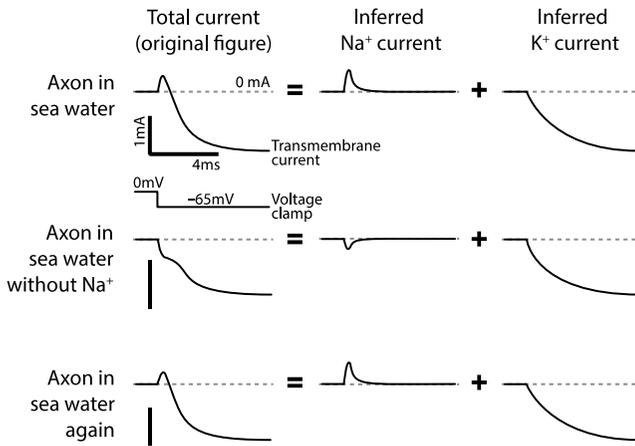
Thus, measurements of the sodium ion concentration inside the axon found that sodium was less than 10% that found outside. Following the Nernst equation, this would give a potential of  $\sim 56$  mV at  $10^\circ$  Celsius, which was close to the potential reached during the peak of the action potential. This would produce a robust positive current flowing into the axon because the sodium ions are positively charged and would move down their concentration gradient, from the high extracellular concentration to the low intracellular concentration. Since the Nernst equation indicates that the potential for sodium depends in part upon its concentration outside the axon, and that was a relatively simple factor to manipulate, Hodgkin and Katz systematically varied it while recording the action potential's peak voltage. They replaced sodium in the solution bathing the nerve with isotonic dextrose, which prevents the nerve from being damaged by an osmotic imbalance. Reducing sodium toward the concentration found inside the axon brought the action potential amplitude toward 0 mV, which would be predicted by the Nernst equation since the log of 1 ( $[\text{Na}]_i = [\text{Na}]_o$ ) is 0. They found that for a range of concentrations the peak of the action potential matched the sodium potential predicted by the Nernst equation.

### 3. Abduction in Hodgkin and Huxley's treatment of sodium and potassium currents in nerve membrane

Our review of Hodgkin and Huxley's (1952) treatment of sodium and potassium currents is meant to draw further attention to the role of abduction in the development of the compositional accounts of the resting and action potentials. Stated succinctly, we make the case that Hodgkin and Huxley used abduction as a means of guiding their experiments and framing the quantitative measures they used in their compositional explanation of the action potential. For these purposes, we need not review the entire series of Hodgkin and Huxley's papers from 1952. Instead, it will suffice to draw attention to the arguments in the first part of their first paper of 1952.<sup>16</sup>

Their first experiment depolarized an axon by 65 mV and measured the change in current across the membrane, first, in sodium-containing sea water, then in a sodium-free "choline" sea water solution, then again in a sodium-containing sea water

<sup>16</sup> This is the first paper in their series that is solely authored by Hodgkin and Huxley and is focused on collecting experimental data to later ground their model of the action potential. The actual first paper in the series was coauthored with Bernard Katz and is a description of their experimental apparatus, the voltage clamp.



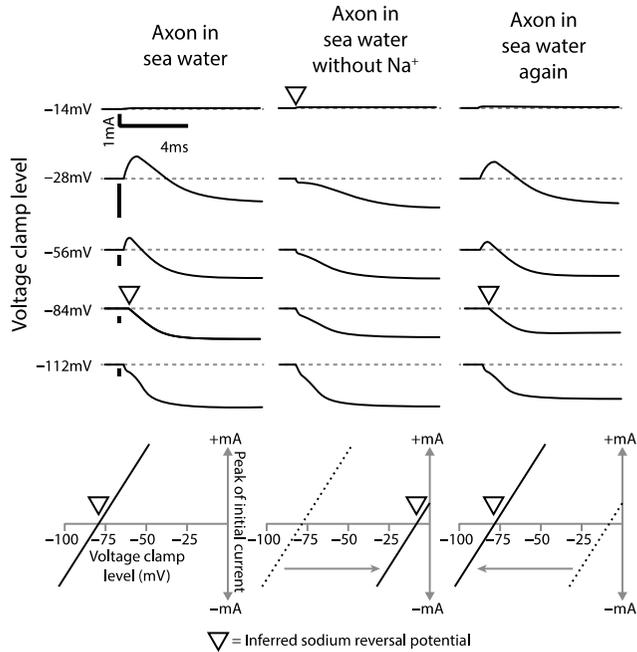
**Figure 1.** Hodgkin and Huxley's measured total current and inferred currents in sodium and sodium-free solutions. (Based on Hodgkin et al. 1952, p. 451, Fig. 1).

solution (see Figure 1). The most obvious feature of these panels is that the initial inward current (represented by the initial rise at the left of the curve) disappears when the axon is placed in the sodium-free solution. From earlier research, Hodgkin and Huxley (1952) had concluded that the initial inward current was due to sodium.<sup>17</sup> So, while they do not write as much at this point in their paper, we take them to believe that what would explain this feature is sodium ions being responsible for the transient inward current seen in the top and bottom panels.

Their second experiment is a more complex variation on the first, wherein the same sodium-containing, sodium-free, sodium-containing protocol is used, but with variations in the strength of the depolarization (see Figure 2). The sodium component of the membrane current will be proportional to the driving force for sodium ions across the membrane, which is the difference between the membrane potential (controlled by Hodgkin and Huxley [1952]) and the sodium equilibrium potential. By bringing the membrane potential closer to the sodium equilibrium potential they should offset the inward driving force of the sodium ion concentration gradient, thereby reducing the transient inward flux of sodium ions even to the point of completely eliminating the influx in favor of an efflux. This prediction is borne out in the first and third columns of Figure 2 at all depolarizations greater than -15 mV. Again, the implicit assumption is that the pattern in the data might be explained by the hypothesis that the currents are due to the influx of sodium ions driven by the sodium ion concentration gradient. In addition, of the second column of Figure 2, Hodgkin and Huxley write:

When the axon is placed in a sodium-free medium, such as the “choline sea water”, there can be no inward flux of sodium, and the sodium current must always be outward. This will account for the early hump on the outward current

<sup>17</sup> For example, “there are experiments which show that the rate of rise and amplitude of the action potential are determined by the concentration of sodium in the external medium (e.g., Hodgkin and Katz, 1949; Huxley and Stämpfli, 1951).” (Hodgkin, Huxley, and Katz 1952, 424).



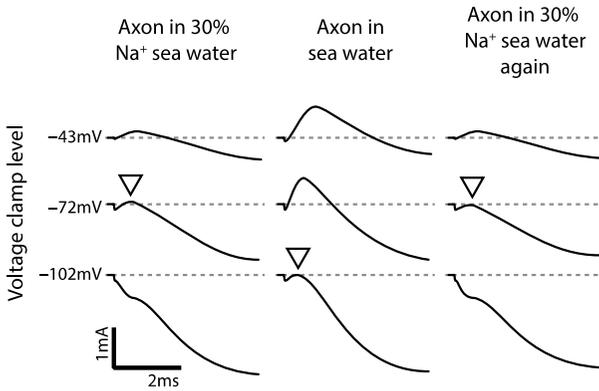
**Figure 2.** Hodgkin and Huxley's measured current at multiple voltage clamp levels along with inferred sodium reversal potential (Based on Hodgkin and Huxley 1952, p. 451, Fig. 2).

which is seen at all but the lowest strength of depolarization in the center column of Fig. 2. (Hodgkin and Huxley 1952, 454)

If we take “account for” to be synonymous with “explain,” then there is reason to think that Hodgkin and Huxley understand themselves to be doing something like abductively inferring that sodium ion fluxes are responsible for the observed transient inward currents. Note that they did not experimentally inject sodium into the axon, nor did they verify that the ions mediating the current were in fact sodium. Instead, sodium ion fluxes explained the constellation of facts already known about the molecular constitution of axoplasm, the Nernst equation, and their experimental circumstances.

Finally, their third experiment varies both a) the concentration of the sodium in the sea water to 30% and 10% of normal and b) the strength of the depolarization. Figure 3 shows the results of the protocol placing the axon in 30% sodium sea water, then sea water, then 30% sodium sea water. Using the Nernst equation applied to the sodium ion, Hodgkin and Huxley were able to predict the depolarization at which there would be neither influx nor efflux of current and found that the predicted values were close to those measured.

To this point, we have tried to emphasize Hodgkin and Huxley's use of abduction, but properties come to the fore in what Hodgkin and Huxley themselves describe as a “quantitative test” of their hypothesis regarding the role of sodium ions in the rising phase of the action potential (the initial hump in their current traces, such as the middle panel of Figure 3). So, they let  $E_{Na}$  denote the sodium potential in sea water and  $E'_{Na}$  denote the sodium potential in a reduced sodium solution. Using the Nernst



**Figure 3.** Hodgkin and Huxley's measured currents over multiple voltage clamp levels and sodium sea water concentrations.

equation, one can predict the difference between these potentials,

$$\begin{aligned}
 E'_{Na} - E_{Na} &= \frac{RT}{F} \ln \frac{[Na]_i}{[Na]'_o} - \frac{RT}{F} \ln \frac{[Na]_i}{[Na]_o} \\
 &= \frac{RT}{F} \left\{ \ln \frac{[Na]_i}{[Na]'_o} - \ln \frac{[Na]_i}{[Na]_o} \right\} \\
 &= \frac{RT}{F} \ln \frac{[Na]_o}{[Na]'_o}
 \end{aligned}
 \tag{3}$$

Since they did not have experimental access to  $E_{Na}$  (they could only estimate it from the Nernst equation), it had to be calculated from the axon's resting potential plus the amount of additional depolarization required to reverse the direction of the initial hump of (putative) sodium current. This sum reflects the absolute membrane potential at which the sodium current would reverse,  $E_{Na}$ . The difference between it in the normal and reduced sodium sea water could then be compared with that predicted by the modified form of the Nernst equation immediately above. As shown in the reproduced table, the Observed and Theoretical values for this difference in potential were remarkably consistent. We cannot assume that all numerical values are values for properties; they might instead be rates of processes, for example. Caution noted, the values of  $R$ ,  $T$ ,  $G$ ,  $[Na]_o$ ,  $[Na]'_o$ , are properties and they play an essential role in Hodgkin and Huxley's argument. Indeed, a general moral would seem to be that, when scientists rely on quantitative arguments, one should consider whether or not they are trafficking in properties.

Axon no.	Temp. (°C)	$\frac{[Na]'_o}{[Na]_o}$	$V_{Na}$ (mV)	$V'_{Na}$ (mV)	( $E_r - E_r$ ) (mV)	Observed (mV)	Theoretical (mV)
20	6.3	0.3	-105	-78	+3	+30	+28.9
20	6.3	0.1	-96	-45	+4	+55	+55.3
21	8.5	0.1	-100	-48	+4	+56	+55.6
21	8.5	0.1	-95	-45	+4	+54	+55.6

Having made their abductive case for the sodium ions being responsible for the earliest phase of the current—the transient inward current—Hodgkin and Huxley turn their attention to the falling inward current (represented by the portion of the curve following the peak on the left of Figure 1) and the persistent outward current (represented by the rightmost portion of the curve). To begin to support their theory of this outward current they write, “An outward current which arises with a delay after a fall in the membrane potential is clearly what is required in order to explain the falling phase of the action potential” (Hodgkin and Huxley 1952, 456). Again, this passage gives us reason to believe that Hodgkin and Huxley take themselves to be doing what we claim they are doing, namely, using abduction to determine what underlying processes are responsible for the falling phase of the current. Unlike the sodium case, they do not perform experiments to determine whether potassium is the mediator of the late outward current. Instead, they use a piecemeal collation of results to abductively make that conclusion. In a previous review (Hodgkin 1951), Hodgkin had made this case by drawing upon radiotracer studies, changes in membrane conductance, the effects of nerve stimulation on potassium transport across the membrane, and the sensitivity of the late phase of the action potential to extracellular potassium concentration.

Having postulated that the late falling phase of the current is mediated by potassium, Hodgkin and Huxley next consider two facts about the potassium current,  $I_K$ , as found in sea water versus as found in the sodium-free “choline” sea water. The first fact is that the maximum outward current in the sodium-free solution is 10–20% smaller than in sea water and the second is that this maximum is achieved in a shorter period of time. They offer three factors that they assume explain these facts, beginning with this:

Part of the difference in amplitude is explained by the difference of resting potential. Since the resting potential is greater in the low-sodium medium, a higher strength of depolarization is needed to reach a given membrane potential during the voltage clamp. This difference can be allowed for by interpolation between the actual strengths employed in one of the solutions. In most cases, this procedure did not entirely remove the difference between the amplitudes. (Hodgkin and Huxley 1952, 456)

So, in the first three sentences here, they explicitly appeal to an hypothesis about the concentrations of intracellular and extracellular sodium to explain the two facts. This provides further confirmation for the role of sodium in the resting potential. Moreover, they note that the whole of the 10–20% difference in the maximum outward current cannot be explained by the difference in resting potential, at which point they move to two additional hypotheses. There is an effect of not using “compensated feed-back” and a “polarization.” For present purposes, we need not go into the details of what these effects are. Instead, we want to draw attention to their comment that, “We do not know enough about either of these effects to estimate the amount by which they may have reduced the potassium current. It does seem at least possible that they account for the whole of the discrepancy, and we therefore assume provisionally that substituting choline sea water for sea water has no direct effect on the potassium current” (Hodgkin and Huxley 1952, 456–57). Hodgkin and Huxley’s handling of these facts is by no means among their most important scientific accomplishments, but

they serve our purpose of highlighting the role of abduction. They do not perform inter-level experiments; instead, they attempt to understand what is going on in the nerve by what they think would explain the measured features of the neurons.

#### 4. Conclusion

We have proposed that scientists sometimes use abduction to justify claims about what things are compositionally related. In addition, we provided a type of scenario in which we think the scientific use of abduction makes sense. Scientists use abduction to determine what things are compositionally related when they lack the technical means to directly detect features at lower levels. We tried to illustrate this practice in action through a case study, namely, the experimental work that led to the establishment of ion fluxes as the compositional bases of the action potential.

What, one might ask, is the larger significance of this? After all, the theoretical analysis is modest. It focuses on just two features of abduction that make it apt for reasoning about compositional hypotheses. Moreover, the conclusion that scientists sometimes use abduction to justify claims about what things are compositionally related is again modest. The rationale for this modesty is that even these modest analyses should suffice to suggest further philosophical inquiry. It should suffice to raise new questions or provoke further reflection on existing answers. To put the matter in another way, while our theoretical analysis and philosophical conclusion are modest, our expectation is that they will be fruitful. Consider this in a bit more detail.

Obviously, neither abduction nor compositional explanation is an entirely new topic in the philosophy of science. They have, however, largely been explored independently. “Compositional abduction,” one might say, is a largely unexamined topic. A common metaphor for this is that we have “siloed” research programs. Research on abduction and research on compositional explanation are conducted in relative isolation. In the abstract, compositional abduction may not seem like a compelling direction for philosophical research, so we should consider some new questions that should immediately come to mind in light of the modest analyses that have been presented.

Consider, first, what additional attention to abduction might mean for the justification of compositional hypotheses. One feature of abduction is that it does not require an intervention at one level, then a detection at a lower level. Abduction is, thus, applicable where interlevel experiments are unavailable. Craver says that interlevel experiments are “indispensable” (Craver 2007, 147), whereas Craver and Darden claim that “interlevel experiments play a crucial role in integrating findings at multiple levels into a single mechanism schema in which more and more gray boxes can be turned to glass boxes.” (Craver and Darden 2013, 125). But what do they mean by “indispensable” and “crucial”? And are these claims true? Were Hodgkin and Huxley awarded the Nobel Prize for their theory of the action potential, even though they were unable to perform interlevel experiments? If so, what impact does this have on our understanding of the importance of interlevel experiments? How does the recognition of a role for abduction influence our assessment?

Consider a second related line of inquiry. The Hodgkin–Huxley (1952) paper shows that sometimes scientists rely on abduction, rather than interlevel experiments, to confirm compositional hypotheses. How often, then, do scientists rely on abduction, rather than interlevel experiments? The question might be made more pressing by

noting another limitation on interlevel experiments as usually understood. It is commonly assumed that there must be a) an intervention on the activity of an individual (an  $S \Psi$ -ing) and the detection of something regarding the activity of a part ( $x_i \Phi_i$ -ing) and b) an intervention on the activity of that kind of part (i.e., on  $x_i \Phi_i$ -ing) and a detection of something regarding the activity of the first individual ( $S \Psi$ -ing). In other words, the  $S \Psi$ -ing and the  $x_i \Phi_i$ -ing must be the same in two experiments. Are there scientific episodes in which scientists cannot carry out such “perfectly paired” interlevel experiments? If so, what do scientists do? Do they somehow use abduction? Do they do something else? Or, maybe the limitation on interlevel experiments as usually understood needs to be revised.

Consider a third line of inquiry. There is a scientific practice of using interlevel experiments as outlined by Craver (2002). Beginning with Craver (2007), however, it has been common to attempt to provide a theoretical account of this practice in terms of one or another set of conditions of “mutual manipulability.” It has been argued, however, that this sort of analysis invokes conditions that cannot be satisfied. (e.g., Romero 2015; Baumgartner and Gebharter 2016; Baumgartner and Casini 2017; Baumgartner, Casini, and Krickel 2018). Once we recognize a role for abductive inferences, however, we have a new option. Perhaps the “logic” of interlevel experiments is not based on conditions of mutual manipulability, but on a “logic” of abduction.

Consider, now, what additional attention to compositional explanation might mean for the study of abduction. Section 1 proposed that abductive inference can involve the postulation of entities that are qualitatively distinct from what is cited in the observations that form the premises of the abductive inference. It further noted that the qualitative distinctness of the compositional relata indicates that this type of abductive inference does not fit Schurz’s pattern of “micro-part abduction.” But Schurz (2008, 201) proposed “a classification of different patterns of abduction . . . which intends to be as complete as possible.” If it does not fit this pattern, does it fit some other pattern? If so, which? Or perhaps it does not fit in any of Schurz’s patterns, in which case we would need to recognize (at least?) one new pattern. More broadly, is Schurz’s omission symptomatic of the research on abduction not incorporating the research on compositional explanation?

## References

- Baumgartner, M., and A. Gebharter. 2016. “Constitutive Relevance, Mutual Manipulability, and Fat-Handedness.” *The British Journal for the Philosophy of Science* 67:731–56.
- Baumgartner, Michael, and Lorenzo Casini. 2017. “An Abductive Theory of Constitution.” *Philosophy of Science* 84:214–33.
- Baumgartner, Michael, Lorenzo Casini, and Beate Krickel. 2018. “Horizontal Surgicality and Mechanistic Constitution.” *Erkenntnis* 85:1–14.
- Bernstein, Julius. 1971. “Investigations on the Thermodynamics of Bioelectric Currents.” In *Founders of Experimental Physiology Biographies and Translations*, ed. John W. Boylan. Munich: J. F. Lehmanns Verlag. First published in 1902.
- Bernstein, Julius. 1912. *Elektrobiologie: Die Lehre von den elektrischen Vorgängen im Organismus auf moderner Grundlage dargestellt*. Braunschweig: Vieweg.
- Campos, Daniel G. 2011. “On the Distinction between Peirce’s Abduction and Lipton’s Inference to the Best Explanation.” *Synthese* 180:419–42.
- Cole, Kenneth S. 1968. *Membranes, Ions and Impulses*. Berkeley, CA: University of California Press.
- Craver, Carl. 2002. “Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory.” *Philosophy of Science* 69:83–97.

- Craver, Carl. 2007. *Explaining the Brain*. Oxford: Oxford University Press.
- Craver, Carl. 2008. "Physical Law and Mechanistic Explanation in the Hodgkin and Huxley Model of the Action Potential." *Philosophy of Science* 75:1022–33.
- Craver, Carl, and Lindley Darden. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press.
- Curtis, Howard J., and Kenneth S. Cole. 1942. "Membrane Resting and Action Potentials from the Squid Giant Axon." *Journal of Cellular and Comparative Physiology* 19:135–44.
- Douven, Igor. 2017. "Abduction." In *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), ed. Edward N. Zalta. <https://plato.stanford.edu/archives/sum2017/entries/abduction/>.
- Gebharder, Alexander. 2017. "Uncovering Constitutive Relevance Relations in Mechanisms." *Philosophical Studies* 174:2645–66.
- Gillett, Carl. 2010. "Moving beyond the Subset Model of Realization: The Problem of Qualitative Distinctness in the Metaphysics of Science." *Synthese* 177:165–92.
- Gillett, Carl. 2016. *Reduction and Emergence in Science and Philosophy*. New York: Cambridge University Press.
- Glennan, Stuart. 2005. "Modeling Mechanisms." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36:443–64.
- Harbecke, Jens. 2010. "Mechanistic Constitution in Neurobiological Explanations." *International Studies in the Philosophy of Science* 24:267–85.
- Harbecke, Jens. 2015. "The Regularity Theory of Mechanistic Constitution and a Methodology for Constitutive Inference." *Studies in History and Philosophy of Biological and Biomedical Sciences* 54:10–19.
- Hodgkin, Alan L., A. F. Huxley, and B. Katz. 1952. Measurement of current-voltage relations in the membrane of the giant axon of *Loligo*. *The Journal of Physiology* 116 (4):424–48.
- Hodgkin, Alan L. 1951. "The Ionic Basis of Electrical Activity in Nerve and Muscle." *Biological Reviews* 26:339–409.
- Hodgkin, Alan L., and Andrew F. Huxley. 1939. "Action Potentials Recorded from inside a Nerve Fiber." *Nature* 144:710–11.
- Hodgkin, Alan L., and Andrew F. Huxley. 1952. "Currents Carried by Sodium and Potassium Ions through the Membrane of the Giant Axon of *Loligo*." *The Journal of Physiology* 116:449–72.
- Hodgkin, Alan L., and Bernard Katz. 1949. "The Effect of Sodium Ions on the Electrical Activity of the Giant Axon of the Squid." *The Journal of Physiology* 108:37–77.
- Huxley, Andrew, and Robert Stämpfli. 1951. "Effect of Potassium and Sodium on Resting and Action Potentials of Single Myelinated Nerve Fibres." *The Journal of Physiology* 112:496–508.
- Levy, Arnon. 2014. "What Was Hodgkin and Huxley's Achievement?" *The British Journal for the Philosophy of Science* 65:469–92.
- Lipton, Peter. 1991. *Inference to the Best Explanation*, London: Routledge.
- Lipton, Peter. 2004. *Inference to the Best Explanation* (2nd Ed.), London: Routledge.
- Nernst, Walter. 1889. "Die Elektromotorische Wirksamkeit der Ionen." *Zeitschrift für Physikalische Chemie*, 4:129–81.
- Ostwald, Wilhelm. 1890. "Elektrische Eigenschaften halbdurchlässiger Scheidewände." *Zeitschrift für Physikalische Chemie* 6:71–82.
- Overton, Ernest. 1902. "Beiträge zur allgemeinen Muskel- und Nervenphysiologie." *Archiv für die gesamte Physiologie des Menschen und der Tiere* 92:115–280.
- Romero, Felipe. 2015. "Why there is Not Inter-Level Causation in Mechanisms." *Synthese* 192:3731–55.
- Schurz, Gerhard. 2008. "Patterns of Abduction." *Synthese* 164:201–34.
- Thagard, Paul. 1978. "The Best Explanation: Criteria for Theory Choice." *The Journal of Philosophy* 75:76–92.