Anatomy's Role in Mechanistic Explanations of Organism Behaviour

Aliya R. Dewey

Centre for Philosophy & Artificial Intelligence Research (PAIR)

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Explanations in behavioural neuroscience are often said to be mechanistic in the sense that they explain an organism's behaviour by describing the activities and organisation of the organism's parts that are "constitutively relevant" to organism behaviour. Much has been said about the constitutive relevance of working parts (in debates about the so-called "mutual manipulability criterion"), but relatively little has been said about the constitutive relevance of the organising relations between working parts. Some New Mechanists seem to endorse a simple causal-linking account: organising relations are constitutively relevant to organism behaviour if and only if (and because) they are causal relations that link the working parts that are constitutively relevant to organism behaviour. In this paper, I argue that the causal-linking account is inadequate because it neglects the constitutive relevance of anatomical relations that organise the working parts of a behaving organism. I demonstrate this by considering a case study where the anatomical organisation of the barn owl (*Tyto alba*) is constitutively relevant to their mechanism for sound localization. The anatomical organisation of this mechanism is best understood as the back-and-forth flow of task information across 7 "levels of anatomy" (a notion that I distinguish from levels of mechanism). A further implication, I conclude, is that at least some of the interlevel structure of neuroscientific explanation is accounted for by levels of anatomy, not levels of mechanism.

Keywords: behavioural neuroscience; mechanistic explanation; constitutive relevance; anatomy; organisation; multilevel explanation

Explanations of organism behaviour in neuroscience are often said to be *mechanistic*: they explain an organism's behaviour by describing the activities and organisation of the organism's parts (e.g., Bechtel & Richardson, 1993; Machamer et al., 2000; Craver, 2007; Bechtel, 2008; Kaplan, 2011; Kaplan & Craver, 2011; Piccinini & Craver, 2011; Piccinini, 2020). Of course, though, not every part of an organism is relevant to the organism's behaviour: e.g., action potentials in auditory cortex that process ambient sounds aren't relevant to performance on an image recognition task. Likewise, not every relation between parts is relevant to the organism's behaviour: e.g., the spatial relation that puts the primary visual cortex in a medial and posterior position with respect to the secondary visual cortex probably isn't relevant to performance on an image recognition task either. So, mechanistic explanation requires (at least) two standards for relevance—one for working parts and another for the relations that organise them. This notion of relevance is often called *constitutive relevance* (Craver, 2007, 139–160: Kaiser & Krickel, 2017: 752–753).[1]

Most attention has been paid to accounting for the constitutive relevance of working parts. The most influential such account is Craver's (2007) *mutual manipulability criterion*: one working entity (a ϕ-ing X) is constitutively relevant to another working entity (a ψ-ing S, such as a behaving organism) if and only if (a) X is part of S and (b) X's ϕ-ing and S's ψ-ing are *mutually manipulable*—i.e., holding certain background conditions fixed, (i) there is an ideal (bottom-up) intervention to X's ϕ-ing with respect to S's ψ-ing that makes a difference to S's ψ-ing and (ii) there is an ideal (top-down) intervention to S's ψ-ing with respect to X's ϕ-ing that makes a difference to X's ϕ-ing. There has been significant debate about this criterion. Some dispute the notion of working entities (Glennan, 1996, 2002; Tabery, 2004; Psillos, 2004; Campaner, 2006; Torres, 2008; Kaiser & Krickel, 2017), some dispute the notion of parthood (Kaiser, 2017; Kaiser & Krickel, 2017), and still others dispute the notion of mutual manipulability (Leuridan, 2021; Romero, 2015; Baumgartner & Gebharter, 2016; Baumgartner & Casini, 2017; Kästner, 2017). In response, some have proposed revisions to this criterion (Kaiser & Krickel, 2017; Krickel, 2018; Craver et al., 2021).

By comparison, much less attention has been paid to accounting for the constitutive relevance of organising relations. A simple position is that organising relations are *constitutively relevant* to behaviour if and only if (and because) they are the causal relations that *link together* working parts that are themselves constitutively relevant to behaviour. For example, Bechtel (2008: 17) suggests this when he says that "organisation is important… [because] the functioning of the mechanism requires the different operations to be linked to one another." However, New Mechanists often say that spatial and temporal organisation (which are distinct from causal links) are also relevant to mechanistic explanation (e.g., Craver, 2007: 251–254; Bechtel, 2008: 119–127). Perhaps, New Mechanists only think that these other forms of organisation are relevant insofar as they are *reliably indicative* of causal links. Or perhaps, they allow that other forms of organisation can be constitutively relevant to organism behaviour *per se*. An account is needed to clarify whether, which, and how organising relations that aren't causal links can be constitutively relevant.

One way to develop such an account is to examine how constitutive relevance for organising relations is determined in real neuroscientific explanations of organism behaviour. To do this, it's important to choose case studies carefully. For example, a central case study for Craver (2007) is hippocampal place cells in spatial navigation, but this wouldn't do for our purposes. After all, much is known about the relationship between place cells and spatial navigation behaviour but relatively little is known about the relationship between place cells and other working parts of the spatial navigation mechanism.[2] The focus on case studies like this might partially explain why the constitutive relevance of working parts has received so much more attention than the constitutive relevance of organising relations. To counter this, it is important for us to consider case studies

---

[1] The notion of constitution used by New Mechanists is distinct from the notion of constitution used by metaphysicists in other literatures (Kaiser & Krickel, 2017). It is used to refer to the non-causal form of relevance that some working parts bear to the whole mechanism's activity.

[2] This claim might strike some readers as implausible, given how much focus the rodent and primate hippocampus and spatial navigation have received in behavioural neuroscience. The key is the modifier 'relatively'. By the end of this paper, we'll see that *much less* is known about the anatomical organisation of spatial localization mechanisms than is known about the anatomical organisation of sound localization mechanisms.

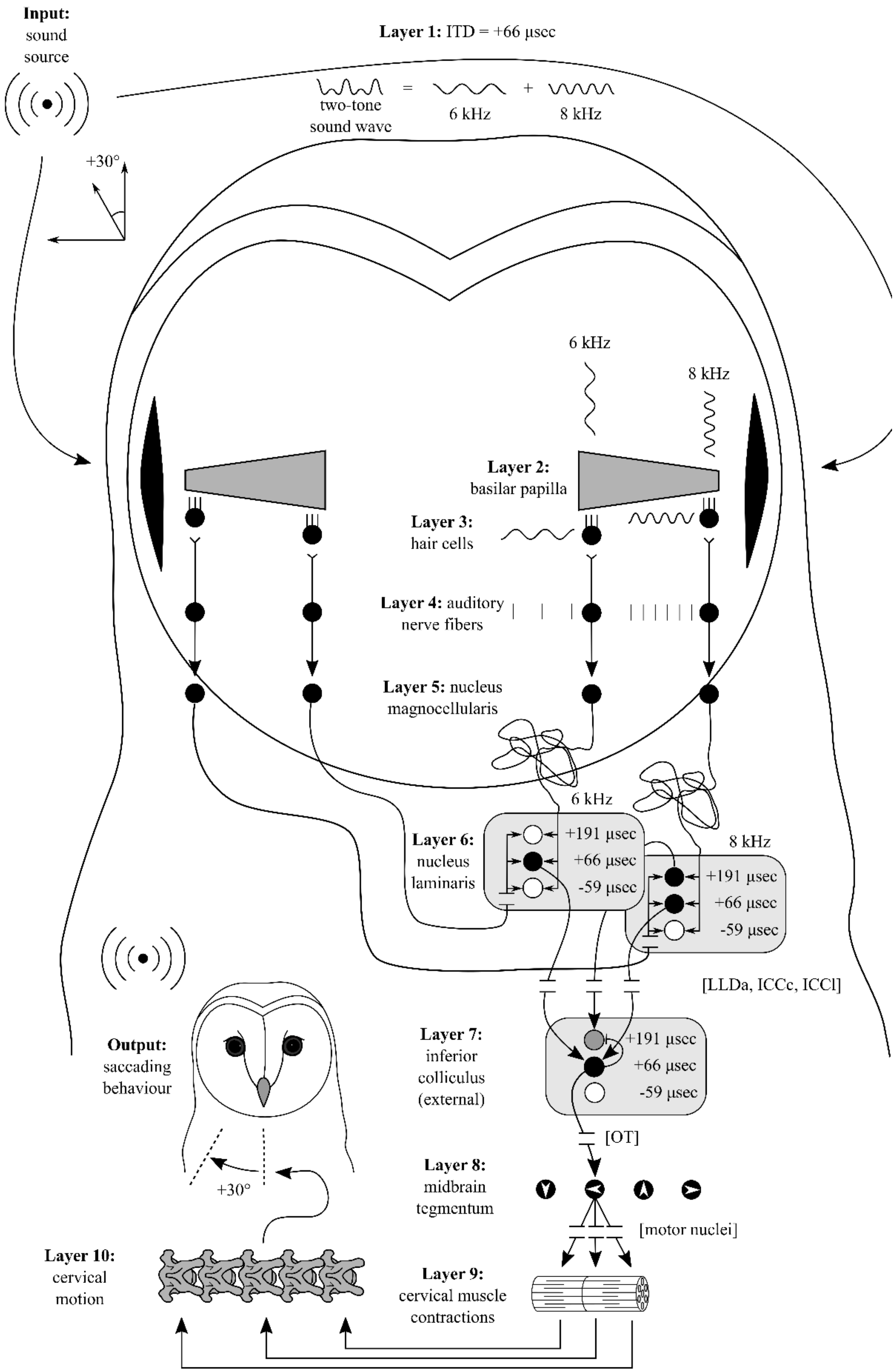

Figure 1. A complete figure depicting all the 10 working parts that are constitutively relevant to the basic FASL mechanism for reference. Figures 2–5 depict different sections of this figure. (Note: all illustrations are original.)

where much is known about the organising relations between parts of behavioural mechanisms.

An ideal example for this purpose is the mechanism for sound localization in the barn owl (*Tyto alba*), which has been completely mapped out from input-to-output between the 1970s and 1990s, primarily by the neuroethologists Eric Knudsen and Mark Konishi.

In this paper, my goal is to rationally reconstruct a comprehensive (but still incomplete) diagnostic account of constitutive relevance for both working parts and organising relations from the nearly-complete mechanistic explanation of sound localization in the barn owl. In §1, I'll review our case study of sound localization in the barn owl. In §2, I'll rationally reconstruct a novel diagnostic account of constitutive relevance for the working parts of the sound localization mechanism. In §3, I'll develop a novel—but incomplete—diagnostic account of constitutive relevance for the anatomical organising relations of the sound localization mechanism. We'll find that an anatomical conception of levels distinct from Craver's (2007) mechanist conception of levels provides a useful framework for revealing the anatomical relations that are constitutively relevant to behaviour. In §5, I'll conclude by cautioning against premature generalization from our case study.

§1. Frontal Azimuthal Sound Localization

Let's start by reviewing the mechanistic explanation for sound localization in the barn owl (*Tyto alba*). With a few exceptions that I'll note in §3, this review is meant to be representative of other reviews of the FASL mechanism (and its more complicated variants): see Konishi (1993), Peña & Konishi (2000), Peña & DeBello (2010), and Peña & Gutfreund (2014). A couple features of this system make it ideal for our purposes. First, neuroethologists have *completely* mapped the system from task inputs to behavioural outputs. Thus, they have used tacit criteria to individuate both the working parts and organising relations that are constitutively relevant to sound localization. This creates an ideal opportunity for us to rationally reconstruct comprehensive criteria for constitutive relevance from their complete mechanistic explanation of sound localization.

Second, the mechanism for sound localization in the barn owl is organised in a remarkable way: its anatomical organisation is much more complicated than its causal-linking organisation. After all, its causal-linking organisation is simple: (a) it is a feedforward network (it has no feedback connections), (b) it is a Markov chain (it has no memory and no skip connections), (c) it has only one processing stream (it is a perfect sequence of processing layers), (d) it only uses elementary arithmetic computations, (e) its activities are totally task-driven (it doesn't make any decisions), and (f) it contains only 10 working parts (which we'll call *layers*) that are constitutively relevant to behaviour (Fig. 1). By comparison, its anatomical structure is remarkably complex. In §3, we'll see that its 10 constitutively relevant working parts are distributed across 7 "levels of anatomy". In §3, I'll also define levels of anatomy and distinguish them from levels of mechanism.

To single out a mechanism with these useful properties, we need to precisely individuate our *explanandum*. It is the barn owl's successful performance on the simplest version of the sound localization task: the barn owl faces forward in perfect darkness and saccades (orients) their head to within a few degrees of a single broadband sound played from any position in the left-to-right

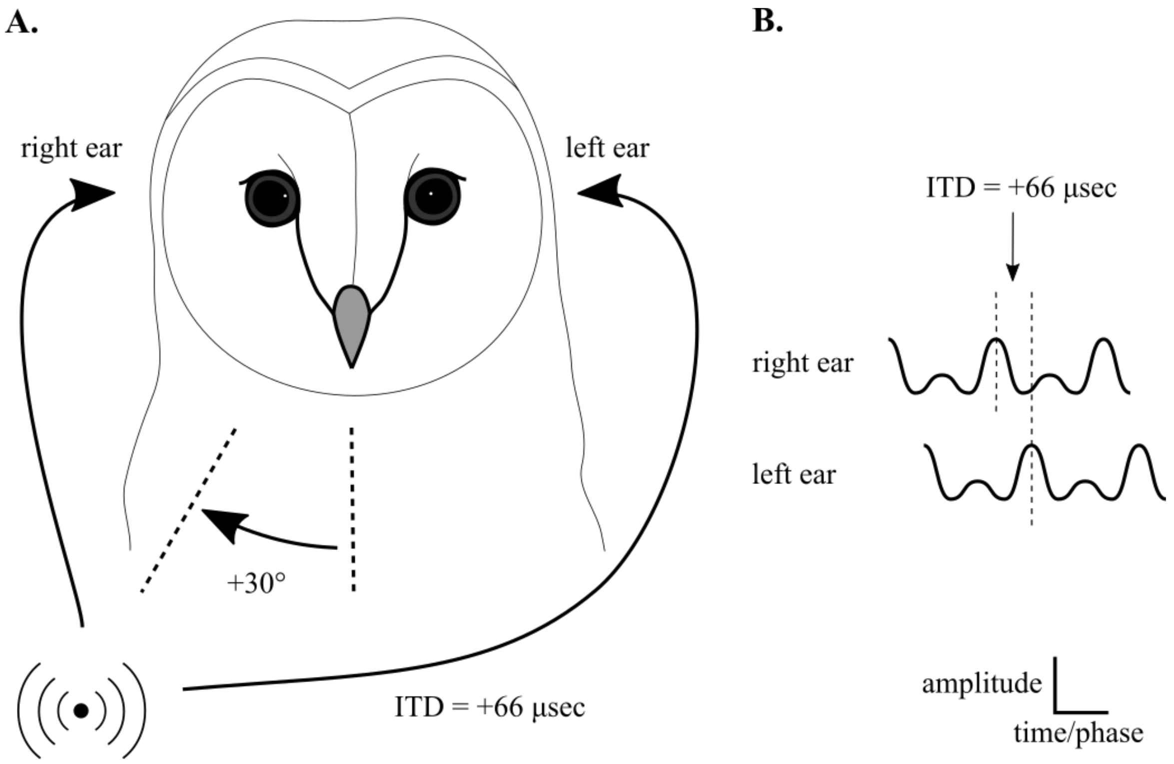

Figure 2. (A) Sound waves from sources at, e.g., +30° reach the left ear 66 µsec after the right ear. (B) The interaural (between-ears) time delay (ITD) is encoded by the delay in the propagation of the sound waves through the owl's left and right ears. Note that amplitude is on the y-axis and time (or phase) is on the x-axis.

plane (forward azimuth) (Knudsen et al., 1979).[3] [4] We'll call the mechanism that this basic task individuates the *frontal azimuthal sound localization* (FASL) system. Since there is only one variable input, there is no need for attention or decision-making in the basic FASL mechanism. Hence, multivariable tasks (e.g., sound localisation in visibility conditions) can individuate more complicated versions of the FASL mechanism that include attentional, multimodal, and decision-making components. We'll set aside such systems in this paper.

§1.1. Sensory Processing

Alright, let's begin. The only input to the basic FASL mechanism is information about the location of the sound source in the frontal azimuth. Unfortunately, location isn't the kind of thing that can directly stimulate sensory organs, like the ears. So, Layer 1 of the FASL mechanism re-encodes this information into a stimulus code: an interaural (between-ears) time delay (ITD) between the left and right ears. For example, if a sound originates at +30°, then the sound will travel (a) +60° around the right side of the barn owl's head and reaches their right ear first and (b) -120° around the front and left side of the head and reaches their left ear 66 µsec later. In other words, the geometry of the barn owl's head converts location information into an ITD code (Fig. 2). Since the ITD code can be encoded by sensory organs in the ears, it counts as a stimulus code.

---

[3] Requiring the barn owl to face forward at the beginning of each trial holds the proprioceptive inputs fixed.

[4] 'Saccading' usually refers to eye movements between fixation points, but it refers to head movements between fixation points in the barn owl literature (Masino & Knudsen, 1990), because barn owls cannot saccade with or even move their eyes due to the fact that their eyes are tubular, not spherical (Pettigrew & Konishi, 1976).

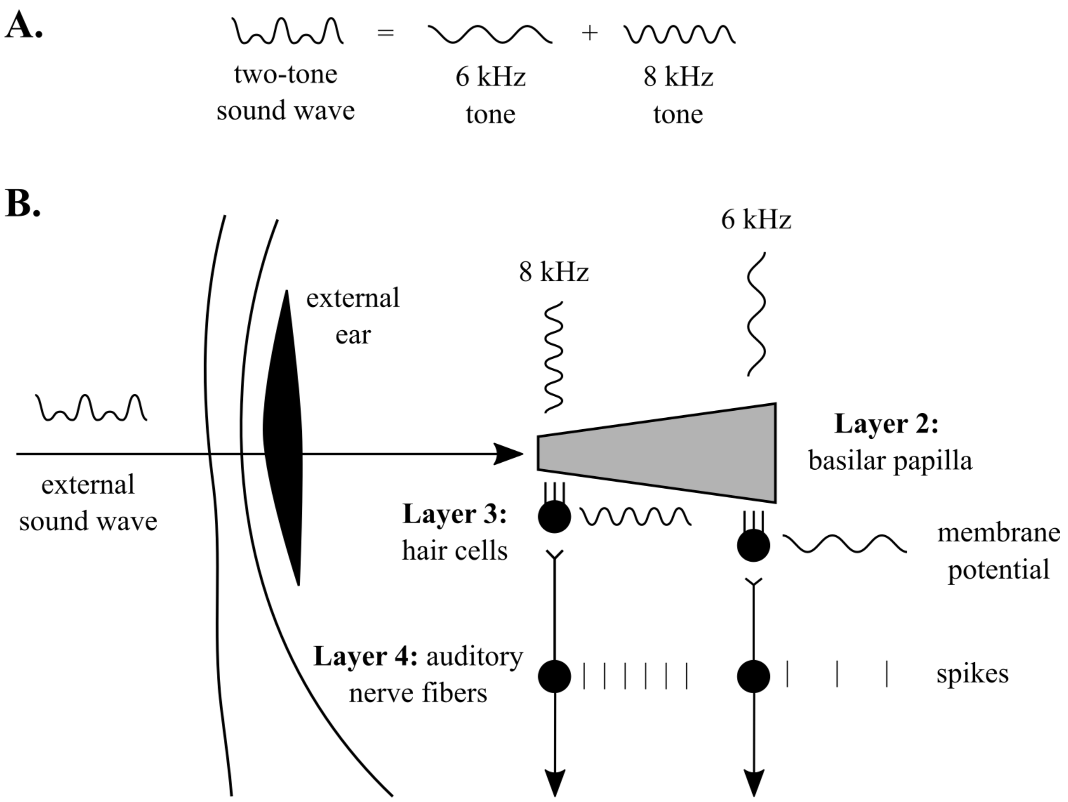

Figure 3. (A) Every sound wave is composed of single-frequency components. (B) The soundwave enters each ear and strikes each BP. Each BP separates out the two tones and HCs transduce the signals in separate frequency channels along the length of the BP. These frequency channels are preserved by the ANFs, which use spikes to encode the timing of peaks in the membrane potentials of HCs.

Most sound waves are composed of multiple frequency components (Fig. 3A) and the FASL mechanism requires those frequency components to be separated for downstream processing in Layer 7. Layer 2 satisfies this need for frequency separation in advance during stimulation. It involves soundwaves striking a stiff membrane known as the *basilar papilla* (BP) in the left and right ears at an ITD. Each BP is tapered. Since higher-frequency sound waves have more energy than lower-energy ones, the higher-frequency components concentrate at the thick, rigid ends (where their higher energy quickly dissipates) and lower-frequency components concentrate at the thin, flexible ends. So, a frequency map forms along the length of each BP at an ITD during stimulation from the sound wave.

Since neurons communicate with electrical signals, the FASL mechanism requires task information to be *transduced*: to be encoded in an electrical code. Layer 3 satisfies this need for transduction by lining hair cells along the length of each basilar papilla (BP).[5] Each peak in a sound wave causes the hair cells (HCs) on each BP to collide into the adjacent tectorial membrane and each collision causes hair-like fibers (*stereocilia*) on the HCs to bend against the membrane, opening ion channels and depolarizing the HCs. This encodes phase information from each sound wave in the continuous fluctuations of the membrane potential (HCs don't generate action

---

[5] I redefine most acronyms in each paragraph for readability, since this section contains a lot of unfamiliar terms.

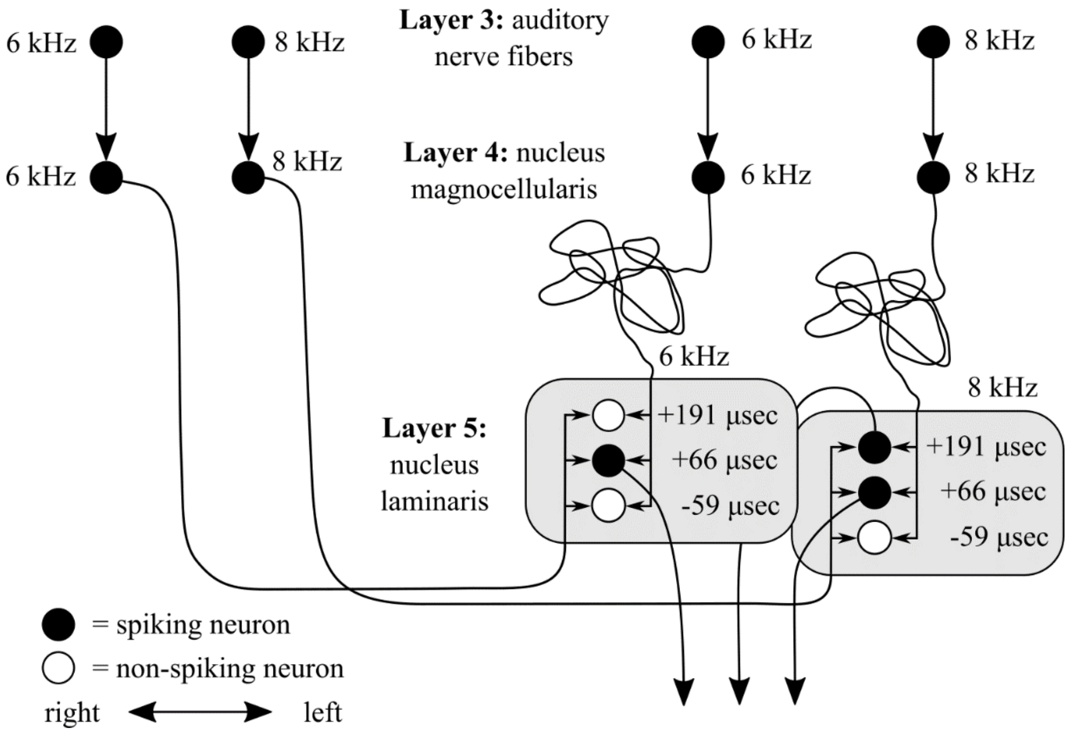

Figure 4. Axonal projections from left and right NM cells are the same length, so that they preserve the ITD between right and left auditory signals. 6 kHz signals are far enough apart that they cross only one signal (at the 66 µsec layer) from the other ear before clearing the NL. But 8 kHz signals are so close together that they cross two signals (at the 66 µsec layer and the 191 µsec (66 µsec + 125 µsec = ITD + time between phase-locked signals) layer before clearing the NL. This creates phase ambiguity: the true ITD (66 µsec) appears in both the 6 kHz and 8 kHz frequency channels, and the phantom ITD (191 µsec) appears in the 8 kHz channel.

potentials). Also, HCs transduce unique positions on the frequency map, such that the frequency map of the BP is copied over by the population of HCs (Fig. 3).

Discrete action potentials require more energy and encode less information than graded potentials, but they have the critical advantage of preventing signals from attenuating as they propagate over longer distances (Jack et al., 1975; Laughlin et al., 1998; Sarpeshkar, 1998; Sengupta et al., 2014). To secure this advantage, the FASL mechanism requires task information to be *discretized*: encoded with action potentials (vs. graded potentials). Layer 4 satisfies this need by having hair cells (HCs) synapse onto auditory nerve fibers (ANFs), which generate action potentials at the peak of each signal from the sound wave (Sullivan & Konishi, 1984; Köppl, 1997b). Moreover, ANFs transduce unique positions on the frequency map of the HCs, creating “frequency channels” that remain until Layer 7 (hence, I won’t mention them further till our discussion of Layer 7).

Auditory nerve fibers (ANFs) use a *rate code* to encode information about soundwave amplitude: higher noise levels increase the frequency of their action potentials. This is a problem because the FASL mechanism doesn’t need amplitude information, such that encoding amplitude information with the rate code prevents it from encoding task-relevant information. So, the FASL mechanism requires this information to be discarded. Layer 5 is known to satisfy this need but how exactly this is achieved isn’t fully understood yet. What we do know is that it is achieved between ANFs and the nucleus magnocellularis (NM) (Sullivan & Konishi, 1984). Meanwhile, the NM also re-encodes phase information from the ANFs (Köppl, 1997a).

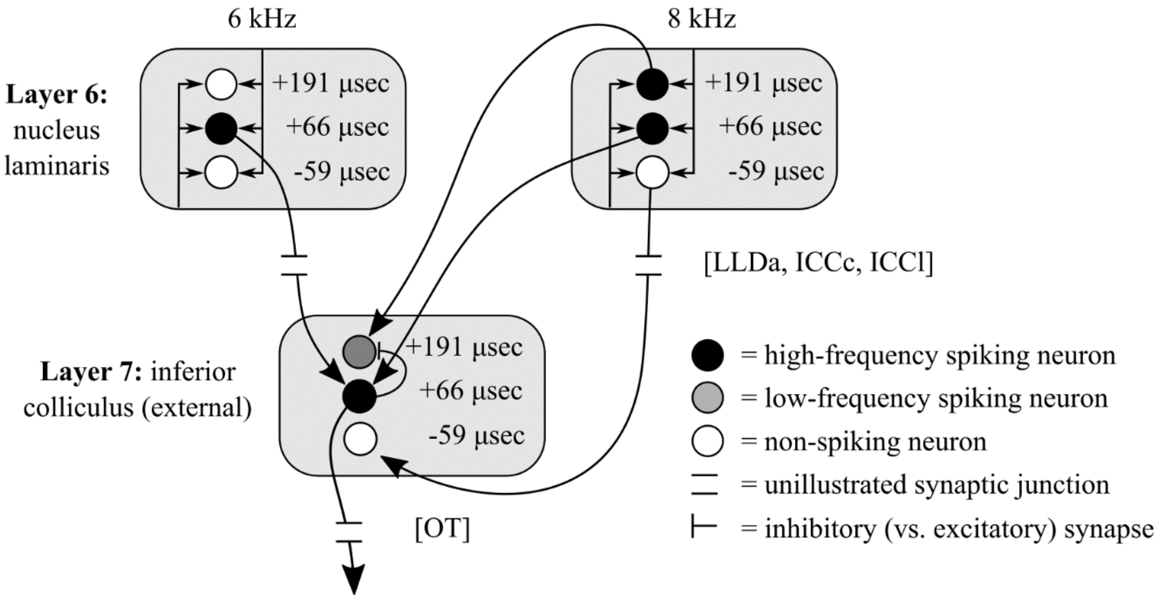

Figure 5. The ICX receives the true IPD signal (e.g., 66 µsec) across all frequency channels and phantom IPD signals (e.g., 191 µsec) in specific frequency channels (e.g., 8 kHz). It uses lateral inhibition between frequency channels to single out the strongest signal: the true IPD signal will overpower, inhibit, and silence the phantom IPD signals. Then the ICX delivers the mode-ITD signal to the optic tectum (OT).

Encoding and re-encoding task information in an ITD code requires the temporal resolution of a few microseconds, which is costly for the FASL mechanism and thereby creates the need for a cheaper, non-temporal code. Layer 6 solves this problem using an impressive circuit. Each nucleus magno-cellularis (NM) cell sends axonal projections to each nucleus laminaris (NL): e.g., (a) the left NM sends axonal projections into the dorsal (upper) surface of the left NL, makes synapses on several NL cells as it passes straight through the left NL, and ends at the ventral (lower) surface; and (b) the right NM sends axonal projections into the ventral (lower) surface of the left NL, makes synapses on several NL cells as it passes straight through the left NL, and ends at the dorsal surface of the left NL (Fig. 4; Takahashi & Konishi, 1988; Carr & Konishi, 1998, 1990).

The axonal projections from each nucleus magnocellularis (NM) cell to each NL are *size-matched* (Seidl et al., 2010; Carr et al., 2016). This way, the signals from a sound at 0° (ITD = 0) from each ear would reach each nucleus laminaris (NL) at the same time, cross in the middle of each NL, and activate the coincidence detector cells in that middle position (Carr & Konishi, 1988, 1990; Fischer et al., 2008). But the signals from a sound at +30° (ITD = 66 µsec) would reach the left NL through the right NM 66 µsec *before* they reach the left NL through the left NM, giving the right NM signal a 66 µsec head-start to travel through the NL. As a result, the signal from the right NM will reach closer to the dorsal surface of the NL before it crosses the signal from the left NM, such that both will activate the coincidence detector cells in that dorsal position. Hence, these projections form a spatial map in the NL that encodes ITDs (Sullivan & Konishi, 1986; Carr & Konishi, 1988, 1990; Peña et al., 2001).

While this circuit is impressive, it has a serious problem: *phase ambiguity*. In the explanation that I just gave, I pretended that there were only two discrete signals (one left and one right) locked to the same phase of a single sound wave. In reality, though, there is a continuous stream of sound waves. Unfortunately, higher frequency sound waves are so short that the time between phase-locked monaural signals (e.g., 167 μsec for 6 kHz; 125 μsec for 8 kHz) is often less than the time it takes for the signals to pass through the nucleus laminaris (NL) (~200 μsec) (Carr & Konishi, 1988). Thus, the first pair of signals won't clear the NL before they each encounter the second pair of signals. Instead, the first pair of signals will coincide with the second pair of signals at a pair of different positions in the left NL. As a result, the NL will encode three crossing events: one that corresponds to the real ITD (the true position in the frontal azimuth) and two that correspond to a "phantom" ITD (false positions in the frontal azimuth). This is known as *phase ambiguity*.

Fortunately, though, there is a key difference between the true ITD and every phantom ITD. The true ITD is frequency-invariant because it's the result of the location of the sound source in the frontal azimuth, which is frequency-invariant. By comparison, every phantom ITD is frequency-variant because it's the result of the wavelengths of certain high-frequency soundwaves (which are frequency-variant) being shorter than the distance between the dorsal and ventral surfaces of the nucleus laminaris (NL). But the NL separately encodes these crossing events for each frequency channel (Sullivan & Konishi, 1986; Takahashi & Konishi, 1988), so it preserves information about the true ITD: it's the only ITD that appears in every frequency channel, whereas the phantom ITDs will only appear in a few frequency channels.

The NL encoding crossing events in separate frequency channels does solve the problem of phase ambiguity, but the FASL mechanism requires the selection of the true signal and the discarding of the phantom signals: it requires *phase disambiguation*. Layer 7 satisfies this need by using the external nucleus of the inferior colliculus (ICX) to sum up the signals across frequency channels for each position on the map of the ITDs (and hence, of the frontal azimuth) (Fig. 5; Knudsen, 1983, 1984; Takahashi & Konishi, 1986; Wagner et al., 1987; Peña & Konishi, 2000).[6] Thus, the strongest signal is the true signal, and any weaker signals are phantom signals. Then the ICX uses inhibitory inter-neurons to facilitate the true signal across frequency channels (primary peak facilitation) and suppress the phantom signals (secondary peak suppression) (Moiseff & Konishi, 1981; Peña & Konishi, 2000). Overall, this ensures that the ICX approximates the mode (not the sum) of ITDs across frequency channels.[7]

§1.2. Motor Processing

The external nucleus of the inferior colliculus (ICX) encodes task information in a place code, but saccading movement requires a hybrid code in the muscles: a place code for the specific group of muscles that are supposed to contract and a rate code for the strength of contraction in each muscle. After all, further azimuthal positions (e.g., +30° vs. +10°) require more force to saccade the head to. Layer 8 satisfies this functional need by using the optic tectum (OT), which copies the frontal azimuthal map from the ICX (more on the OT in §3), to send *convergent* projections through the

---

[6] My explanation here bypasses two nuclei: the central nucleus of the inferior colliculus (ICC) and the dorsal nucleus of the lateral lemniscus (LLD). The functions of these nuclei are uncertain and appear to involve *noise reduction* (see Peña & Debello, 2010 for review). I'll argue in §3 that they aren't constitutively relevant to sound localization, so I'll pre-emptively exclude them here to save space.

[7] Unless the sound is a tone (which contains one frequency component), the mode signal across frequency channels will always be the true signal. Saberi et al. (1999) confirmed that FASL mechanisms cannot resolve phase ambiguities in tones: the barn owls saccade to the hypothetical sources of the phantom IPDs.

medial efferent pathway to synapse onto premotor neurons in nuclei in the medial dorsomedial region of the midbrain tegmentum (mTeg) (du Lac & Knudsen, 1990; Masino & Knudsen, 1992; Cazettes et al., 2018). These junctions translate the 2D (azimuth and altitude) multi-sensory map into a set of four coordinates (up, down, right, left), which are place-coded and so interpreted as a *head-centred motor map* (Masino & Knudsen, 1990, 1993). Further azimuthal positions are encoded with higher spiking rates in order to build more excitatory drive to the muscles (Masino & Knudsen, 1992, 1993).

Task information in the mTeg is encoded in a place code (a set of coordinates)—which is analogue, spatial, and mechanical—and a rate code (the firing rates at each coordinate)—which is discrete, temporal, and electrical. Saccading behaviour is a continuous motion in space, though, so task information needs to be completely encoded in an analogue, spatial, and mechanical format. In other words, the FASL mechanism needs to (a) *de-discretize* the code, by converting the signal from a discrete rate code back into a continuous graded code, and (b) to *de-transduce* the code, by converting the signal from an electric code into a mechanical code. Layer 9 satisfies both needs at once by using the *motor nuclei*—which make simple reconfigurations to the space and rate codes (e.g., boosting the rate codes to amplify excitatory drive) from the mTeg (Masino & Knudsen, 1992)—to re-encode task information in continuous, mechanical contractions of up to 31 pairs of cervical muscles (Masino & Knudsen, 1990).

Finally, the FASL mechanism needs to convert the muscular contractions into movement. Layer 10 achieves this with the contraction of cervical muscles against the resistance of the cervical skeleton (and the rest of the barn owl's body) to orient the barn owl's head into the same frontal azimuthal position that the sound source is located in. If all 10 layers achieve their functions (and if the sound contains more than one frequency component within the barn owl's frequency register), then the FASL mechanism will achieve its function: it will cause the barn owl's head to saccade to the frontal azimuth of the sound source, successfully performing the FASL task and so earning a food reward.

This explanation of the basic FASL mechanism is complicated but it *excludes* a tremendous amount of information in producing a behavioural response to the FASL task: e.g., it doesn't mention any computations that occur inside the neurons, the local inhibitory circuits in each brain region (except for the ICX in Layer 7), any control mechanisms, etc. Nevertheless, neuroethologists who study the FASL mechanism regard this explanation as *complete* (or almost so): research on the components of the *basic* FASL mechanism mostly ended when Saberi et al. (1999) confirmed that the ICX performs phase disambiguation. Today, some research continues on the internal functioning of the FASL mechanism's working parts at lower levels of mechanism (e.g., Christianson & Peña, 2007; Kuokkanen et al., 2010) and on more advanced versions of the FASL mechanism, such as those capable of learning. This suggests that neuroethologists regard our mechanistic explanation (or something like it) as successfully identifying all of the active parts and organising relations that are relevant to sound localization in the barn owl on the basic FASL task, despite excluding so much information.

§2. Constitutive Relevance for Working Parts

Let ‘FASLx’ refer to our mechanistic explanation of the barn owl’s performance on the FASL task. FASLx is built with implicit, unprincipled diagnostic criteria for which working parts are constitutively relevant to FASL behaviour. However, since these diagnostic criteria aren’t explicit and principled, it shouldn’t be surprising that they are *imperfect*: we’ll find that FASLx includes some irrelevant information and excludes some relevant information. The goal of this section is to rationally reconstruct from FASLx explicit, principled criteria that are diagnostic of constitutive relevance for working parts in the FASL mechanism. We’ll find that this is possible without appealing to criteria for constitutive relevance of organising relations—suggesting that constitutive relevance for working parts is *prior to* constitutive relevance for organising relations. Instead, we’ll find that these criteria appeal to the informational, normative, spatial, temporal, and physical properties of the working parts. Finally, we’ll discuss how these diagnostic criteria might fit with the mutual manipulability criterion for the constitutive relevance of working parts.

§2.1. Six Competencies

One salient feature of FASLx is that it focuses on the many difficulties that the FASL mechanism faces during performance on the FASL task: it has to move information about the task state across far reaches of space and time, through the extremely noisy and lossy channels of the barn owl’s brain and body, and then encode that information in a behavioural response that is appropriate to the task state. Another salient feature of FASLx is that it singles out the working parts that make distinctive contributions to the process by which the FASL mechanism overcomes each of these difficulties. I’ll refer to these problem-solving contributions loosely as *competencies*, but I don’t mean to be ontologically committed to competencies (these are just diagnostic criteria after all). By my count, FASLx individuates working parts that manifest six general kinds of competencies. I’ll show these can be assimilated into three distinct kinds of competencies and each working part that is constitutively relevant to FASL performance manifests at least one of them.

One competency is that the FASL mechanism achieves a sustained flow of information through all 11 of its layers. There are many notions of information (Piccinini, 2020), but the relevant notion here is *natural semantic information*: a state S carries natural semantic information that a fact F obtains if and only if S raises the probability that F. Note that S raises the probability that F if and only if the conditional probability of F given S is greater than the unconditional probability of F, or $P(F|S) > P(F)$. In other words, natural semantic information is just correlation between information-carrying states and facts: a state S carries natural semantic information that a fact F obtains if and only if S is *correlated* with F obtaining.

A second competency is that the FASL mechanism achieves a flow of natural semantic information about three properties in particular: (a) frontal azimuthal position, (b) frequency components in the sound wave, and (c) the amplitude of each frequency component in the sound wave. Only the first property is essential to the task: the frontal azimuthal position of the sound source is *necessary* and *sufficient* to determine which frontal azimuthal position the barn owl *should* saccade to (i.e., the same position). However, the FASL mechanism cannot achieve the task unless it also encodes *ancillary* information about the frequency components of the sound wave so that it can perform phase disambiguation in Layer 7 (the ICX). Likewise, it must also encode information about the amplitude so that its mechanical energy can be used for transduction and then discarded at Layer 4 (the ANFs). Thus, the FASL mechanism achieves the encoding of task information that is (a) necessary for any solution to the task and (b) unnecessary for all solutions to the task but necessary for the FASL mechanism's particular solution to the task.

A third competency is that the FASL mechanism causally produces a series of states that not only encode natural semantic information about the three task-related properties but do so in a way that ultimately produces a response that is *appropriate* to the task state. After all, FASLx focuses on tracking the *functional* (or *correct*) flow of essential and ancillary task information through all 11 layers of the FASL mechanism, which results in its *successful* performance. Each component has the *function* to carry natural semantic information about the task in a particular way. For example, FASLx reports that the interaural time delay (ITD) at the nucleus magnocellularis (NM; Layer 6) being 66 μsec was correlated with the sound source being at +30° because FASLx only considered its functional state. So, FASLx implicitly claims that the ITD of the NM has two *functions*: to carry natural semantic information that (a) the source of the sound is at a particular frontal azimuthal position and (b) the sound itself has a particular frequency distribution.

Before we continue, I should note that informational functions are associated with informational teleosemantics, which is the thesis that informational functions are sufficient for representation. In other words, a state S having a function to carry natural semantic information that a fact F obtains is sufficient for S to represent that F (Neander, 2017; Shea, 2018; c.f., Dretske, 1981). However, informational teleosemantics is orthogonal to our purposes. After all, informational teleosemantics aims to account for what makes a state a representation. However, our purpose here is to account for something different: what makes a working part constitutively relevant to FASL behaviour. My proposal is that having a function to encode task-related information in some particular state is diagnostic of a working part being constitutively relevant to FASL behaviour. Whether this is true is independent of the matter whether informational functions are necessary or sufficient to ground representations.[8]

A fourth competency is that the FASL mechanism progressively re-encodes task information in states with novel spatiotemporal properties. This is extraordinarily difficult: doing so requires the most sophisticated and specialized anatomical equipment in the FASL mechanism. For example, take the interaction between the nucleus magnocellularis (NM; Layer 5) and the nucleus laminaris (NL; Layer 6) when the sound source is at +30°. The NM has the function to encode this information by entering a state that is *spatiotemporally discontinuous*: task information is encoded in the temporal distance of 66 μsec between the right NM and activity in the left NM, but no information is encoded in the space between the right and left NM or in the subintervals of time (<66 μsec) between activity in the right NM and activity in the left NM. By comparison, the NL has the function to re-encode this information by entering a state that is *spatiotemporally*

---

[8] Informational teleosemantics entails a controversial claim for the FASL mechanism: that the barn owl's behaviour, the pattern of contractions in their muscles, etc. all *represent* the location of the sound source. For example, the barn owl's head position (the output layer) has the function to enter a state that is not only correlated with but equal to the angular position of the sound source: to saccade to +30° if and only if the angular position of the sound source is +30°. If informational functions are sufficient for representation, then this notion of representation isn't distinctive to brains (or minds): it's shared by non-neural systems like barn owl bodies. Advocates of embodied cognition (such as myself) may embrace this implication, but others may be inclined to count this as a problem for informational teleosemantics.

continuous: activity in the NL occurs at the single location where signals in size-matched delay lines from the right and left NM cross.

There is some precedent for singling out processes that encode information in states with different spatiotemporal properties. Piccinini (2015) proposes that *physical computation* is "the processing of vehicles by a functional mechanism according to rules that are sensitive solely to differences between spatiotemporal parts of the vehicles". For example, he notes that digital computation uses rules that are only sensitive to the spatial or temporal position of each digit in a sequence of digits. Likewise, the working parts of the FASL mechanism follow rules that are only sensitive to their spatial or temporal properties: e.g., temporal distance between neural activity in the left and right NM, spatial distance of neural activity between the dorsal and ventral surface of the NL.[9] The fact that constitutively relevant activities in the FASL mechanism change the spatiotemporal properties of the states that encode task-related information is evidence, I think, that Piccinini's account of physical computation is singling out a relevant explanatory kind.

Although I don't purport to be offering a general account of behavioural mechanisms, I will flag that this fourth competency often instantiates differently in other behavioural mechanisms besides the FASL mechanism. After all, the FASL mechanism has an unusual property: information about all three properties of the task is encoded together in states of the same parts—until ancillary information is discarded. For example, bilateral hair cells (HCs; Layer 3) encode sound location in their ITD, the frequency composition of the sound in the position of active HCs along the basilar papilla, and the amplitude of each frequency component in the amount of depolarization for each active HC. If anything, this is the exception, not the rule. Many behavioural mechanisms use different parts to carry information about different properties of the task—often because they transduce different task properties in different modalities.

For example, the hippocampus uses a rate code by place cells to encode relevant positions (e.g., reward location, boundaries) in egocentric space. It's often thought that spatial information flows through the dorsal stream to the allocentric map in the entorhinal cortex and object information flows through the ventral streams until it is integrated with the entorhinal map to form the egocentric map in the hippocampus (Suzuki et al., 1997; Gaffan, 1998; Manns & Eichenbaum, 2006; Knierim et al., 2006; Diana et al., 2007; c.f., Connor & Knierim, 2017). If this is true, then the hippocampus takes task information encoded in *spatiotemporally separated* representations in ventral and dorsal streams and then recodes it in *spatiotemporally integrated* representations in the hippocampus. If we wanted to produce diagnostic criteria that were applicable to all behavioural mechanisms, we'd need to cash out this notion of spatiotemporal competency in a way that could capture both the competency exhibited by the hippocampus and the NL. We aren't aiming for generality in this paper, though, so this is unnecessary for our purposes.

---

[9] The notion of *medium-independence* might be helpful here too (Haugeland, 1985; Piccinini, 2015): the barn owl could still perform the FASL task if their working parts were replaced so long as the new working parts (a) have the same spatial and/or temporal degrees of freedom and (b) are appropriately interfaced with neighboring working parts. For example, the NL is medium independent insofar as the FASL mechanism would continue to operate if it were replaced with an AND gate in the same spatial position (same degrees of freedom) with input microelectrodes embedded in the bilateral NM delay lines (appropriately interfaced to receive the signals) and output microelectrodes embedded in the central core of the inferior colliculus (appropriately interfaced to deliver the signals).

Layer	Name	Competency	Notes
input	task		
1	head	1	Location information is encoded in ITD
2	basilar papillae (BP)	2	Frequency information is encoded in spatial map
3	hair cells (HCs)	1, 2	Mechanical work is transduced into electrical work
4	auditory nerve fibers (ANFs)	1, 2	Membrane potential (info stored in properties of one event) is discretized into rate code (info stored in temporal relations between events)
5	nucleus magnocellularis (NM)	3	Amplitude information is discarded
6	nucleus laminaris (NL)	1	ITD is translated into spatial map
7	external nucleus of the inferior colliculus (ICX)	1, 3	Phase disambiguation and frequency information is discarded
8	midbrain tegmentum (mTeg)	1	Spatial map is partly translated into a rate code
9	cervical muscles	1	Electrical work is de-transduced into mechanical work and rate code is de-discretized into analogue code
10	cervical movement	1	Rate code is translated into a space code
output	saccading behaviour		

Table 1. A table depicting the layers that are constitutively relevant to FASL behaviour in the barn owl, the basic kinds of competencies they manifest (which are diagnostic of their constitutive relevance), and notes about the particular ways that they manifest these basic kinds of competencies.

A fifth competency is that the FASL mechanism re-encodes task information in states with novel *physical* properties. For example, transduction in hair cells (HCs; Layer 3) takes task information encoded in mechanical work (movement of basilar papillae) and re-encodes it in electrical work (movement of ions). Likewise, de-transduction in the cervical muscles (Layer 9) takes task information encoded in electrical work and re-encodes it back into mechanical work. I refer to this kind of change as a change in “physical properties”. By “physical properties”, then, I mean to refer to properties relevant to physics, like mechanical and electrical work. I don’t mean to refer to “physically realized” properties, which include very many properties that are irrelevant to physics: e.g., being a neuron and other anatomical properties. For our purposes, nothing really turns on the terminology that we choose to refer to these properties in our diagnostic criteria.

A sixth competency is that the FASL mechanism discards ancillary information after it’s been used, reducing noise and freeing up dimensions to represent other information. For example, consider the interaction between the auxiliary nerve fibres (ANFs; Layer 4) and the nucleus magnocellularis (NM; Layer 5) when the sound source is at +30°. The bilateral ANFs and the bilateral NM have the function to encode this information by entering a spatiotemporally extended state where activity in the right nucleus precedes activity in the left nucleus by 66 μsec. But the bilateral ANFs encode information about the amplitude of each frequency component in the sound and the bilateral NM do not. This difference matters for two reasons. First, amplitude information is only needed for transduction, which is performed by hair cells in the previous layer. Second, amplitude information is encoded by spiking rates, which are needed for encoding ITD in the NL, so it must be discarded to free up spiking rates to encode ITD.

§2.2. Competency Criterion

Some of these six competencies are subsumed by others and some are distinct, so let's take a tally. I propose that there are three basic kinds of competencies, which are collectively manifested by all the constitutively relevant working parts of the FASL mechanism (see Table 1):

> **Competency 1:** encoding natural semantic information about states that are *essential* to the FASL task (i.e., necessary for the non-accidental successful performance of the task by any possible system) in states of the barn owl's FASL mechanism with novel spatiotemporal or physical properties.
>
> **Competency 2:** encoding natural semantic information about states that are *ancillary* to the FASL task (i.e., necessary for the non-accidental successful performance of the task by the barn owl's FASL mechanism but unnecessary for other possible mechanisms) in states of the barn owl's FASL mechanism with novel spatiotemporal or physical properties.
>
> **Competency 3:** discarding natural semantic information about states that are ancillary to the FASL task when that information is *unnecessary* for the non-accidental successful performance of the task by the barn owl's FASL mechanism.[10]

Moreover, I propose that these three basic competencies are not only sufficient for constitutively relevant parts but also necessary: any working part that doesn't manifest at least one of these basic competencies is constitutively irrelevant to FASL performance. This implies that FASLx includes some irrelevant information: e.g., that the bending of stereocilia in hair cells (HCs; Layer 3) twists their ion channels at their base into an open conformation, causing the influx of positive charge from the extracellular environment and thus, depolarization in HCs. After all, this working part doesn't manifest any of Competencies 1–3. In a sense, then, it's too "low-level" to be constitutively relevant to the FASL performance. Thus, my proposal is a *rational* reconstruction—not a *perfectly accurate* reconstruction—of the relevance criteria used to assemble FASLx.

To save space, I pre-emptively excluded information from FASLx that is constitutively irrelevant to FASL performance given our criteria but that features prominently in other reviews of the FASL mechanism (e.g., Konishi, 1993; Peña & Konishi, 2000; Peña & DeBello, 2010; Peña & Gutfreund, 2014): activities in the central nucleus of the inferior colliculus (ICC), the dorsal nucleus of the lateral lemniscus (LLD), the optic tectum (OT), and the cervical motor nuclei. Notably, for example, the OT contains a striking map of the visuo-auditory field that integrates visual and auditory information (Knudsen, 1982, 1984). Without visual inputs, though, the OT and the ICX (Layer 7) encode task information in states with the same spatiotemporal and compositional properties: both use a rate code to plot the sound source on a one-dimensional map of the frontal azimuth.[11] This is true for the ICC, LLD, and motor nuclei too. If FASLx were fully representative

---

[10] These are specific competencies that are manifested in the barn owl's FASL mechanism, but they may be indicative of general kinds of competencies that are manifested in other relevantly similar behavioural mechanisms.

[11] Of course, this isn't to say that there aren't differences in the non-spatiotemporal properties of the place codes in the ICX and the OT. For example, place codes in the ICX encode task information in the distance in space between the anterior and posterior edges of the ICX (Knudsen & Konishi, 1978, 1979) whereas place codes in the OT encode

of other reviews in the literature, it would have included this information, and our account of constitutive relevance for working parts would have implied that it was mistaken for doing so: these working parts don't manifest Competencies 1–3.

Overall, then, I propose that these three basic competencies diagnostically indicate all and only the working parts that are constitutively relevant to the barn owl's performance on the FASL task: a working part is constitutively relevant to the barn owl's FASL mechanism if and only if the working part has the function (which it can fail to achieve, of course) to manifest any of these three basic competencies. I'll refer to this as the *competency criterion* for the constitutive relevance of working parts in the barn owl's FASL mechanism.

§2.3. Mutual Manipulability Criterion

Our diagnostic account of constitutive relevance for working parts indicates that it is possible to individuate all the working parts that are constitutively relevant to FASL performance without individuating the organising relations between the working parts that are constitutively relevant to FASL performance. This supports current efforts by New Mechanists to start by developing accounts of constitutive relevance for working parts, such as the mutual manipulability criterion. Nevertheless, our competence criteria for constitutive relevance of working parts are categorically different from the mutual manipulability criteria. They appeal to the informational and normative relations between the working parts and the FASL performance of the whole barn owl and to spatiotemporal and physical properties of the working parts themselves. They do not appeal to any relations of mutual manipulability between the activities of the parts and the FASL performance of the whole barn owl. This raises the question: what makes these criteria so different?

At least one reason why these criteria are different (perhaps the only reason) is that they are meant to be diagnostic in different contexts. Mutual manipulability criteria are meant to be diagnostic in the context of *scientific experimentation*, when neuroscientists are performing (a) top-down interventions on the barn owl's FASL performance in search of differences in the activities of their parts or (b) bottom-up interventions on the activities of the barn owl's parts in search of differences in their FASL performance (Craver, 2007). By comparison, competence criteria are meant to be diagnostic in the context of *scientific review*, where scientists (a) draw results from dozens or more experimental studies, (b) filter these results for relevance, (c) assemble them into a concise but comprehensive explanation of organism behaviour, and then (d) publish the resulting explanation as a scientific review article. Explanations like FASLx in scientific review papers tend to abstract away from experimental methodology, so it isn't surprising that the diagnostic criteria for the constitutive relevance of working parts don't appeal to ideal interventions.

Another potential reason for the difference between these criteria is that they might possibly make *contradictory* claims about which working parts are constitutively relevant to FASL performance. I won't speculate whether this is the case here. However, I will say that if mutual manipulability criteria pick out different working parts of the barn owl than competence criteria, then this is a problem for mutual manipulability criteria, not the other way around. After all, we should treat

---

task information in the space between the dorsal and ventral edges of the OT (Knudsen, 1982). Even so, these are both *place codes*, just with a difference in the anatomical properties.

FASLx as *imperfectly authoritative* in the sense that a plausible account of constitutive relevance must provide a plausible rational reconstruction of FASLx. This is necessary to ensure that the philosophical notion of mechanistic explanation stays accountable to the neuroscientific notion of mechanistic explanation. Now, we've shown that competence criteria provide a plausible rational reconstruction of FASLx. Therefore, the onus would be on advocates of the mutual manipulability criteria to show that mutual manipulability criteria provide a plausible rational reconstruction of FASLx—even though they diverge from the plausible rational reconstruction of FASLx provided by competence criteria.

§3. Constitutive Relevance for Organising Relations

At first, FASLx might appear to lend some support for the simple causal-linking position (which I mentioned in the introduction) that organising relations are *constitutively relevant* to behaviour if and only if (and because) they are the causal relations that *link together* working parts that are themselves constitutively relevant to behaviour. After all, FASLx clearly emphasises the causal organisation of the 10 working parts: they are linked together by a sequence of causal relations into a single-stream, feedforward, Markov chain. Spatial, temporal, and anatomical organising relations are mentioned, but they do seem to take on a secondary role in FASLx. In this section, I'll argue that this a defect of FASLx. Instead, we'll see that careful reconsideration of the FASL mechanism indicates that anatomical organisation plays an underestimated role in a mechanistic explanation of FASL behaviour. This indicates that the causal-linking account of constitutive relevance for organising relations is false, and a more inclusive account is needed.

The first indication that FASLs underestimates the role of anatomical organisation is that it leaves open what I'll call the *organizational question*: why does the FASL mechanism re-code the exact same task information so many times in so many types of states—10 times for frontal azimuthal information, 7 times for frequency information, and 4 times for amplitude information?[12] The simple causal organisation of the FASL mechanism makes this organizational question even more perplexing: since every working part is linked together in a single-stream, feedforward, Markov chain, isn't it *redundant* for the FASL mechanism to re-code the exact same task information so many times? An answer to the organisational question must explain what is changing between the constitutively relevant working parts that makes 10 constitutively relevant working parts *non-redundant*. In this section, I'll argue that the hidden source of non-redundancy is the anatomical organisation of the FASL mechanism. I'll argue that a plausible account of constitutive relevance for organising relations must include the relations that constitute this anatomical organisation.

§3.1. Levels of Anatomy

Talk of *levels* is pervasive in behavioural neuroscience (Craver, 2007). Some (not all) talk of levels is talk of what I'll call *levels of anatomy*, where each level is described by an anatomical type but the relation between levels is left implicit. For example, different measurement techniques are said to target different levels in this sense: e.g., fMRI record BOLD activity at the level of brain regions,

---

[12] To emphasize the force of the question, consider a hypothetical alternative. Moiseff & Konishi (1981: 43) found that the anatomy of the barn owl's head translates frontal azimuth into an *interaural time difference* (ITD) via this equation: ITD = 2.16 × frontal azimuth + 1.40. Why couldn't the FASL mechanism use just three layers: one to encode the ITD, a second to directly solve for the frontal azimuth, and a third to generate the appropriate behavioural response?

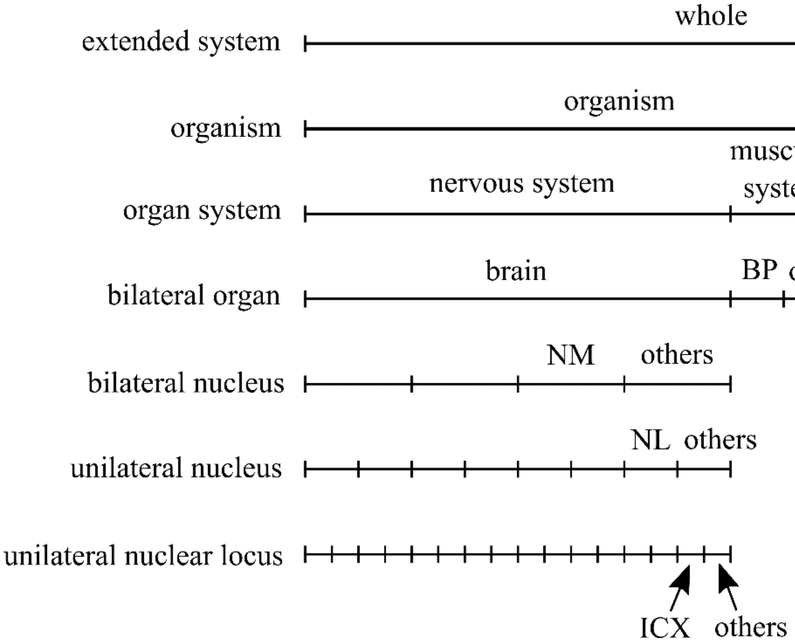

Figure 6. A partition of the anatomical organisation of the barn owl down to 7 levels of anatomy. The highest level is trivial (it partitions the whole system into itself) and the lowest level could be indefinitely refined by further partitions.

electrocorticography records local field potentials at the level of cell clusters, patch clamps record current at the level of ion channels. Likewise, an ongoing debate in behavioural neuroscience is whether levels of anatomy are more relevant to behavioural explanation and if so, which: e.g., the level of neurons (Barlow, 1972), neural assemblies (Buzsáki, 2006), neural circuits (Douglas et al., 1989; Harris & Shephard, 2015; Braganza & Beck, 2018), neural populations (Saxena & Cunningham, 2019; c.f., Barack & Krakauer, 2021), and whole-brain networks (Deco et al., 2015).

Despite its popularity, the notion of levels, including but not limited to levels of anatomy, is fraught (e.g., Craver, 2007; Craver & Bechtel, 2007; Potochnik, 2021). One is that levels-of-anatomy talk rarely (if ever) makes explicit reference to the relations that obtain between levels. For example, behavioural neuroscientists might say that the level of cortical lobes is higher than the level of gyri and sulci, but they leave it implicit what relation *makes* cortical lobes a level higher than gyri and sulci. A second issue is that talk of anatomical levels might seem to suggest that anatomical types form monolithic levels, which is clearly false. For example, the level of gyri and sulci extends across cortex, but not to subcortical regions, which lack gyrus/sulcus formations. A third issue is that individuating parts of an organism by their anatomical level doesn't seem optimal for mechanistic explanations of organism behaviour: wouldn't it be better to individuate parts by their causal relationships with behaviour? Responses to these issues include abandoning anatomical levels (ibid.), treating them as imperfect heuristics (Eronen, 2013; Brooks & Eronen, 2018), or somehow trying to rescue (or reconstruct) them.

Despite these issues, I propose that a rough and incomplete conception of levels of anatomy can be useful for answering our organizational question about the FASL mechanism. Let me explain. Once we've individuated the 10 layers of the FASL mechanism, we can categorize each layer under the anatomical type that it falls under: (a) the input (sound source) and output (saccade) layers are states of the extended barn owl system, (b) Layer 1 (ITD) and Layer 10 layer (movement)

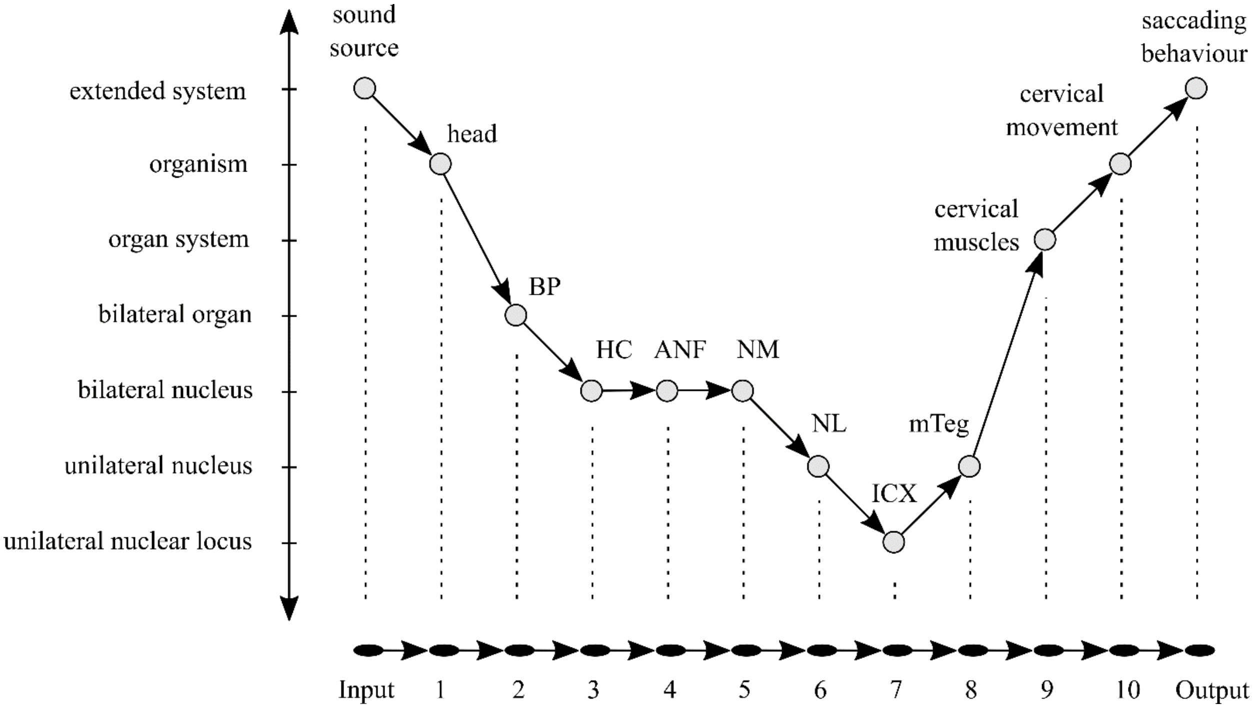

Figure 7. This level plot (named, on the top) takes the single-stream, feedforward, Markov network diagram of the FASL mechanism (numbered, on the bottom) and projects it onto the level axis (left).

are states of the organism, (c) Layer 9 (muscular contraction) is a state of an organ system, (d) Layer 2 (BP) is a state of the bilateral organs, (e) Layers 3–5 (HC, ANF, NM) are states of bilateral nuclei, (f) Layers 6 (NL) and 8 (mTeg) are states of unilateral nuclei, and (g) Layer 7 (ICX) is a state of unilateral nuclear loci. Some of these anatomical types are good candidates for natural kinds: e.g., an organ system (like the muscles in Layer 9) has meaningful anatomical boundaries. Others are bad candidates for natural kinds: e.g., a unilateral nuclear locus (like the ICX in Layer 7) is a cluster of neurons that are anatomically equivalent to their neighbouring neurons but are distinguished from them by their long-range connectivity with other parts of the brain.

In a convenient coincidence—thanks to my cherry-picking the example!—these 8 anatomical types of the 11 working parts that are constitutively relevant to FASL performance happen to form a meaningful sequence. Each entry on the list (except for the final one) can be partitioned into parts of the next entry: we start with (a) the extended system, partition it into (b) an organism and their environment, partition the organism into (c) their organ systems, which can be partitioned into (d) bilateral organs, and so on, into (e) bilateral nuclei, (f) unilateral nuclei, and finally, (g) unilateral nuclear loci (Fig. 6). Since these anatomical types form a meaningful sequence, it is suitable in this case to refer to them as *levels of anatomy*. In fact, some philosophers have argued that a sequence of objects related by parthood forms a series of levels (e.g., Oppenheim & Putnam 1958; Wimsatt, 1976; Kim, 1993).

To be clear, though, I'm not claiming that parthood is the defining relation of these levels of anatomy. After all, each entry on the list is also *larger than* and *spatially-materially includes* the next entry. Likewise, some other philosophers have argued that a descending series of levels is a sequence of objects related by descending scale (e.g., Wimsatt, 1976; Churchland & Sejnowski, 1988; Kim, 1998; Potochnik & McGill, 2012) or spatial/material inclusion (for critical reviews, see Craver, 2007; Potochnik, 2021). My point is neutral to which of these relations (if any) is definitive of levels of anatomy. My point is also neutral to whether levels of anatomy is a useful, coherent notion in general. I'm only claiming that it is fitting to describe the particular sequence of anatomical types that the working parts constitutively relevant to FASL performance fall under as "levels of anatomy". Thus, my claim only turns on the assumption that differences in scale, composition, spatial inclusion, material inclusion, etc. are *imperfectly diagnostic* of differences in anatomical level.

§3.2. Level Plotting

Despite all these caveats, levels of anatomy in the barn owl are useful because they provide a novel diagram that I'll call a *level plot*: we can plot each of the 10 layers of working parts onto each of the 7 levels of anatomy that they occupy (Fig. 7). Level plotting reveals (some of) the systemic anatomical structure of the FASL mechanism, which answers our organizational question: task information is re-coded by 10 layers of working parts *because* 10 layers are needed to form this triangular trajectory across 7 levels of anatomy.[13] What does this triangular trajectory represent? It involves three features that correspond to three difficulties that the barn owl has to overcome during FASL performance: (a) the inflection point at the ICX (Layer 7), (b) the gradual descent from the level of the whole extended system to the level of the unilateral nuclear locus (from the input layer to Layer 7), and (c) the gradual (but faster) ascent from the level of the unilateral nuclear locus back to the level of the whole extended system (from Layer 7 to the output layer). These three features are competencies that are distinct from, but complementary to, Competencies 1–3.

The first feature is the inflection point: task information is encoded at a low level of anatomy—the level of unilateral nuclear loci in the ICX (Layer 7). This feature overcomes the difficulty that the premotor system is an extremely noisy environment: it's a bottleneck that every signal from the sensory systems must pass through to get encoded as a behavioural response. It turns out that *sparse codes* (which encode task information at lower levels of anatomy) are much less vulnerable to noise than distributed codes (which encode task information at higher levels of anatomy) (e.g., Billings et al., 2014; Cayco-Gajic et al., 2017, 2019; Litwin-Kumar et al., 2017). To pass task information through the premotor system, then, the FASL mechanism has to encode the task signal in a sparse code, which is least vulnerable to noise—a process often known as *pattern separation*, *decorrelation*, or *untangling* (DiCarlo & Cox, 2007).[14] This problem is addressed by the ICX. As such, the ICX manifests a *relational* competency (distinct from Competencies 1–3) in virtue of being at the lowest level of anatomy in the FASL mechanism.

The second feature is the gradual descent across 7 levels of anatomy to the inflection point. This feature overcomes the difficulty of moving task information down each level of anatomy. After all, the FASL mechanism receives task information at a very high level (i.e., the angle between the barn owl and the sound source) but it needs to re-encode that information at a relatively low level of anatomy (i.e., firing rates on a single position in a neural map in the ICX) so that the task

---

[13] Likewise, the optic tectum is the inflection point for the more complicated audiovisual version of the FASL mechanism. This could justify the critical role that the optic tectum has played in explanations of most FASL mechanisms.

[14] It's unclear what makes the level of unilateral nuclear locus "low enough" for the FASL mechanism and whether it's possible for other behavioural mechanisms to encode task information at even lower levels of anatomy.

information can survive the noisy environment of the sensorimotor system. Clearly, though, it's difficult to move information across levels of anatomy: information can only be moved down a level or two at a time. Thus, the basic FASL mechanism requires 7 layers of sensory processing to encode task information at the relevant level of anatomy. Once it's encoded in a maximally sparse code by the ICX (Layer 7), it's finally ready to be sent to the noisy environment of the premotor system (Layer 8 to the output layer). As such, Layers 1–7 collectively manifest a second relational competency (distinct from Competencies 1–3) in virtue of forming a descending sequence along levels of anatomy in the FASL mechanism.

The third feature is the gradual ascent across 7 levels of anatomy back up to the system level. This feature overcomes the difficulty of moving task information up each level of anatomy. After all, the FASL mechanism needs to produce an appropriate behavioural response to the task: this involves encoding task information in the behavioural response at the level of the entire extended system. Surprisingly, though, the FASL mechanism requires only 3 motor layers to re-encode task information in progressively more distributed codes—without destructive interference from other signals. This is less than half of the 7 layers required for sensory processing, suggesting it's much easier to move information up levels of anatomy than down them, at least for the FASL mechanism. For the FASL mechanism, there is a simple explanation for this: motor processing has to "undo" some of the changes made by sensory processing (e.g., de-transduction, de-digitization) but not all of them (e.g., converting ITDs into space codes, phase disambiguation).[15] As such, Layers 7–10) collectively manifest a third relational competency (distinct from Competencies 1–3) in virtue of forming an ascending sequence along levels of anatomy in the FASL mechanism.

Of course, my point isn't that all behavioural mechanisms (or sensorimotor mechanisms) exhibit this particular triangular structure or any other anatomical organization. No doubt there is a lot of diversity in the anatomical organization of behavioural mechanisms.[16] Rather, my point is that the organisation that is constitutively relevant to organism behaviour goes well beyond the causal-linking organisation of working parts. Levels of anatomy happen to be a useful way to reveal the complex anatomical organisation of a simple chain of working parts that are constitutively relevant to FASL performance. Obviously, though, levels of anatomy won't be available for all behavioural mechanisms: often the constitutively relevant working parts fall under anatomical types that don't form a meaningful sequence. We'll need to find a different way of analysing the contingent anatomical organization in those behavioural mechanisms. Therefore, future work is also needed to develop a more general form of level plotting that can analyse the non-sequential anatomical organization in more complicated behavioural mechanisms.

---

[15] I suspect that something like this is a general feature of behavioural mechanisms, but further level plotting would be needed to test this claim.

[16] For example, some have argued that higher-order cognition (e.g., planning, decision-making, navigation) involves internalized sensorimotor feedback loops that decouple it from external cues (e.g., Buzsáki, 2013, 2016). If anything like that is true, task information might zig-zag between higher and lower levels of anatomy, creating a sequence of inflection points. Comprehensive mechanistic explanations of such mechanisms would have to identify difficulties raised by higher-order cognitive tasks (but not simple behavioural tasks) that zig-zag trajectories (but not triangular trajectories) are sufficient to overcome.

§3.3. Anatomical Organisation

Overall, these systemic anatomical features implement three basic competencies that are distinct from Competencies 1–3 and highly specific to the barn owl's FASL mechanism:

> **Competency 4:** natural semantic information about states that are *essential* to the FASL task must be encoded in a state of a unilateral nuclear locus causally (and temporally) prior to re-encoding that information in a state of the pre-motor system.
>
> **Competency 5:** natural semantic information about states that are essential and ancillary to the FASL task must be re-encoded in states of anatomical entities at progressively lower levels of anatomy until information encoded at the level of the whole system is encoded at the level of the unilateral nuclear locus prior to encoding in the pre-motor system.
>
> **Competency 6:** natural semantic information about states that are essential to the FASL task must be re-encoded in states of anatomical entities at progressively higher levels of anatomy until information encoded at the level of the unilateral nuclear locus in the pre-motor system is encoded at the level of behaviour.

Competencies 4–6 indicate that the constitutive relevance criteria for the organising relations is dependent on the constitutive relevance criteria for the working parts, but not vice versa. Thus, it is not only possible but also necessary to individuate the constitutively relevant working parts *prior to* individuating the constitutively relevant organising relations. This provides some vindication for the initial focus by New Mechanists on constitutive relevance for working parts.

While I'm optimistic that Competencies 1–3 exhaust the intrinsic competencies of the working parts of the FASL mechanism, I doubt that Competencies 4–6 exhaust all the organising relations of the FASL mechanism that are constitutively relevant to FASL performance. After all, they only concern the flow of natural semantic information, yet there seem to be organisational properties involved in changes in non-informational properties too. For example, it seems like a constitutive organisational feature of the FASL mechanism that it has working parts in motor processing that "undo" changes made to information encoding in sensory processing. In particular, hair cells on the basilar papillae (HCs; Layer 3) transduce the task signal and auditory nerve fibres (ANFs; Layer 4) subsequently discretize the task signal whereas cervical muscles (Layer 10) "de-transduce" and "de-discretize" the task signal. These features seem relevant to a more complete answer to the organisational question: the FASL mechanism has 10 layers partly because there are later layers needed to "undo" certain changes (not all changes) made by earlier layers. For these reasons, I propose that Competencies 4–6 are sufficient but *unnecessary* for constitutive relevance for anatomical organising relations.

While I won't provide necessary criteria for constitutive relevance for anatomical organisation here, I think they would be useful. After all, some anatomical relations clearly aren't constitutively relevant to FASL performance. For example, (a) the nucleus magnocellularis (NM, Layer 5) has a dorsal position relative to the nucleus laminaris (NL; Layer 6), (b) the ipsilateral (i.e., same-side) projections from the NM penetrate the dorsal surface of the NL, and (c) the contralateral (i.e., opposite-side) projections from the NM penetrate the ventral surface of the NL. This is anatomical organisation, but it clearly isn't relevant to FASL performance. After all, (a) the NM could just as well have occupied a lateral position relative to the NL, (b) the ipsilateral projections could just as well have penetrated the ventral surface of the NM, and (c) the contralateral projections could just as well have penetrated the dorsal surface of the NL. These differences wouldn't have necessitated a different causal organisation of the 10 working parts, nor would they have necessitated a different anatomical organisation of the 10 working parts across the 7 levels of anatomy. Therefore, neuroscientists would benefit from a principled basis for excluding such information from their mechanistic explanations of organism behaviour.

I think that New Mechanists are well-positioned to provide a principled basis for identifying and excluding anatomical organising relations that are constitutively irrelevant to organism behaviour. Unfortunately, though, the mechanism-specificity of Competencies 4–6 suggest that it is unlikely that there are *general* diagnostic (or definitive) criteria for the constitutive relevance of organising relations—especially by comparison to the relative mechanism-generality of Competencies 1–3, which suggests that there are *general* diagnostic (and even definitive) criteria for the constitutive relevance of working parts. This mechanism-specificity will surely make it more difficult for New Mechanists to develop necessary conditions for constitutive relevance for organising relations in general philosophical terms. Nonetheless, I think it is worthwhile for New Mechanists to address these difficulties head-on because I think there is a distinct opportunity here for New Mechanists to make a direct positive contribution to mechanistic explanations in neuroscientific practice.

§3.4. Levels of Mechanism

I have tried to press a subtle critique of the New Mechanists in this paper: they have focused too much on the constitutive relevance of working parts and not enough on the constitutive relevance of anatomical organising relations. This critique can be pressed most vigorously on the topic of levels. After all, New Mechanists have complained that the anatomical conception of levels that I appealed to in §3.1 is flawed. For example, Craver (2007: 164–165) observes that explanations in neuroscience "span multiple levels" yet complains that "the levels metaphor is ambiguous". He stresses that "the multilevel structure of neuroscientific explanations" is a central *explanandum* for his account of mechanistic explanation and is "a consequence of the mechanistic structure of neuroscientific explanations" (ibid.). In this subsection, I'll argue that this is false: at least some of the interlevel structure of neuroscientific explanations is just a consequence of the anatomical organisation of the brain.

Craver (2007) starts by introducing a mechanistic conception of levels. On this conception, levels are sequences of working entities ordered by the following rule: one working entity (a ϕ-ing X) is lower-in-level than another working entity (a ψ-ing S) if and only if the ϕ-ing X is *constitutively relevant* to the ψ-ing S (Craver, 2007). Notice that this is an attempt to account for levels in terms of the constitutive relevance of working parts (rather than the constitutive relevance of organising relations). He notes that this mechanistic conception of levels doesn't account for all of popular levels-talk in neuroscience: it is inconsistent with the suggestion that "levels are monolithic strata in nature; that things at different levels interact causally; and that levels, fields, and theories correspond to one another" (165). Instead, he argues that the mechanistic conception of levels provides a *rational reconstruction* that is "consistent with many presumed features of levels… but inconsistent with others…" (2007: 165). This is meant to be a core achievement of his project:

"My emphasis on constitutive mechanistic explanation should not lead one to forget that I am primarily interested in defending a multilevel view of explanation" (2007: 162).

However, the FASL mechanism clearly indicates that the multilevel structure of neuroscientific explanations may be orthogonal to the constitutive relevance of working parts. After all, the 10 layers of the FASL mechanism are all constitutively relevant to FASL performance but not to each other, so they all form a *single* level of FASL mechanism *below* the level of FASL performance. Still, the working entities at this single level of FASL mechanism are distributed across 7 levels of anatomy. I've conceded that the anatomical conception of levels is problematic, but I've shown that it is still helpful (if not indispensable) for revealing the complex anatomical organisation of the FASL mechanism—i.e., the triangular distribution of 10 working parts across the 7 levels of anatomy. Clearly, then, there is a useful, anatomical conception of levels that is fully orthogonal to levels of mechanism and the constitutive relevance of working parts. Future work should try to solve the problems with this anatomical conception of levels, rather than reject it.[17]

Moreover, this anatomical conception of levels better rationalises ordinary levels-talk than the mechanistic conception of levels does. First, it rejects that levels are monolithic strata in the trivial sense that, e.g., the level of the cell only spans across systems composed of cells. Still, it implies that levels are defined in mechanism-independent ways and so can span multiple mechanisms. Thus, it makes sense to compare the cell level in the FASL mechanism with the cell level in an image recognition mechanism. Second, it implies that "things at different levels interact causally" (ibid.). For example, information encoding by the basilar papillae (BP; Layer 2) at the bilateral organ level causes information encoding by the hair cells (HCs; Layer 3) at the bilateral nucleus level. Third, it implies that "levels, fields, and theories correspond to one another" (ibid). For example, theories of organism are appropriate for explaining information encoding by the barn owl's head (Layer 1) at the level of organism whereas theories of neural assemblies are appropriate for explaining information encoding by cells in the external nucleus of the inferior colliculus (ICX; Layer 7) at the level of the unilateral nuclear locus.

Therefore, I propose that popular levels-talk in neuroscience aims to refer to levels of anatomy, not levels of mechanism. If so, then at least some part of the interlevel structure of neuroscientific explanations is *not* "a consequence of the mechanistic structure of neuroscientific explanations" (*contra* Craver, 2007: 164). Once we admit that levels of mechanism are different from the more robust and no less useful anatomical conception of levels, the mechanistic conception of levels may start to seem contrived. Eronen (2013) echoes this sentiment when he claims that "the account of levels of mechanisms is in fact an account of mechanistic composition: it relies entirely on the component-mechanism relation and simply labels whole mechanisms as being at higher 'levels' and their components as being at lower 'levels'" (1047). I'm sympathetic to this: it's misleading

---

[17] Perhaps, Potochnik (2021) is right that nests is a better notion than levels. Then perhaps, the way forward here is to develop some form of "anatomical nest plotting". If we find that anatomical nests can be adequately related by some form of partition, then this would involve forming a tree of anatomical nests and then plotting the nodes of a mechanistic model onto that tree. The result would be difficult to illustrate or visualize, but that would just be a graphics problem, not a conceptual problem.

for mechanists to appropriate levels talk to describe the structure of mechanistic explanation, when levels talk seems more fitting to refer to levels of anatomy.[18]

This point is important because it suggests that New Mechanists have tried to account for too much using an account of constitutive relevance for working parts—and not enough using an account of constitutive relevance for anatomical organisation. We'd do well to expand the New Mechanist toolbox to include a more sophisticated account of constitutive relevance for organising relations. I have stopped short of providing general criteria for such an account, and I've canvassed some reasons to be sceptical that there are any truly general criteria to be identified. However, if New Mechanists are "primarily interested in defending a multilevel view of explanation", as Craver (2007: 162) suggests, then it is imperative that they pay more attention to developing some kind of account for the constitutive relevance of anatomical organisation in behavioural mechanisms. Such an account could prove to be very useful to practicing behavioural neuroscientists. After all, such criteria are responsible for the significant improvements we made to FASLx in this section.

§4. Conclusion

In conclusion, I've argued that an account of constitutive relevance is needed for anatomical organising relations and that it asymmetrically depends on an account of constitutive relevance for working parts, among other things. To do this, I reviewed a complete mechanistic explanation of frontal azimuthal sound localization in the barn owl. Next, I used it to rationally reconstruct a diagnostic account of constitutive relevance for both working parts and anatomical organising relations. I argued that an adequate account of constitutive relevance for anatomical organisation can't be a mere consequence of constitutive relevance for the working parts. In particular, I showed that anatomical organisation of the barn owl's sound localization mechanism involves a triangular trajectory of task information down and back-up an astonishing 7 levels of anatomy, which is independent of the causal-linking organisation of the working parts. Finally, then, I argued that the interlevel structure of *at least one* neuroscientific explanation is a consequence of the constitutive relevance of anatomical organisation, rather than a consequence of the constitutive relevance of working parts (contra Craver, 2007).

I've avoided premature generalization (even in the form of speculation) in this paper because the risk of over-generalization is too serious. Our diagnostic criteria for constitutive relevance for working parts and/or anatomical organisation may be limited to FASL performance, or they could extend to all sensorimotor behaviour, all behaviour on tasks with unique solutions, all organism behaviour. They could be specific to the explanatory standards of neuroethology, or general to the behavioural neurosciences, behavioural sciences, and cognitive sciences. They could be specific to behavioural mechanisms with working parts that fall under a meaningful sequence of anatomical levels, or they could be general to behavioural mechanisms with working parts that fall under a more complicated structure of anatomical types. Future work will be needed to move beyond this case study to (a) develop a more general account of constitutive relevance for anatomical and other

---

[18] Obviously, though, I disagree with Eronen proceeds that "explanations in neuroscience are 'multilevel' only in the sense that they refer to robust properties and generalizations at various stages in the compositional hierarchy and at different (size) scales" (1049) (c.f., Potochnik, 2021). I think the FASL mechanism clearly shows that mechanisms do exhibit anatomical organisation that satisfies a robust notion of level that is both genuinely explanatory and orthogonal to the mechanistic sense of level.

non-causal-linking forms of organisation, (b) develop a cogent conception of anatomical levels, and (c) reevaluate the interlevel structure of other neuroscientific explanations.

Acknowledgements: I thank Sara Aronowitz for her patient discussion and extensive feedback on too many versions of this draft. This paper was exceptionally difficult to write, and I might have given up on it but for her sustained enthusiasm. I thank Carl Craver, Nicholas Shea, and especially Stuart Glennan for their incisive, constructive comments that were pivotal for the paper converging onto its final form. I thank the editors and reviewers at a few journals for their extensive feedback. I thank several people at the neuroscience program at the University of Arizona for workshopping an early draft of this paper in 2020 and for many formative discussions about neuroscience. Finally, I thank the audiences at the 2023 meeting for the International Society for the Philosophy of the Sciences of the Mind (ISPSM), the 2022 meeting for the Philosophy of Science Association (PSA), and the 2021 meeting for the Society for Philosophy and Psychology (SPP) for their helpful feedback.

References

Barack, D. L., & Krakauer, J. W. (2021). Two views on the cognitive brain. *Nature Reviews Neuroscience, 22*(6), 359–71. https://doi.org/10.1038/s41583-021-00448-6

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception, 1*(4), 371–94. https://doi.org/10.1068/p010371

Baumgartner, M. & Casini, L. (2017). An abductive theory of constitution. *Philosophy of Science, 84*(2): 214–33. https://doi.org/10.1086/690716.

Baumgartner, M. & Gebharter, A. (2016). Constitutive relevance, mutual manipulability, and fat-handedness. *British Journal for the Philosophy of Science*, *67*(3), 731–56. https://doi.org/10.1093/bjps/axv003

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton University Press.

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. Lawrence Erlbaum Associates.

Billings, G., Piasini, E., Lorincz, A., Nusser, Z., Silver, R.A., Lőrincz, A., Nusser, Z., & Silver, R.A. (2014). Network structure within the cerebellar input layer enables lossless sparse encoding. *Neuron*, 83: 960–74.

Braganza, O. & Beck, H. (2018). The circuit motif as a conceptual tool for multilevel neuroscience. *Trends in Neuroscience*, *41*(3): 128–36.

Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301069.001.0001

Buzsáki, G. (2013). Time, space and memory. *Nature, 497*, 568–9. https://doi.org/10.1038/497568a

Buzsáki, G. (2016). *The brain from inside out*. Oxford University Press.

Campaner, R. (2006). Mechanisms and counterfactuals: A different glimpse of the (secret?) connexion. *Philosophica*, *77*: 15–44.

Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology & Philosophy*, *27*(1), 49–71. https://doi.org/10.1007/s10539-011-9292-0

Cao, R. (2014). Signaling in the brain: In search of functional units. *Philosophy of Science, 81*(5), 891–901. https://doi.org/10.1086/677688

Carr, C. E., & Konishi, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proceedings of the National Academy of Sciences, 85*(21), 8311–5. https://doi.org/10.1073/pnas.85.21.8311

Carr, C. E., & Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience*, *10*(10), 3227–46. https://doi.org/10.1523/JNEUROSCI.10-10-03227.1990

Carr, C., Ashida, G., Wagner, H., McColgan, T., & Kempter, R. (2016). The role of conduction delay in creating sensitivity to interaural time differences. In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner, & C. Lanting (Eds.), *Physiology, psychoacoustics and cognition in normal and impaired hearing* (pp. 189–196). Springer International Publishing. https://doi.org/10.1007/978-3-319-25474-6_20

Cayco-Gajic, N. A., Clopath, C., & Silver, R.A. (2017). Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks. *Nature Communications, 8*, 1116. https://doi.org/10.1038/s41467-017-01109-y

Cayco-Gajic, N. A., & Silver, R. A. (2019). Re-evaluating circuit mechanisms underlying pattern separation. *Neuron*, *101*(4), 584–602. https://doi.org/10.1016/j.neuron.2019.01.044

Cazettes, F., Fischer, B. J., Beckert, M. V., & Pena, J. L. (2018). Emergence of an adaptive command for orienting behavior in premotor brainstem neurons of barn owls. *Journal of Neuroscience, 38*(33), 7270–9. https://doi.org/10.1523/JNEUROSCI.0947-18.2018

Christianson, G. B., & Peña, J. L. (2007). Preservation of spectrotemporal tuning between the Nucleus Laminaris and the Inferior Colliculus of the barn owl. *Journal of Neurophysiology, 97*(5), 3544–53. https://doi.org/10.1152/jn.01162.2006

Churchland, P. S. & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.

Connor, C.E. & Knierim, J.J. (2017). Integration of objects and space in perception and memory. *Nature Neuroscience, 20*(11): 1493–503. https://doi.org/10.1038/nn.4657

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.

Craver, C. F. (2014). The ontic account of scientific explanation. In M. I. Kaiser, O. R. Scholz, D. Plenge, & A. Hüttemann (eds.), *Explanation in the Special Sciences: The Case of Biology and History* (pp. 27–52). Dordrecht: Springer.

Craver, C. F., & Bechtel, W. (2007). Top-down causation without top-down causes. *Biology & Philosophy*, *22*(4), 547–563. https://doi.org/10.1007/s10539-006-9028-8

Craver, C. F., Glennan, S., & Povich, M. (2021). Constitutive relevance & mutual manipulability revisited. *Synthese*, *199*(3), 8807–28. https://doi.org/10.1007/s11229-021-03183-8

Deco, G., Tononi, G., Boly, M., & Kringelbach, M.L. (2015). Rethinking segregation and integration: contributions of whole-brain modelling. *Nature Reviews Neuroscience*, *16*(7): 430–9. https://doi.org/10.1038/nrn3963

Diana, R.A., Yonelinas, A.P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Science*, *11*, 379–86.

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–341. https://doi.org/10.1016/j.tics.2007.06.010

Douglas, R. J., Martin, K. A. C., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. *Neural Computation*, *1*(4), 480–8. https://doi.org/10.1162/neco.1989.1.4.480

Dretske, F. I. (1981). *Knowledge and the flow of information*. MIT Press.

du Lac, S., & Knudsen, E. I. (1990). Neural maps of head movement vector and speed in the optic tectum of the barn owl. *Journal of Neurophysiology*, *63*(1), 131–46. https://doi.org/10.1152/jn.1990.63.1.131

Eronen, M. I. (2013). No levels, no problems: Downward causation in neuroscience. *Philosophy of Science, 80*: 1042–1052.

Brooks, D. S. & Eronen, M. I. (2018). The significance of levels of organization for scientific research: A heuristic approach. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 68: 34–41.

Fischer, B. J., Christianson, G. B., & Peña, J. L. (2008). Cross-correlation in the auditory coincidence detectors of owls. *Journal of Neuroscience*, *28*(32), 8107–15. https://doi.org/10.1523/JNEUROSCI.1969-08.2008

Gaffan, D. (1998). Idiothetic input into object-place configuration as the contribution to memory of the monkey and human hippocampus: a review. *Experimental Brain Research*, *123*, 201–9.

Glennan, S. (1996). Mechanism and the nature of causation. *Erkenntnis*, *44*, 49–71.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, *69*, 342–53.

Harris, K. D., & Shepherd, G. M. G. (2015). The neocortical circuit: Themes and variations. *Nature Neuroscience*, *18*(2): 170–81.

Jack, J. J. B., Noble, D., & Tsien, R.W. (1975). *Electric current flow in excitable cells*. Oxford: Clarendon Press.

Kaplan, D. M. (2011). Explanation and description in computational neuroscience. *Synthese*, *183*(3), 339–373. https://doi.org/10.1007/s11229-011-9970-0

Kaplan, D. M., & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science*, *78*(4), 601–627. https://doi.org/10.1086/661755

Kaiser, M. I. (2017). The components and boundaries of mechanisms. In S. Glennan & P. Illari (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. Routledge.

Kaiser, M. I., & Krickel, B. (2017). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science*, *68*(3), 745–79. https://doi.org/10.1093/bjps/axv058

Kästner, L. (2017). *Philosophy of cognitive neuroscience, causal explanations, mechanisms and experimental manipulations.* Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110530940

Kim, J. (1993). *Supervenience and mind.* Cambridge: Cambridge University Press.

Kim, J. (1998). *Mind in a physical world.* Cambridge, MA: MIT Press.

Knierim, J.J., Lee, I., & Hargreaves, E.L. (2006). Hippocampal place cells: parallel input streams, subregional processing, and implications for episodic memory. *Hippocampus*, *16*, 755–64.

Knudsen, E. I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *Journal of Neuroscience*, *2*(9), 1177–1194. https://doi.org/10.1523/JNEUROSCI.02-09-01177.1982

Knudsen, E. I. (1983). Subdivisions of the inferior colliculus in the barn owl (Tyto alba). *Journal of Comparative Neurology*, *218*(2), 174–186. https://doi.org/10.1002/cne.902180205

Knudsen, E. I. (1984). Auditory properties of space-tuned units in owl's optic tectum. *Journal of Neurophysiology*, *52*(4), 709–723. https://doi.org/10.1152/jn.1984.52.4.709

Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the owl. *Science*, *200*(4343), 795–7. https://doi.org/10.1126/science.644324

Knudsen, E. I., & Konishi, M. (1979). Mechanisms of sound localization in the barn owl (Tyto alba). *Journal of Comparative Physiology*, *133*(1), 13–21. https://doi.org/10.1007/BF00663106

Knudsen, E. I., Blasdel, G. G., & Konishi, M. (1979). Sound localization by the barn owl (Tyto alba) measured with the search coil technique. *Journal of Comparative Physiology*, *133*(1), 1–11. https://doi.org/10.1007/BF00663105

Konishi, M. (1993). Neuroethology of sound localization in the owl. *Journal of Comparative Physiology A*, *173*(1), 3–7. https://doi.org/10.1007/BF00209613

Köppl, C. (1997a). Frequency tuning and spontaneous activity in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. *Journal of Neurophysiology*, *77*(1), 364–77. https://doi.org/10.1152/jn.1997.77.1.364

Köppl, C. (1997b). Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. *Journal of Neuroscience*, *17*(9), 3312–21. https://doi.org/10.1523/JNEUROSCI.17-09-03312.1997

Krickel, B. (2018). Saving the mutual manipulability account of constitutive relevance. *Studies in History and Philosophy of Science Part A*, *68*, 58–67. https://doi.org/10.1016/j.shpsa.2018.01.003

Kuokkanen, P. T., Wagner, H., Ashida, G., Carr, C. E., & Kempter, R. (2010). On the origin of the extracellular field potential in the nucleus laminaris of the barn owl (Tyto alba). *Journal of Neurophysiology*, *104*(4), 2274–2290. https://doi.org/10.1152/jn.00395.2010

Laughlin, S.B., de Ruyter van Steveninck, R.R., & Anderson, J.C. (1998) The metabolic cost of neural information. *Nature Neuroscience*, *1*: 36–41. https://doi.org/10.1038/236

Leuridan, B. (2012). Three problems for the mutual manipulability account of constitutive relevance in mechanisms. *British Journal of Philosophy of Science*, *63*: 1–29.

Litwin-Kumar, A., Harris, K.D., Axel, R., Sompolinsky, H., & Abbott, L.F. (2017). Optimal degrees of synaptic connectivity. *Neuron*, *93*: 1153–64.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25. https://doi.org/10.1086/392759

Manns, J.R. & Eichenbaum, H. (2006). Evolution of declarative memory. *Hippocampus*, *16*, 795–808.

Masino, T., & Knudsen, E. I. (1990). Horizontal and vertical components of head movement are controlled by distinct neural circuits in the barn owl. *Nature*, *345*(6274), 434–37. https://doi.org/10.1038/345434a0

Masino, T., & Knudsen, E. I. (1992). Anatomical pathways from the optic tectum to the spinal cord subserving orienting movements in the barn owl. *Experimental Brain Research*, *92*(2), 194–208. https://doi.org/10.1007/BF00227965

Masino, T., & Knudsen, E. (1993). Orienting head movements resulting from electrical microstimulation of the brainstem tegmentum in the barn owl. *The Journal of Neuroscience*, *13*(1), 351–70. https://doi.org/10.1523/JNEUROSCI.13-01-00351.1993

Moiseff, A., & Konishi, M. (1981). Neuronal and behavioral sensitivity to binaural time differences in the owl. *Journal of Neuroscience*, *1*(1), 40–8. https://doi.org/10.1523/JNEUROSCI.01-01-00040.1981

Neander, K. (2017). *Mark of the mental: A defence of informational teleosemantics.* MIT Press.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. In H. Feigl, M. Scriven, and G. Maxwell (eds), *Concepts, theories, and the mind-body problem, Minnesota Studies in the Philosophy of Science II*. University of Minnesota Press: Minneapolis, 3–36.

Peña, J. L., & DeBello, W. M. (2010). Auditory processing, plasticity, and learning in the barn owl. *ILAR Journal*, *51*(4), 338–52. https://doi.org/10.1093/ilar.51.4.338

Pena, J. L., & Gutfreund, Y. (2014). New perspectives on the owl's map of auditory space. *Current Opinion in Neurobiology*, *24*, 55–62. https://doi.org/10.1016/j.conb.2013.08.008

Peña, J. L., & Konishi, M. (2000). Cellular mechanisms for resolving phase ambiguity in the owl's inferior colliculus. *Proceedings of the National Academy of Sciences*, *97*(22), 11787–92. https://doi.org/10.1073/pnas.97.22.11787

Peña, J. L., Viete, S., Funabiki, K., Saberi, K., & Konishi, M. (2001). Cochlear and neural delays for coincidence detection in owls. *Journal of Neuroscience*, *21*(23), 9455–9. https://doi.org/10.1523/JNEUROSCI.21-23-09455.2001

Pettigrew, J. D., & Konishi, M. (1976). Effect of monocular deprivation on binocular neurones in the owl's visual Wulst. *Nature*, *264*(5588), 753–4. https://doi.org/10.1038/264753a0

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford University Press.

Piccinini, G. (2020). *Neurocognitive mechanisms: Explaining biological cognition*. Oxford University Press.

Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese*, *183*(3), 283–311. https://doi.org/10.1007/s11229-011-9898-4

Potochnik, A., & McGill, B. (2012). The limitations of hierarchical organization. *Philosophy of Science*, *79*(1), 120–140. https://doi.org/10.1086/663237

Potochnik, A. (2021). Our world isn't organized into levels. In D. Brooks, J. DiFrisco & W. C. Wimsatt (eds.), *Levels of organization in biology*. Cambridge, USA: MIT Press.

Psillos, S. (2004). A glimpse of the secret connexion: Harmonizing mechanisms with counterfactuals. *Perspectives on Science*, *12*: 288–319.

Romero, F. (2015). Why there isn't inter-level causation in mechanisms. *Synthese*, *192*(11): 3731–55. https://doi.org/10.1007/s11229-015-0718-0.

Saberi, K., Takahashi, Y., Farahbod, H., & Konishi, M. (1999). Neural bases of an auditory illusion and its elimination in owls. *Nature Neuroscience*, *2*(7), 656–9. https://doi.org/10.1038/10212

Sarpeshkar, R. (1998) Analog versus digital: Extrapolating from electronics to neurobiology. *Neural Computation, 10*: 1601–1638. https://doi.org/10.1162/089976698300017052

Saxena, S., & Cunningham, J. P. (2019). Towards the neuron population doctrine. *Current Opinions in Neurobiology*, 55: 103–11.

Seidl, A. H., Rubel, E. W., & Harris, D. M. (2010). Mechanisms for adjusting interaural time differences to achieve binaural coincidence detection. *Journal of Neuroscience*, *30*(1), 70–80. https://doi.org/10.1523/JNEUROSCI.3464-09.2010

Sengupta, B., Laughlin, S. B., & Niven, J. E. (2014). Consequences of converting graded to action potentials upon neural information coding and energy efficiency. *PLoS Computational Biology*, *10*(1). https://doi.org/10.1371/journal.pcbi.1003439

Shea, N. (2018). *Representation in cognitive science*. Oxford University Press.

Sullivan, W. E., & Konishi, M. (1984). Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. *Journal of Neuroscience*, *4*(7), 1787–99. https://doi.org/10.1523/JNEUROSCI.04-07-01787.1984

Sullivan, W. E., & Konishi, M. (1986). Neural map of interaural phase difference in the owl's brainstem. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(21), 8400–8404. https://doi.org/10.1073/pnas.83.21.8400

Suzuki, W.A., Miller, E.K. & Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology*, *78*, 1062–81.

Tabery, J. (2004). Synthesizing activities and interactions in the concept of a mechanism. *Philosophy of Science*, *71*: 1–15.

Takahashi, T., & Konishi, M. (1986). Selectivity for interaural time difference in the owl's midbrain. *The Journal of Neuroscience*, *6*(12): 3413–22. https://doi.org/10.1523/JNEUROSCI.06-12-03413.1986

Takahashi, T. T., & Konishi, M. (1988). Projections of the cochlear nuclei and nucleus laminaris to the inferior colliculus of the barn owl. *Journal of Comparative Neurology*, *274*(2), 190–211. https://doi.org/10.1002/cne.902740206

Torres, P. J. (2008). A modified conception of mechanism. *Erkenntnis*, *71*: 233–51.

Wagner, H., Takahashi, T., & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the barn owl's inferior colliculus. *The Journal of Neuroscience*, *7*(10), 3105–16. https://doi.org/10.1523/JNEUROSCI.07-10-03105.1987

Wimsatt, W. (1976). Reductionism, levels of organization, and the mind–body problem. In G. Globus, I. Savodnik, and G. Maxwell (eds.), *Consciousness and the Brain* (199–267). New York: Plenum Press.